STATISTICAL MODELING OF PRODUCT PURCHASES AND NEW MEASURES FOR BRAND LOYALTY

TAN SWEE LEONG
(B.SC. (HONS), NATIONAL UNIVERSITY OF SINGAPORE)

A THESIS SUBMITTED
FOR THE DEGREE OF MASTER OF SCIENCE
DEPARTMENT OF STATISTICS AND APPLIED PROBABILITY
NATIONAL UNIVERSITY OF SINGAPORE
2006

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my sincere gratitude to my supervisor, Dr. Chan Yiu Man for all his guidance and invaluable suggestions towards the success of this thesis.

Special thanks go to my wife, Lily and my two daughters Isabelle and Sheredine for their constant and endless support, encouragement and understanding.

I would also like to thank all the staff in my Department for all their help throughout my study and research in National University of Singapore. Last but not least, I thank all my friends who have always been encouraging me throughout the completion of this thesis.

Tan Swee Leong
August 2006

Table of Contents

SUMMARY

In the marketing world, it is always of interest for a firm to study the purchase patterns of its consumers and that of its competitors. I will study the different methods of how we model consumers' purchase patterns for a wide range of products from commonly used products like the purchase of a box of tissue paper to more durable products such as in the purchase of a house or car and finally to those situations when there is only one purchase in the lifetime of an individual.

Furthermore, each firm is always constantly trying to study the brand loyalty of their product, to understand how best to increase their market share. Here, I will also be introducing current as well as new ways of measuring brand loyalty.

Finally, I will apply these theories to the choices made by our own local students in choosing their preferred choice of Post-Secondary Education.

LIST OF TABLES

CHAPTER 1 INTRODUCTION WITH LITERATURE REVIEW ON THE BRAND LOYALTY MEASURES

1.1 Introduction

In panel surveys, the decision maker is the focus of statistical analysis. The statistical population is based on the decision maker. Each decision maker has one set of brand choice probabilities that do not change with time. The brand selected may change from one choice to the next. If there is low consistency in the choices made by a decision maker, he/she will often switch brands over successive choices. Simple terminologies and their definitions will be provided in Section 1.2 will be devoted towards having a better understanding of these terms in the marketing world while Section 1.3 provides a comprehensive Literature Review about the different ways of measuring brand loyalty that are currently found in the marketing world.

Chapter 2 will be devoted to the introduction of some useful methods to model the purchase probabilities in the case of a two-brand problem while Chapter 3 looks their counterparts for a more than two-brand problem.

In Chapter 2, I will be looking at the probability of making a purchase among competing brands that can be modelled using Dirichlet Distribution in the case when there are more than two brands in the same product category. A special case of the Dirichlet Distribution is that of Beta Distribution which caters to the two-brand problem. The characteristics of these two distributions will be discussed in detail so that greater understanding can be made when modeling the probabilities.

In Chapter 3, we look into how Logistic Regression Modeling can be used to make predictions on the probability of choosing one brand over the other using a set of explanatory variables which are thought to be of importance to a consumer. This helps to deal with a very important limitation of Dirichlet

Modeling which does not look at the influence of covariates in predicting repeated purchases.

In Chapter 4, I will be using the Multinomial Dirichlet Distribution to model the purchases for a particular product category where there are more than two brands. The special case for the Multinomial Dirichlet Distribution for the two-brand problem is that of Beta Binomial Distribution. The characteristics of these two distributions will also be discussed in some detail in this thesis. The brand loyalty measures discussed in Chapter 2 will be extended to measure brand loyalty in situations where there are more than two brands.

However, Dirichlet modeling will not be as useful if the number of repeated purchases is rare, in some cases, possibly once or twice in the lifetime of a consumer. In such a situation, I have suggested two methods of dealing with the problem. The first method is to consider the entire household as the same decision maker and the second method is to consider the use of Logistic Regression.

In Chapter 5, we look at how we can estimate the probabilities for choosing the brands if we do not assume independence between consecutive purchases. This is done by first considering the case for two brands, which was then extended to situations where there are more than two brands.

In Chapter 6, I will be looking at other non-parametric approaches of measuring brand loyalty which do not assume any distribution models for the probabilities of purchasing various brands. Instead, I will study the purchase patterns of the consumers buying in a particular product category. I will be providing new methods of measuring brand loyalty using both the behavioural and psychological aspect of a consumer.

I then conclude the thesis with a case study on how we can apply both types of approaches (namely the parametric as well as non-parametric) to a real life problem so as to see whether there are any differences in these approaches in determining the brand loyalty. This will be done in Chapter 7.

