EPSRC Centre for Doctoral Training in Industrially Focused Mathematical Modelling

Short-Heuristics Applied To Purchasing Decisions

Ian Roper

Contents

Heuristics are subconscious simplifications to decision-making problems used to minimise thought power and response time.

1. Introduction

Heuristics are decision-making strategies we use in order to make fast, simple decisions. These are usually subconscious simplifications or educated guesses to come to a fast decision with minimal effort and thus do not completely account for logic or rigorous inference. These heuristic solutions commonly agree with that resulting from a full and logical analysis as if one had consciously thought carefully through the available options, however, sometimes heuristics can cause illogical decisions and large errors.

Three examples of heuristics in retail are:

A. **Scarcity** – when something is scarce, we subconsciously infer its value is increased.

B. The power of **"for free"** – the word "free" excites customers and the free element is subconsciously overvalued.

C. **Social Proof** – we see others buying a product and subconsciously assume they know as much as we do and are making good decisions and so it is reasonable for us to make the same decisions.

Previous models of consumer product choice have been based on rational decisions using a utility function. The utility function is the gain which the customer hopes to get from a certain choice. The choice the customer then makes is that which maximises their utility function. However, this assumes the consumer will spend the effort thinking about a difficult, but sometimes unimportant, decision. We seek to understand how consumers make *irrational* decisions and so this framework of utility functions no longer holds true.

While the use of heuristics in general life situations and business has been qualitatively observed, mainly in psychological and economic studies, little is understood about the role of heuristics in retail. Moreover, there has been no quantitative research towards modelling the use of heuristics by consumers. Despite this, attempts to influence customers based on these heuristics have already been put in place in online retail. Travel companies such as Expedia use 'urgency messaging' to warn people when the number of rooms in a hotel or seats on a flight are becoming low. This is to cue the use of certain heuristics by restricting the factors the consumer considers to urge them to make a quick and easy purchase. This was, however, perfected by trial and error and no quantitative analysis has been done.

Understanding the decision-making processes and sometimes irrational behaviour of customers would be extremely useful to Tesco and other retail companies. Influencing the purchasing decisions made by consumers to help them make better purchasing choices could improve customer satisfaction and ultimately help retain customers, which is a high priority in Tesco's business model. Furthermore, the lack of previous research in this area means any progress we make will potentially spark the growth of a new area of applied mathematical research.

Our aims are:

i. to propose a general mathematical framework to describe how a person or population may use heuristics to make purchase decisions,

ii. to devise techniques to detect these heuristics in purchasing data,

iii. to apply the framework to purchasing data from Tesco to exhibit the use of the framework in practice.

Following this introduction, we first outline the mathematical framework we have developed to describe the use of heuristics by a population of people. We then analyse the purchasing data made available to us by Tesco and show results of the framework being applied to this data. Finally, we conclude the findings of this project and propose further work in this field.

2. Mathematical Framework

The chance of each customer deploying each heuristic is given by a mixed strategy.

Representing Mixed Strategies on Simplexes

We assume that every consumer has a *mixed buying strategy* over a set of possible heuristics, which is the probability of deploying each heuristic to make a purchasing decision. In practice, the full on utility-based logical decision and not buying at all should be included as valid strategies but we will assume for now that the consumers buy based on one of the three heuristics A, B, C, introduced above.

These mixed strategies can be expressed as a point on a surface (or higher dimensional shape when more than three heuristics are available), known as a simplex, in which all the probabilities sum to one. This is a necessary criterion for mixed strategies as a choice must be made. An example of a simplex for the use of three heuristics is shown in figure 1.

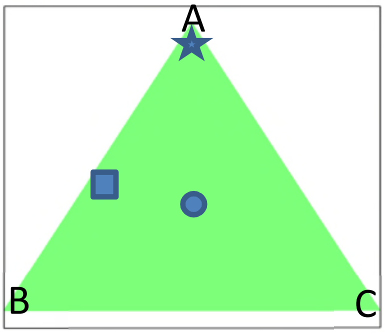

Figure 1: A simplex of mixed strategies with shapes showing example interpretations (star, square and circle) of mixed strategies at certain points.

