Information Collection for Subscription Business

Jiacheng Zhong, Fei Gao

April 12, 2019

KELLEY SCHOOL
OF BUSINESS
INDIANA UNIVERSITY

Subscription Box

Birchbox : a box of four to five selected samples of makeup, or other beauty related products.

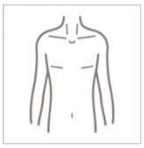
Stitch Fix : Stitch Fix is a personal styling service that sends individually picked clothing and accessories items for a one-time styling fee. Customers fill out a survey online about their style preferences. A stylist at the company picks five items to send to the customer. Stylists pick items based on a customer's survey

VineBox : sends its subscribers three glasses of wines every month. The three tubes of wine are curated by a team of specialists around the world. The company aims to simplify wine by providing tasting notes without the inconvenience of a long wine list

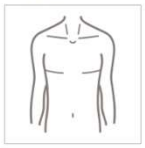
Information Collection

How would you describe your body type?

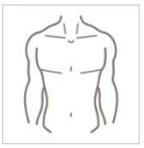
SLIM



AVERAGE



ATHLETIC



HUSKY

Which best describes your shaving habits?

- ☐ I shave my whole face, every day
- ☐ I wear some stubble, so I shave occasionally
- ☐ I sport a mustache or other styled facial hair
- ☐ I have a full beard, but keep it trim and tidy
- ☐ I don't shave at all

What type of wine do you prefer?

We'll take your preferences into account but may throw a curveball or 2 at you. Discovery is part of the experience and tastes change.

- Love Both
- Prefer Red
- Prefer White

Goal

- Is collecting information from customers an effective way to raise profit for retailers?
  - collecting information ⟺ asking customers more questions
- If so, until what extent does the information gathering become ineffective?

vs.

Literature

- Hotelling Line
  - (Neven 1985), (Böckem, S. 1994)
- Probabilistic Sellings
  - (Fay and Xie 2008), (Jerath, Netessine et al. 2010), (Fay and Xie 2015), (Zhang, Joseph et al. 2015)
- Subscription Box
  - (Woo and Ramkumar 2018)

A **Hotelling model** refers to a monopolistic competition model in economics that demonstrates consumer preference for particular brands of goods and their locations.

Hotelling Line

1. "Linear city" is the interval [0,1]
2. Consumers are distributed uniformely along this interval.
3. There are 2 products, located at each extreme who sell the same good. The unique difference among firms is their location.

Probabilistic products

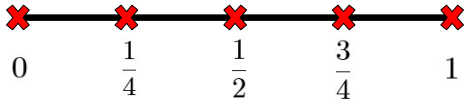
Priceline offers "opaque" hotel rooms in which a buyer specifies dates, city, and approximate quality

…..A probabilistic good is not a concrete product or service but an offer involving a probability of getting any one of a set of multiple distinct items. Under the probabilistic selling strategy, a multi-item seller creates probabilistic goods using the existing distinct products or services and offers such probabilistic goods as additional

purchase choices. The probabilistic selling strategy allows sellers to benefit from introducing a new type of buyer uncertainty, i.e., uncertainty in product assignments.

Base Model

- Customers



$$u_1(x, p_1) = v - tx - p_1$$
$$u_2(x, p_2) = v - t(1 - x) - p_2$$
$$u_o(x, p_o) = \phi(v - tx) + (1 - \phi)\big(v - t(1 - x)\big) - p_o$$

- Retailer

$$\Pi(p_1, p_2, p_o) = q_1(p_1 - c) + q_2(p_2 - c) + q_o(p_o - c)$$

1. Go through notations : v, x, p, ϕ
    1. v : when both players have the same v, the number of customers preferring product 1 is equal to the number of customers preferring product 2 (market symmetry)
    2. t : when t small, neighboring customers share similar preference. When t large, neighboring customers share different preference (customer heterogeneity)
2. Emphasize u_o
    1. ϕ : there is probability ϕ a customer will receive Product 1, and probability $1 - \phi$ to receive product 2
    2. Decision variables for the retailer: price and ϕ
3. Why probabilistic goods improve profit?
    1. Separate the customers who are indifferent between Product 1 and Product 2 from the rest of the customers. In this way, the retailer can set higher prices for those who are not indifferent between Product 1 and Product 2.
4. Sub-game perfect Nash Equilibrium
    1. Sequence of events
    2. Decision variables

Including Information Collection and Market Asymmetry

- Information Collection

start here

Information Collection Effort (# of segments)	Line Partitioning	Available Products
$n = 1$		$1, o, 2$
$n = 2$		$1, o1, o2, 2$
$n = 3$		$1, o1, o2, o3, 2$
$n = 4$		$1, o1, o2, o3, o4, 2$
$n = 5$		$1, o1, o2, o3, o4, o5, 2$

1. Why partition?
    1. For Birchbox, they have to send out over 1 million boxes every month, but they only have around one hundred types of boxes. It means they will a lot of customers share the same type of box. In our setting, customers from the same segment share the same set of products, which are {product 1, product 2, a segment-specific product}
    2. I think this is the most natural way to modeling the process of information collection. In our case, information increases gradually. When n=1, there is no information, and when n=5, retailer has the most precise information of the customers
    3. there may be smart ways to partition the line, such as combinatorial optimization. The goal is not to find optimal partitioning method in this five-customer abstract mode. This kind of smart optimization approaches can be applied to some specific type of problems, but they are definitely not the most general approach. My goal is to find the a general way to model the process of information gathering (the information increases from zero to a lot).