I will begin my discussion with some definitions that will be helpful in understanding of my thesis.

1.2 Simple Terminologies and Their Definitions

We will now look at some of these marketing terms which are useful for us throughout the thesis.

Panel

A panel is a field work method in which data is collected from the same sample of decision makers on several occasions, usually at regular intervals. At each wave of the panel, it is normal for data to be collected using the same questions and survey instruments. Panels provide longitudinal data. Cross sectional surveys can also provide retrospective longitudinal data but their validity and reliability are limited by the knowledge and recall of the decision makers. The panel length refers to the number of purchases being taken into the study. Another related term is that of panel size, which refers to the number of individuals included in the study.

Product Category

A product category is a set of product or service brands that provides similar benefits to the consumer. For example, when we are topping up the petrol in our car at a petrol station, we can choose among the stations that offer “Shell”, “Mobil” or “SPC” petrol. These names are some of the brands under a wider product category of petrol used in Singapore. In the case of choosing a local telephone network, we can choose between “SingTel”, “M1” or “StarHub”. “IDD1521” and “IDD1516” are also telephone networks but cater only to overseas calls. Thus, they belong to a different product category as the local telephone networks.

Decision Maker

In this thesis, the decision maker is used to refer to members in the same household. For example, the choice in question is that of which local universities to go to. Then, there will only be at most one or two choices been made if we were to consider each member of the family as a decision maker. However, if we were to consider the entire household as a decision maker, there will definitely be more choices being made.

Heterogeneity in brand choice

When individuals choose the brand for a product prior to a purchase, they choose it according to their differences in features between the brands (e.g. flavour, speed, price, availability, etc.). Heterogeneity in brand choice is the extent of the variation in the choices of brands. This variation includes differences between successive choices by one decision maker and differences among decision makers.

Choice Models

Choice Models are a class of models where the dependent construct is choice between discrete alternatives. In this thesis, I will be discussing two types of models – Dirichlet model and the Multinomial logit model. The similarity of the Dirichlet and logistic regression models is that, in each model the individual customer has a specific set of choice probabilities. In the Dirichlet, these probabilities are generated by a beta or Dirichlet distribution. In the regression models, they are generated from covariates by regression relationships. Accordingly, the regression models are more flexible, and less restrictive.

Brand Choice Probability

Every decision maker has a probability of selecting each brand. These are the brand choice probabilities. In this thesis, the choice models under consideration are conditional on one brand being selected. Thus, the sum of the set of brand choice probabilities, across all brands, must be one.

Market Share

The percentage of the total market for a product/service category that has been captured by a particular product/service or by a company that offers multiple products/services in that category. In the latter case, the company may choose to look at share on both an individual product/service basis and on a company-wide basis.

Share can be calculated either on a unit basis (i.e., If a company sells 1 million units of mobile phones in a total market of 10 million units, it has a 10 percent share) or on a revenue basis (i.e., If a company sells $1 million worth of mobile phones in a $10 million market, it has a 10% share). Obviously, if a company is able to command a higher price for its product/service than its competitors, it would show a higher market share when calculated on a revenue basis than on a unit basis.

In my thesis, I will be looking at the market share in terms of the units sold by the company rather than in terms of the revenue generated.

Brand Loyalty

Many people had tried giving different definitions of Brand Loyalty. However, by far the most comprehensive one was given by Jacoby and Chestnut (1978). It was defined as the **biased, behavioural response, expressed over time, by some decision-making units** with respect to one or more **alternative brands** out of a set of such brands **and is a function of psychological (decision making evaluative) processes.**

By a **biased response**, it means that there has to be a systematic tendency to buy a certain brand or group of brands. Each brand cannot be chosen independently of the consumer's past purchase decisions. Also, Brand loyalty involves **actual purchases** of a brand. Verbal statements of preference towards a brand are therefore not sufficient to ensure brand loyalty.

For individuals to be loyal to a brand, some **consistency** is needed during a certain **time span**. This suggests that one should not only consider the number of times a specific brand is purchased during a period of time but also the purchase pattern over successive purchase occasions.

A **decision-making unit** may either be an individual, a household or a firm. The decision unit does not have to be the actual purchaser. For example, the purchases of a household are often made by one of the parents, but other members of the household may also be involved in the decision process.