The vertices of the simplex represent the heuristics A, B and C as in the introduction. The position in the simplex gives the probabilities P_A, P_B, P_C that the customer will use each heuristic. If a customer always used the scarcity heuristic in their choices, their mixed strategy would be at the vertex representing that heuristic for example the star in figure 1. If they never use a particular heuristic, their mixed strategy would be on the opposite side to the vertex representing that heuristic (square) and if they were equally likely to use any of the three; their mixed strategy would lie in the middle of the simplex (circle).

The distribution of a population's mixed strategies can be represented as a density plot on a simplex.

Distributions of Mixed Strategies

We choose to use a Dirichlet distribution [1] to model the distribution of mixed strategies of a population of customers. We use this distribution for mathematical convenience; however, any other distribution over a simplex could be used. The distribution of the mixed strategies of a population can be represented by a density plot over the simplex. The higher the density at a given point, the more likely a customer in the population has the mixed strategy given by that point. The density depends on three "α" parameters, one for each heuristic. Examples of this are shown in figure 2.

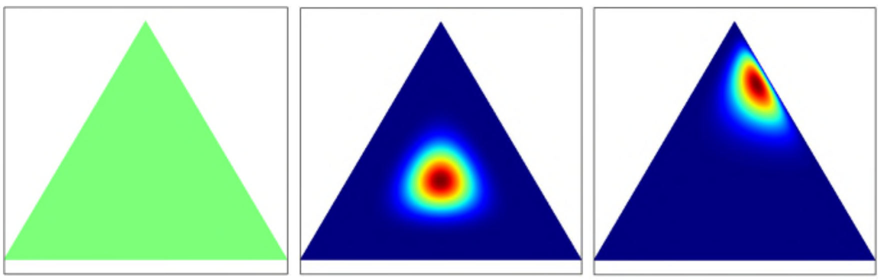

Figure 2: Examples of density plots of Dirichlet distributions; red areas represent high density, blue areas represent low density. Left distribution has all parameters equal to 1 and is flat with all mixed strategies equally likely. Middle distribution has all parameters equal but greater than 1 (10, 10, 10) meaning the centre of mass is in the centre of the simplex and has a higher density (red) than the rest of the simplex (blue). Right distribution has α greater for the heuristic at the top vertex than for the other two heuristics (15,2,5).

Properties of the Dirichlet distribution include:

- The greater the value of α for a certain heuristic, the more dense the distribution is near that vertex and so the higher the probability of a random customer in that population having a large probability of using that heuristic.
- The centre of mass of the distribution (the mean) can be found by normalising the α values by the total of all the α values so that they sum to one.
- The larger the sum of the α values, the more tightly the distribution is concentrated around the centre of mass.

We model changes in the buying habits of consumers by changes in the α values (and therefore the distribution) of mixed strategies for that population of consumers.

Dynamic Distributions

We now model a shift in the distribution of consumers' buying strategies. This is motivated by the possibility of consumers' attitudes changing seasonally or due to external cues, encouraging the use of certain heuristics. The 'urgency messaging' carried out by travel companies is an example of such cues. The dynamics can be modelled as a discrete change between two time periods or as a continuous change over time.

The cue encourages the consumers to use that heuristic more thus the probability of using that heuristic in their mixed strategies has increased. The distribution of mixed strategies has therefore shifted towards the vertex of that cued heuristic and we can model that by the α value for that heuristic increasing, which will move the centre of mass of the distribution towards that heuristic. An example of this is shown in figure 3.

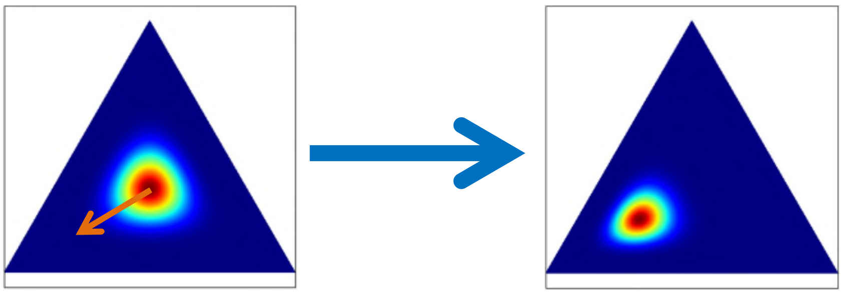

Figure 3: A distribution for a population in which most consumers have an equal chance of choosing each of the three heuristics (α values = (10,10,10)) changing to a distribution with the centre of mass shifted towards heuristic B by an external cue (α values = (10,25,10)).