2. Facts

1. The two specific products: Product 1 and 2 are available to all customers
2. Besides, there is a segment specific product that is only available to those on the same segment.

Including Information Collection and Market Asymmetry

- Customers in an Asymmetrical Market

$$u_1(x, p_1) = v - tx - p_1$$
$$u_2(x, p_2) = \alpha v - t(1 - x) - p_2$$
$$u_{oi}(x, p_{oi}) = \phi(v - tx) + (1 - \phi)(\alpha v - t(1 - x)) - p_{oi}$$

- Retailer
  - For each n, choose the optimal product combination(s)
  - Decide an n that gives the highest profit

1. Market Asymmetry
   1. when $\alpha = 1$, the number of customers prefer product 1 is equal to the number of customers prefer product 2
   2. when $\alpha \neq 1$, more customers prefer Product 1 to Product 2

Including Information Col … arket Asymmetry

$$\max_{p_1, p_{oi}, p_2, \phi_{oi}} \Pi(p_1, p_{o1}, p_{o2}, p_{o3}, p_2)$$

$$\text{s.t.} \quad u_1(0, p_1) \geq 0$$
$$u_1(0, p_1) > u_2(0, p_2)$$
$$u_1(0, p_1) > u_{o1}(0, p_{o1})$$
$$u_{o2}\left(\frac{1}{4}, p_{o2}\right) \geq 0$$
$$u_2\left(\frac{1}{4}, p_2\right) < u_{o2}\left(\frac{1}{4}, p_{o2}\right)$$
$$u_1\left(\frac{1}{4}, p_1\right) < u_{o2}\left(\frac{1}{4}, p_{o2}\right)$$
$$u_{o3}\left(\frac{1}{2}, p_{o3}\right) \geq 0$$
$$u_2\left(\frac{1}{2}, p_2\right) < u_{o3}\left(\frac{1}{2}, p_{o3}\right)$$
$$u_{o3}\left(\frac{1}{2}, p_{o3}\right) > u_1\left(\frac{1}{2}, p_1\right)$$
$$u_{o4}\left(\frac{3}{4}, p_{o4}\right) \geq 0$$
$$u_{o4}\left(\frac{3}{4}, p_{o4}\right) > u_1\left(\frac{3}{4}, p_1\right)$$
$$u_2\left(\frac{3}{4}, p_2\right) < u_{o4}\left(\frac{3}{4}, p_{o4}\right)$$
$$u_2(1, p_2) \geq 0$$
$$u_1(1, p_1) < u_2(1, p_2)$$
$$u_2(1, p_2) > u_{o5}(1, p_{o5})$$

Information Collection Effort (# of segments)		…inations
$n = 1$	{1, o, o, o, 2}, {1,	…l, 2, 2, 2, 2} .....
⋮		
$n = 5$	{1, o2, o3, o4, 2},	…2, 2}, {1, o2, 2, 2, 2} .....

INDIANA UNIVERSITY

For each information collection effort n, the retailer will find product combinations that give the highest profit. Finally, the retailer will decide the optimal information collection effort.

Here is an example about the optimization problem the retailer needs to solve given a product combination {1, o2, o3, o4, 2}

- explain the constraints

Equilibria

$\alpha = 1$	$0 < \alpha < 1$		
	$t < 2v - 2\alpha v$	$t = 2v - 2\alpha v$	$t > 2v - 2\alpha v$
			

red : profit is greater than the previous level
grey : profit remains the same as the previous level

- $0 < \alpha < 1$
  - $t < 2v - 2\alpha v$: under market a symmetry, the retailer increases his profit through dividing the customers into two groups: the majority of the customers who prefer product 1 and the minority of the customers who prefer product 2. The retailer can raise the price of product 1 for those who prefer product 1; similarly, raise the price of product 2 for those who prefer product 2. **Basically, information collection helps the retailer group the customers with similar preference together so that he can charge a different price for each group.** But at the very last stage, information collection does not help because of the low customer heterogeneity, which means neighboring customers share similar preference. Therefore, treating similar customers differently isn't really better than treating them in similar ways at the last stage.
  - $t > 2v - 2\alpha v$: customers heterogeneity
  - $t = 2v - 2\alpha v$: there exist a threshold such that information collection from the customers is not profitable beyond the threshold and before the threshold. This short window is interesting because it asks the retailer to put just right amount of effort.
- $\alpha = 1$

- Under market symmetry, information collection isn’t profitable until the retailer puts his maximum effort. The number of customers preferring product 1 equals to the number of customers preferring product 2.

Conclusion and Further Research

- What if there is a cost to collect information from customers?
  - Customers feel annoyed when they are asked to answer more questions.
- Customers continuously distribute along the Hotelling line instead of the 5 discrete locations.

Thank You!