The fifth condition is that one or more brands are selected from a set of brands. This condition implies that consumers may actually be loyal to **more than one brand**, a phenomenon observed by many researchers (e.g. Ehrenberg (1972) and Jacoby (1971). If more than one brand is acceptable, an individual might be indifferent between them and exhibit loyalty to a group of brands rather than to a single brand. A problem with multi-brand loyalty is that it is hard to distinguish this kind of behaviour from brand switching, especially if there are only a few brands available. In this thesis, I will not be considering multi-brand loyalty.

Brand loyalty is a **function of psychological** (decision-making, evaluative) **processes**. Brands are chosen according to internal criteria resulting in a commitment towards the brand. Although consumers do not always seek information actively, they do receive some information, e.g. due to advertising campaigns, which may be used to form certain beliefs about brands. Based on these prior beliefs, brands are evaluated and some are preferred over others. In time, the consumer may develop a commitment towards a brand and become brand loyal. Hence, brand loyalty implies consistent repurchase of a brand, resulting from a positive affection of the consumer towards that brand.

Definition of Consistency of Brand choice or Purchase Consistency

Consistency of Brand choice is a much simpler concept as compared to brand loyalty. It looks at the purchase patterns of the brand. If there is consistency of a particular brand, it means that there is a certain purchase pattern over successive purchase occasions. In a simple two-brand problem, if the two brand choices have high consistency, Individuals will either choose Brand A on repeated occasions or will choose Brand B on repeated occasions. On the other hand, if the two brand choices have low consistency, then there will be small variation in the purchase probabilities among individuals. It is to be noted that It is possible for a brand to have a large market share (say 75%) but yet experiences low consistency of brand choice. In such a situation, there are few individuals who will choose the same brand on all 10 occasions. Most of the individuals will choose the same brand on about 75% of the times. Likewise, it is possible for a brand to have a low market share (say 10%) but yet have high consistency of the brand. Most of the individuals who purchase the brand tend to consistently purchase the same brand. Some measures of brand loyalty involve simply the measure of the consistency of the brand choice. An example of such a measure of brand loyalty is to measure it using the proportion of individuals who purchase a particular brand consistently for the last couple of times.

1.3 Current Approaches to the Measuring of Brand Loyalty

Many people had tried finding different ways of measuring brand loyalty. These measures can however be classified into four groups, based on the following two dimensions:

(1) attitudinal versus behavioural measures

(2) brand-oriented versus individual-oriented measures

1.3.1 Attitudinal versus Behavioural measures

Behavioural measures define brand loyalty in terms of the actual purchases being made over a certain period of time. These types of measures focus on brand loyalty being a biased behavioural response, expressed over time. Their advantages are that they are: (1) based on actual purchases, which are what the firm is interested in as it is related to the performance of the firm; (2) not likely to be by chance as it is based on purchases over a period of time; and (3) easy to compute as the data to be collected comes from actual realization of purchase.

However, one key limitation of behavioural measures is that they do not differentiate between brand loyalty and repeat buying, and therefore may contain false impression of what the true loyalty of the brand is. Also, behavioural measures can easily be affected if there is shortage of the stock during a period of time. The brand loyalty had not changed but the behavioural measures seem to suggest otherwise. Finally, no information is collected on the underlying reason for a particular behaviour and thus it is hard to select the right decision unit.

On the other hand, attitudinal measures are able to differentiate between brand loyalty and repeat buying. They are based on purchase intentions of the consumers and on preferences. As brand loyalty involves a decision being made of one brand with respect to one or more alternative brands, attitudinal measures will be able to take that into consideration. Also, it takes into consideration the cognitive elements of brand loyalty. If attitudinal measures are used, it might be easier to choose the right decision unit. As they are in most instances based on surveys, it may be possible to get data from the decision maker rather than the purchaser (who need not be the decision maker and may represent a group of decision makers) by asking questions to the right individual. Attitudinal measures are also not influenced by short term fluctuations in stock supply as they measure the intrinsic value that the decision maker places on the brand.

However, attitudinal measures may not be an accurate representation of reality as they are not based on actual purchases. An individual may not have a favourable attitude towards a particular brand of car but still purchases the brand as it is the only brand within his/her budget. Finally, attitudinal measures are usually collected at a particular instant of time and it does not reflect possible changes due to changes in income level and changes in preferences over time.

1.3.2 Brand-oriented versus individual-oriented measures

Brand loyalty is the result of a consumer's mental impression of the brand's features. Thus, Brand loyalty can be seen as a property of the brand's features (Aaker (1991); Rossiter and Perrcy (1987)) or may be considered more as a characteristic of the respective consumers who process the information (Hafstrom, Chae and Choung (1992); Sproles and Kendall (1986)). If brand-oriented measures are used, a value of brand loyalty is obtained for each brand. Difference between the brand loyalty of each individual is not as important since the value of the brand loyalty is an aggregated one. These types of measures are less suited to study the influence of individual's characteristics on brand loyalty. On the other hand, if an individual-oriented measure is used, the loyalty of specific customers is estimated, and it is of less importance to what specific brand that individual is loyal. These types of measures are less suited to make comparison between brands.