Mixtures of Distributions

We expect there to be groups of consumers with similar buying strategies and with one strategy as their 'default' which they use a lot. To model this, we combine the distributions of populations that have a heuristic that most people usually use by simply adding the distributions together with certain weightings, known as mixing probabilities. These weightings are based on the proportion of consumers in this new population coming from each smaller population.

The dynamics of these mixtures could either come from the strategies of each group changing or from people switching from one group to another. We believe the latter is more probable in reality and this distribution shift is shown in figure 4.

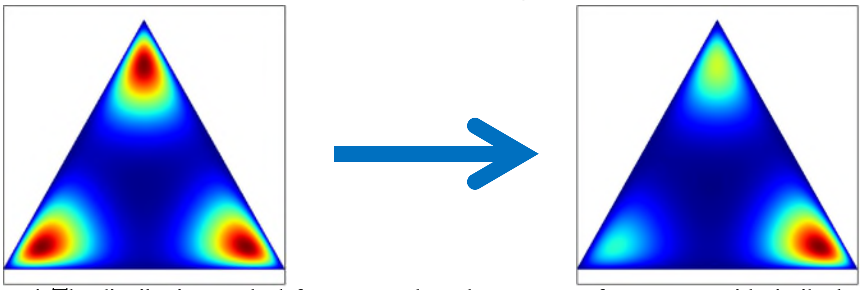

Figure 4: The distribution on the left corresponds to three groups of consumers with similar buying tactics and who usually use the same heuristic. The shift represents consumers changing to the group which usually use the bottom right heuristic due to a cue. Here the mixing probabilities are shifted from (0.33, 0.33, 0.33) to (0.6, 0.3, 0.1).

Scarcity

An example of a 2-dimensional Dirichlet distribution is the distribution of the probability of flipping heads given a coin is randomly sampled from a bag. Most will be approximately 0.5 but there will be slight variations due to manufacturing faults.

We now consider the problem "Does scarcity affect sales rates?" and reformulate the problem as one in which there are two choices for consumers: 'buy' and 'do not buy' the product.

The distribution of the mixed strategies can be described by a 2-dimensional Dirichlet distribution. In this case, the distribution only depends on two α values and it is the probability, p, that a customer will buy the product that is distributed. We expect the distribution of buying to change with the number of units of product on the shelf as more or fewer customers use the scarcity heuristic. We show two possible density functions in figure 5.

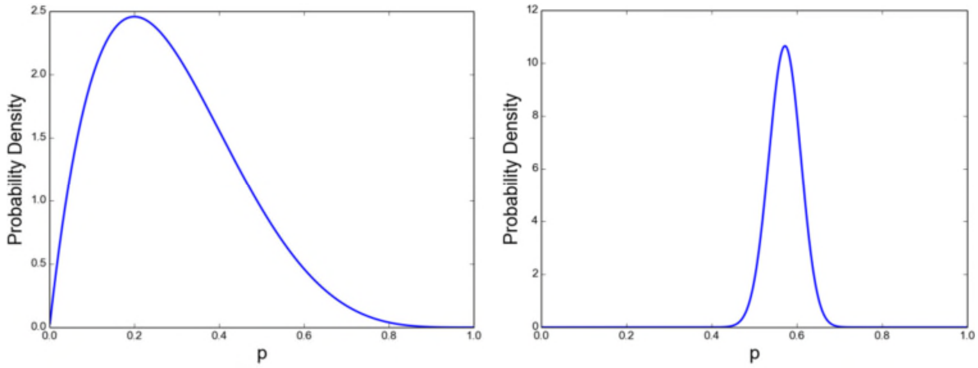

Figure 5: Graphs showing the probability density of a buying strategy against the probability of buying the product, p, for two sets of αs: (2,5) & (100,75).

3. Data

Tesco provided purchasing data and basket data for fruit and bread sections for August 2015 and March 2016. The data included all the items bought in each basket, along with the time of transaction. This included the quantity of the item bought and a description of the product. We calculate the sales rate (number bought per time period) by grouping the purchases and baskets into time periods.