Crossing the above mentioned dimensions, four categories can be defined:

A: Brand-oriented attitude measures (e.g. the percentage of consumers who want to purchase brand A).

B: Individual-oriented attitude measures (e.g. the level of agreement or disagreement with the statement "I like to be loyal to the most well-known brands "; see Jacoby, 1971; Raju, 1980 as cited in Sergio Brasini, Marzia Freo, Giorgio Tassinari, 2003).

C: Brand-oriented behavioural measures (e.g. the percentage of buyers that, having already purchased brand A, repurchase it; see Guadagni and Little, 1983; Colombo and Morrison, 1989; Krishnamurthy *et al*., 1992 as cited in Sergio Brasini, Marzia Freo, Giorgio Tassinari, 2003).

D: Individual-oriented behavioural measures (e.g. a consumer is brand-loyal if he/she buys brand A belonging to a specific product category in more than half of the purchasing episodes; see Cunnigham, 1956 as cited in Sergio Brasini, Marzia Freo, Giorgio Tassinari, 2003).

As there are many different measures of brand loyalty done by different experts, I will only be presenting a few more recent and interesting ways of measuring brand that can be found in Literature.

1. Brand Loyalty Measure using a mixture of B and D by Simon Knox, David Walker, 2001

In this study, 191 individuals were recruited onto a panel and they were to record their purchases in at least two of the three product categories in order to produce an effective sample size of 463. Information about current brand usage, stated preferred brands and background information on shopping behaviour and demographic information was elicited from a self-completion questionnaire, which was administered at the beginning of the panel recording period. Fourteen items for assessing involvement and two further items specifically about brand commitment were also included in the questionnaire. As cited in Simon Knox, David Walker, 2001, one of the two commitment scales followed Traylor (1981) and was a simple five-point scale, while the other was a modification of the scale used by Cunningham (1967), which was used because it expresses the psychological construct of commitment set in a behavioural context.

Simon Knox, David Walker developed a measure of brand buying behaviour that reflected the degree to which purchasing within a product category was devoted to a limited set of brands from the greater number that were available in the market place. Such an index was derived ‘using data on respondents’’ purchasing throughout the full 16 week recording period, which is expressed mathematically as

Brand support index =

$$\sum\nolimits_{\text{brands in a set}} \left\{ \frac{(\text{purchase of brand } n)^2}{(\text{total purchase of products})^2} \right\} \log(\text{total purchases})$$

The main part of the equation is derived from the classical Hirschman-Herfindahl index (Hirshman, 1987). The log total purchases multiplier was introduced in order to comply with the requirement for a non-random response. This also has the effect of reducing the weight of the index for respondents who only made a small number of purchases in the category over the 16 week recording period.

If brand commitment and support for each respondent is plotted in matrix format, groupings of cases are being done using a simple *K*-means clustering procedure. Four clusters were specified in order to identify the characteristics and the number of respondents in each is outlined below:

(1) Cluster 1, high commitment/high support, named “loyals”

(2) Cluster 2, low commitment/high support, named “habituals”

(3) Cluster 3, high commitment/low support, named “variety seekers”

(4) Cluster 4, low commitment/low support, named “switchers”.

2. Brand Loyalty Measure using D (Scaled Probability Of Purchase-SPOP) by Terry Elrod, 1988

Suppose for the moment that we have a good estimate of a household's purchase probabilities for a set of J brands of products. The household's loyalty towards the i^{th} brand in the set is defined as:

$$L_i = J\left(P_i - \frac{1}{J}\right) = JP_i - 1$$

In the construction of SPOP measure, one begins with subtracting $\frac{1}{J}$ from all purchase probabilities to give a meaningful origin to the measure. A household that buys a brand with probability exceeding $\frac{1}{J}$ is buying the brand more often than the average brand and therefore shows some degree of loyalty to the brand. This household receives a positive SPOP score. On the other hand, a household buying a brand with probability less than $\frac{1}{J}$ is disloyal and receives a negative SPOP score for that brand.

The maximum value for the household's loyalty towards the i^{th} brand is $J - 1$ while the minimum value is $- 1$. Thus, maximum attainable loyalty increases with the number of brands in the analysis. This is conceptually pleasing since always buying brand out of a larger set of competing brands is a stronger (and rarer) indication of loyalty to the brand. Thus, it should be reflected in a larger maximum attainable score for this brand loyalty measure.

The SPOP measure of brand loyalty assumes that we do have a good estimate of a household's probability of buying each brand. One such estimate is to use a Bayes estimate of a household's purchase probability for each brand. A Bayes estimate recognizes that the household is sampled from a population and uses information about the population to yield an improved estimate of the household's purchase probability.

If the purchase probabilities are distributed with a Dirichlet distribution having parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_J)$, then the distribution of the purchase probability for i^{th} brand for households observed to buy the brand r_i times on N purchase occasions is Beta with parameters $\alpha_i + r_i$ and $s + N - \alpha_i - r_i$ where $s = \sum_{i=1}^{J} \alpha_i$.

The mean of this distribution is the Bayes estimate of purchase probability. It is the minimum variance estimator and it is given by:

$$PB_i = \frac{\alpha_i + r_i}{s + N}$$

However, this method in measuring brand loyalty does have its shortcomings. It tends to favour a brand which has a higher market share as compared to another which has a smaller market share. The brand which has a higher market share may not necessary mean a higher level of brand loyalty as the consumers may have randomly chosen it due to convenience and might actually choose another preferred brand if it had been available!

3. Other Measures of Brand Loyalty

Price until switching (Pessemier 1960): Suppose a particular brand A is purchased by n_1 individuals at the current price p_1. Prices for brand A is then gradually increased to prices $p_2, p_3, \ldots, p_k$, the number of individuals will reduce to $n_2, n_3, \ldots, n_k$. The demand curve can then be drawn for brand A for prices greater than p_1. To draw the other half of the demand curve, the prices for other brands as well as brand A is reduced simultaneously by the same amount. This is done instead of just simply reducing the price of brand A, while keeping the prices of the other brands fixed. Under the later circumstances, switching would not have been the result of a secondary preference for Brand A so much as the result of a more direct price appeal possessed by Brand A when compared to other brands. The price until switching can be obtained, giving a sense of the level of loyalty for the brand.

Brand Allegiance (Hammond 1996):

Respondents were asked to indicate the length of time they had been with the main brand purchased in the telecommunications market. In this way, the longer the time an individual had with the brand, the greater is the sense of loyalty.

Elasticity (Krishnamurthi 1991)

Customers whose brand repurchase is driven by intrinsic product attributes are generally high-value customers because they exhibit a high predisposition to stay with the brand and have low price elasticity (e.g., their sales volume is relatively unaffected by an increase in price). Conversely, customers whose brand repurchase is driven primarily by price/promotion sensitivity are generally low-value consumers because they exhibit low predisposition to stay with the brand through price fluctuations (e.g., high price elasticity).

Market Share Loyalty (Cunningham, 1956)

The measure of brand loyalty chosen for this analysis was drawn from both *single-brand loyalty,* or the proportion of total purchases represented by the largest single brand used; and *dual-brand loyalty,* or the proportion of total purchases represented by the two largest single brands used. In addition, variations of these two measures were developed: *single brand minus deals,* obtained by subtracting from total purchases all those sales made on special price inducements or deals and then calculating the percentage represented by the largest single brand among non deal purchases, and similarly *dual brand minus deals*.

Attitude towards the loyal/disloyal act (Sharp, 1997)

Two 0-10 scales were used, where zero was “totally disagree” and 10 was “totally agree” in response to the following statements:

“I would feel uncomfortable moving to the purchase of another brand”

“I would not like to change my current brand”

Verbal Probability (Jacoby, 1978)

A 0-10 scale, where zero was “no chance or almost no chance” and 10 was “certain, practically certain” that respondents would not change from a given brand in a given time period.

Commitment or Attitude towards the brand measures (Hawkes 1994 and Sharp 1997)

Respondents indicate which of 3 statements best described their feeling towards each brand. Statements included “There are many good reasons to continue to use and no good reasons to change”, “There are many good reasons to continue to use but also many good reasons to change” and “There are few good reasons to use but many good reasons to change”.

Brand Preference (Guest 1944, Guest 1945)

Respondents were asked to indicate the brand they most preferred within a product category. A value of 80% for Brand A means that 80% of the customers prefer A as compared to the other brands within the same category.