We want to infer if sales rates change due to scarcity and nothing else. Therefore, we normalise the sales rates of a product by the rate of baskets passing through the tills which contain any products from the section (fruit or bread); we call these sector baskets. This is to account for changing footfall of customers affecting sales rate but also to account for differences in buying habits of customers over the day (for example, we observe that less fruit is bought in the evening than in the morning).

The sales rates for 'Tesco Loose Bananas' and 'Hovis Soft White Medium Bread 800G' when normalised by sector baskets are shown in figure 6. These products were chosen as they are the most popular in their respective sections and so should provide as many data points as possible.

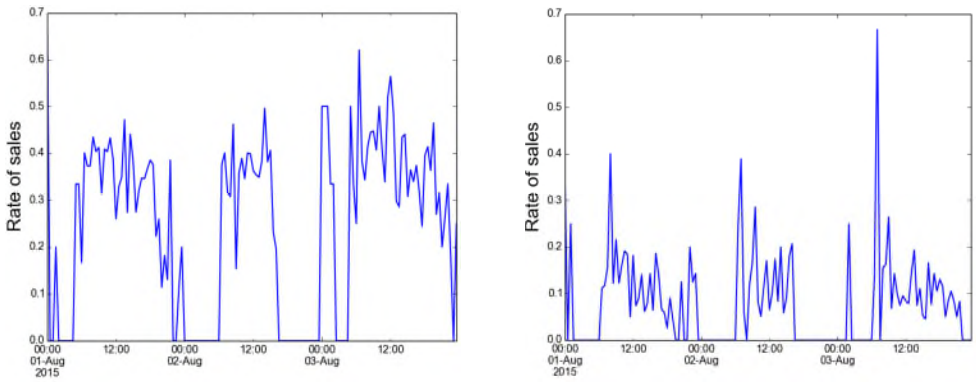

Figure 6: Sales rates normalised by sector baskets for 'Tesco Loose Bananas' (left) and 'Hovis Soft White Medium Bread 800G' (right). Data taken from 1st-3rd August 2015.

Based on information from Tesco, we assume that the shelves are only stocked at the beginning of the day and thus, as consumers buy the product, the shelf stock will decrease. We look for trends throughout the day which will correspond to trends with stock levels if this assumption holds. The data is very variable which causes a big problem when attempting to find trends in the sales rates. This can be helped by using larger time periods; however, the data may get smoothed out too much so the trends still cannot be detected.

We use examples with clear trends to demonstrate the application of the mathematical framework. Other examples do not have clear trends, if any at all.

4. Results

We now apply the framework to the purchasing data and investigate the changes in distribution of buying strategies. Example cases are used here to illustrate the uses of the mathematical framework.

Discrete Change

We separate the data for 'Tesco Red Seedless Grapes 500G' into "morning" and "evening" and fit the αs in the Dirichlet model. The sales data are shown in figure 7 (left) and the probability density versus probability, p, in figure 7 (right). The αs of the Dirichlet distributions for morning and evening are found using the maximum likelihood estimation algorithm to find the distribution that is most likely to produce the data that we see.

We can see that the peak for the distribution for the morning is to the right of the peak of the distribution for the evening. This means the most customers in the morning have a higher chance of buying the product than in the evening. This can be seen in the purchasing data as there is a sharp change in sales rate at around 14:30, indicating something could be happpening at this time to influence sales rates. The higher variance (larger spread) in the morning distribution reflects the higher variability in the sales rate in the morning compared with that in the evening. The fitted distributions capture the purchasing habits of customers at different times.

The sharp change in sales for 'Tesco Red Seedless Grapes 500G' is seen at this time every Friday, suggesting something happens in store at this time; perhaps a restock.