CHAPTER 2 MEASURES OF BRAND LOYALTY, PURCHASE CONSISTENCY AND MARKET SHARE IN PROBABILITY CHOICE MODELS FOR TWO-BRAND PROBLEM AND NON-PARAMETRIC CASE

We have briefly defined in Chapter 1 the two types of choice models that can be used to model panel survey data. In this chapter, we will only be considering the independent purchases, i.e. discussing only the Dirichlet model for independent purchases. The non-parametric model for independent purchases will also be considered in this chapter.

In Dirichlet modeling for independent purchases, we make two assumptions regarding purchase incidence in the product category and brand choice probabilities:

(A1) The i^{th} individual's brand choices over a succession of purchases are as if random, with a probability $(p_j)_I$ of choosing brand j from j = 1, …, J brands. These probabilities are fixed over time and brand-choices at successive purchases are assumed independent. The number of purchases of each brand that individual i makes in a sequence of n_j purchases can therefore be modelled by a multinomial with parameters n_j, $(p_1)_I$,…, $(p_J)_I$

(A2) The probabilities $(p_j)_I$ vary among individuals according to a Dirichlet distribution. This is a multivariate Beta-distribution and the joint density function is given in 3.1.2. The special case where J = 2 reduces to the beta distribution.

Some properties of a beta and Dirichlet distribution are described below in 2.1 and 2.2.

2.1 Beta Distribution

A random variable P is said to have a standard beta distribution with parameters α and β if the probability density function is given as:

$$\mathrm{f}(p) = \frac{\Gamma(s)p^{\alpha-1}(1-p)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \qquad 0 \le p \le 1; \alpha, \beta > 0,\ s = \alpha + \beta$$

where $\Gamma(k) = \int_0^\infty y^{k-1} \exp(-y)\,\mathrm{d}y$

Some useful properties of the standard Beta Distribution

a) $\mathrm{E}(P) = \frac{\alpha}{s}$

b) $\mathrm{Var}(P) = \left(\frac{\alpha}{s}\right)\left(1-\frac{\alpha}{s}\right)\left(\frac{1}{1+s}\right)$

2.2 Dirichlet Distribution

The multivariate form of the Beta distribution is known as the Dirichlet distribution. The probability density of the Dirichlet distribution for variables **p** = ($p_1, p_2, \ldots p_J$) with parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_J)$ is defined by

$$\mathrm{f}(p_1, p_2, \ldots p_J) = \frac{\Gamma(s)}{\prod_{j=1}^{J}\Gamma(\alpha_j)} \prod_{j=1}^{J} p_j^{\alpha_j - 1} \text{ where } \sum_{j=1}^{J}\alpha_j = s \text{ and } 0 < p_i < 1 \text{ for all i}$$

Some useful properties of the Dirichlet Distribution

a) $\mathrm{E}(P_j) = \frac{\alpha_j}{s}$

b) $\mathrm{Var}(P_j) = \left(\frac{\alpha_j}{s}\right)\left(1-\frac{\alpha_j}{s}\right)\left(\frac{1}{1+s}\right)$

c) Cov(P_j, P_i) = $-\left(\dfrac{\alpha_j}{s}\right)\left(\dfrac{\alpha_i}{s}\right)\left(\dfrac{1}{1+s}\right)$

2.3 Beta Binomial Distribution

Let R follow a Binomial distribution with parameter N and p. We further assume that p follows a Beta distribution with parameters α and β. Then the marginal distribution of R follows a Beta Binomial distribution. The probability function of the Beta Binomial Distribution with parameters α and β is given by

$$P(R = r) = \frac{\Gamma(s)N!\Gamma(\alpha + r)\Gamma(\beta + N - r)}{\Gamma(s+N)\Gamma(\alpha)r!\Gamma(\beta)(N-r)!} \quad r = 0,1,2,\ldots,N$$

Some useful properties of the Beta Binomial Distribution

a) E(R) = $\dfrac{N\alpha}{s}$

b) Var(R) = $N\left(\dfrac{\alpha}{s}\right)\left(1-\dfrac{\alpha}{s}\right)\left(\dfrac{N+s}{1+s}\right)$

2.4 Dirichlet Multinomial Distribution

The multivariate form of the Beta Binomial distribution is known as the Dirichlet Multinomial distribution. The probability density of the Dirichlet Multinomial distribution for variables **r** = (r_1, r_2, …,r_J) with parameters $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_J)$ is defined by

$$P(R_1 = r_1, R_2 = r_2, \ldots, R_J = r_J) = \frac{\Gamma(s)N!}{\Gamma(s+N)} \prod_{j=1}^{J} \frac{\Gamma(\alpha_j + r_j)}{\Gamma(\alpha_j) r_j!} ,$$

$$r_j = 0, 1, 2, \ldots, N \quad \text{for all} \quad j = 1, 2, \ldots, J, \ \sum_{j=1}^{J} r_j = N \ \text{and} \ \sum_{j=1}^{J} \alpha_j = s$$

Some useful properties of the Dirichlet Multinomial Distribution

a) E(R_j) = $\dfrac{N\alpha_j}{s}$

b) Var(R) = $N\left(\frac{\alpha_j}{s}\right)\left(1-\frac{\alpha_j}{s}\right)\left(\frac{N+s}{1+s}\right)$

c) Cov (R_i , R_j) = $-\frac{N\alpha_i\alpha_j(s+N)}{s^2(s+1)}$

2.5 How Beta and Dirichlet distribution can be applied to our two-brands product category

Suppose there is a product category with only two brands. For example, in the private taxi services, there are currently two main taxi companies, namely "COMFORT" and Trans Island Bus Services (TIBS). Over the population of decision makers, there is a random variable P which has probability density function f(p), represents the probability of a randomly selected decision maker choosing, say COMFORT as a mode of transport.

The probability of selecting the other type of taxi services, TIBS for the same decision maker, is $1-p$. If a sample of k decision makers is taken and the probability of decision maker n choosing COMFORT is p_n, then the probability of him/her choosing TIBS will be $1-p_n$.

Each decision maker is observed to be making N independent choices from the choice set containing the two brands. Over the population of decision makers, there is a random variable R, with a probability distribution function h(R), representing the number of times a randomly selected decision maker choosing, say COMFORT as a mode of transport out of N times. The number of times TIBS is chosen would then be $N-R$.

If a sample of k decision makers is taken and if r_n is the number of times the n^{th} decision maker chooses COMFORT as a mode of transport, where $1 \le n \le k$ and $0 \le r_n \le N$, an estimate of p_n will be $\frac{r_n}{N}$.

A standard distribution for the variable *P*, is the beta distribution. The probability density function had been given in Section 2.1.

The mean of *P* and the variance of *P* can be written as:

E(*P*) = $\frac{\alpha}{s}$

Var (*P*) = $\left(\frac{\alpha}{s}\right)\left(1-\frac{\alpha}{s}\right)\left(\frac{1}{1+s}\right)$

I will use this example to explain how the three important marketing terms – market share, purchase consistency and brand loyalty are related to each other.

Market share:

P is a random variable for the probability of an individual choosing COMFORT while 1 – *P* is a random variable for the probability of an individual choosing TIBS. Thus, E(*P*) will give the expected proportion of individuals choosing COMFORT and 1 – E(*P*) will give the expected proportion of individuals choosing TIBS. The market share for COMFORT can be given as E(*P*) = $\frac{\alpha}{s}$ while the market share for TIBS can be given as 1 – E(*P*) = 1 – $\frac{\alpha}{s}$

Brand Loyalty:

We can rewrite the above expression of Var(*P*) as $\phi = \frac{1}{1+s} = \frac{Var(P)}{\mu(1-\mu)}$ where μ = E(*P*).

The maximum value of ϕ occurs when Var(*P*) equals $\mu(1-\mu)$. This happens when *P* takes on two extreme values 1 and 0 (since P lies between 0 and 1). In order to have E(*P*) = μ, P has to take the values 1 and 0 with probability μ and $1-\mu$ respectively. In such an extreme situation, if an individual decision maker is selected, there will be a probability μ that he will always choose COMFORT and a probability of $1-\mu$ that he will never choose COMFORT but instead will always choose TIBS. In such a situation, we consider the group of individuals as either being totally loyal to COMFORT or totally loyal to TIBS. The other extreme situation occurs when ϕ = 0. In that case, Var (P) = 0, which means that all individuals have the same probability μ of choosing COMFORT. Every individual will choose COMFORT, in an entirely random manner, with their probability of choosing COMFORT being equal to its market share. In this case, we consider the group of individuals as being totally disloyal to COMFORT and TIBS.

Therefore, we can use ϕ as a measure of brand loyalty, with ϕ = 1 being the case when there is extreme loyalty to either COMFORT or TIBS and ϕ = 0 being the case when there is extreme disloyalty to both COMFORT and TIBS.

Purchase Consistency:
We can also interpret the purchase consistency in the same way as that of brand loyalty. Suppose each individual either consistently chooses COMFORT or TIBS (but not both) throughout his/her purchase history, then the value for Var (P) will be $\mu(1-\mu)$, which means that ϕ = 1. On the other hand, if each individual chooses COMFORT or TIBS randomly throughout his/her purchase history with probability μ, we can say that there is little or no purchase consistency in both brands. For such a case Var (P) = 0, which means that ϕ = 0. So there is no difference between brand loyalty and purchase consistency in this case.