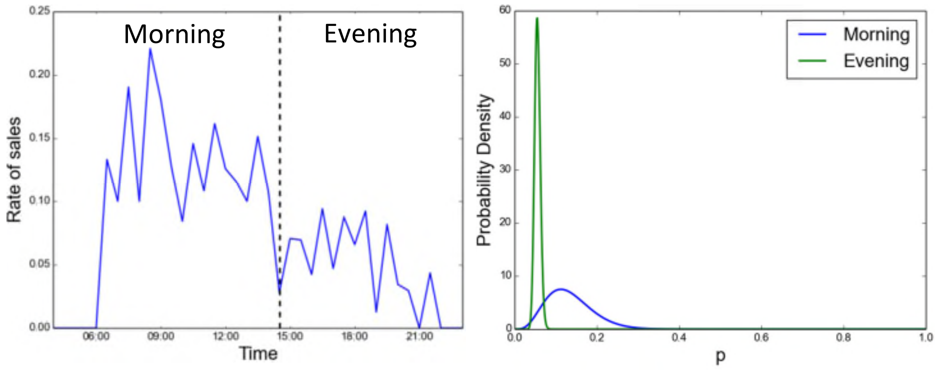

Figure 7: Left: sales rate normalised by section baskets. Right: Discrete shift in Dirichlet distributions for the probability a customer will buy 'Tesco Red Seedless Grapes 500G' between morning (09:00-14:00) and evening (15:00-20:00). Data from Friday 4th March 2016. The time period for grouping the sales was 30 minutes.

Continuous Change

We now model a continuously changing distribution of buying strategies. For this we use an example case of 'Hovis Soft White Medium Bread 800G' and again normalise by sector baskets. The distribution is fitted for every 10 minute time period using the α value for 'buy' (being the number of baskets that underwent transaction containing the product) and the α value for 'do not buy' (being that for the baskets that do not contain the product). We plot the centre of mass for each time period and show the results in figure 8. There is a quadratic trend over the day that can be seen more clearly if the noise is removed. This trend could be due to many reasons other than scarcity, for example different type of bread being popular at different times. However, the distribution fitting and trend shows how we could model changes in consumer behaviour due to heuristics if we could be sure that the cause of the change was indeed heuristic cues.

Smoothing the calculated centres of mass allows the trend to be seen more easily.

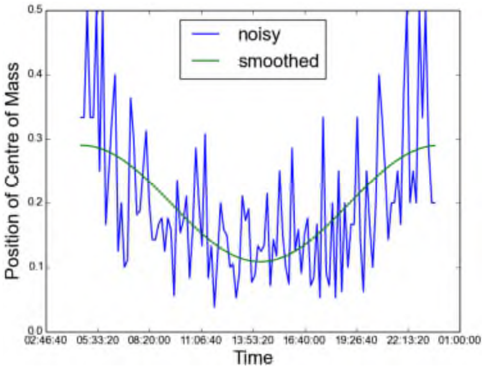

Figure 8: Centre of mass for every 10 minute time interval and a smoothed trend by removing the noise. Data from 1st August 2015 4:00-midnight.

5. Conclusions and Future Work

We have presented a general and versatile framework for modelling consumer choice which we believe can be applied to study the use of heuristics by consumers to make purchasing decisions. We also attempted to detect when consumers use these heuristics in purchasing data made available to us from Tesco. We fitted distributions of the mixed strategies of the consumers at times during the day to observe changes in the distribution caused by changes in stock levels on the shelves.

While the mathematical framework is relatively sound and can be applied to other choice-based applications, detecting when products are purchased based on heuristics, and which heuristics, is very difficult. Because of this, it is difficult to train the model on the data to find the most likely distribution describing the consumers' buying strategies.

Future work on detecting heuristics and the effect of scarcity on sales include:

- Acquiring more data on the stock levels on the shelf either by monitoring this in stores, or carrying out experiments, or obtaining customer-specific data to see how individual strategies change.
- Applying the techniques seen here to other heuristics including the power of "free" and social proof; again, possibly using experiments to detect sales changes.
- Detecting multiple heuristics at once and investigating when consumers use which ones.
- Using these distributions to assist forecasting sales in the future by taking heuristic-based purchases into account.

Potential Impact

Developing a framework for mathematically analysing the use of heuristics is a new area of research so this first attempt can be used to build more robust models on. Using this work, Tesco can improve their understanding of customer buying habits and improve customer satisfaction by encouraging customers to make better purchasing choices. Other companies, such as Expedia, can also benefit from this research by quantitatively understanding the heuristic cues they already use online.

Trevor Sidery, the industrial supervisor on this project, commented on the impact of this work on the digital product department at Tesco:

"Ian has put in place a framework to model heuristics and how to understand what drives customers to make the choices that they do. This was a high risk project as this is a field that has very little prior work, and we now have an excellent base from which future work can rely."

References

1. A Bela, AK Frigyik, and MR Gupta. Introduction to the Dirichlet distribution and related processes. *Department of Electrical Engineering, University of Washington,* 2010.