ϕ can also be interpreted as the consistency of P among the population of all the decision makers. If ϕ = 1 or close to 1, then this implies the decision makers have a very different values of P among themselves. This is due to the brand loyalty. Hence from the company's viewpoint, any additional advertising targeting at the whole group of decision makers would not help much to increase the market share. If ϕ = 0 or close to 0, then this implies that the decision makers have almost the same value of P. There is little or no brand loyalty in this case. Hence from the company's viewpoint, any additional advertising targeting to the whole group of decision makers would likely lead to an increased market share of its product.

We will next look into how the parameters can be estimated using method of moments and maximum likelihood method.

Suppose *N* choices are been observed from one decision maker and brand A is selected *R* times. If the probability that the decision maker selects Brand A is *p* and if across the population of decision makers, *P* follows the beta distribution with parameters *α* and *β*, then the distribution for *R* across the population follows the beta binomial distribution as given in Section 4.3.

The mean and variance of *R* can be written as:

$$E(R) = \frac{N\alpha}{s} = N \times (\text{share (A)})$$

$$\text{Var}(R) = N \times (\text{share (A)}) \times (\text{share (B)}) \times \frac{N+s}{1+s}$$

2.6 Estimation of the parameters α and s in Probability Choice Models

2.6.1 Method of Moments in Beta and Dirichlet Model

By equating the first and second population moments of R to the corresponding sample moments of R, we have

$$\frac{1}{k}\sum_{j=1}^{k} r_j = \frac{\alpha N}{s}$$

$$\frac{1}{k}\sum_{j=1}^{k} r_j^{\,2} = \frac{N\alpha[N(1+\alpha)+(s-\alpha)]}{s(1+s)}$$

Solving the above equations, we obtain an estimate for *s* and *α* as follows.

$$\hat{s} = \frac{N^2\left(\frac{1}{k}\sum_{j=1}^{k} r_j\right) - N\left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)^2 - N\left(\frac{1}{k}\sum_{j=1}^{k} r_j^2 - \left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)^2\right)}{N\left(\frac{1}{k}\sum_{j=1}^{k} r_j^2 - \left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)^2\right) + \left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)^2 - N\left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)}$$

and

$$\hat{\alpha} = \frac{\hat{s}\left(\frac{1}{k}\sum_{j=1}^{k} r_j\right)}{N}$$

I will now illustrate this by a simple hypothetical example:

A sample of 20 people were selected and asked to record the choices that are being made on the taxi services that they will use in the next 30 occasions. The results are shown below:

Table 1: Data collected on the choices on taxi services.

Individual	1	2	3	4	5	6	7	8	9	10
Comfort (r_1)	18	24	6	25	17	10	12	26	28	20
TIBS (r_2)	12	6	24	5	13	20	18	4	2	10

Table 1 (Continued)

Individual	11	12	13	14	15	16	17	18	19	20
Comfort (r_1)	15	18	20	10	12	25	20	18	25	23
TIBS (r_2)	15	12	10	20	18	5	10	12	5	7

The analysis of the results is shown below:

$\bar{r}_1 = 18.6$, $\frac{1}{20}\sum_{r=1}^{20} r_{1,j}^2 = 382.5$

$\hat{s} = 5.529$, $\hat{\alpha}_1 = 3.428$ and $\hat{\alpha}_2 = 2.101$

$E(P_1) = 0.62$ = share (A) and $E(P_2) = 0.38$ = share (B)

$Var(P_1) = Var(P_2) = 0.0361$

A 90% confidence interval for $\frac{\alpha}{s}$ can be obtained using the bootstrap method (see Efron and Tibshirani (1993)).

Using Bootstrapping method on the above sample 500 times, we can obtain the 90% confidence interval for the $\frac{\alpha_1}{s}$ as (0.548, 0.693)

<u>Remark</u>

While the method of moments estimate are easy to calculate, they can be quite inefficient in some cases, as compared to maximum likelihood estimate (see Fisher, 1921)

<u>2.6.2 Maximum Likelihood estimate of the parameters in the Beta and Dirichlet Model</u>

Another method that can be used in the estimation of the parameters is that of the use of Maximum Likelihood method. The estimate of α and s can be obtained in the following way:

The likelihood function $L(\alpha, s)$ is given by

$$L(\alpha,s)=\prod_{i=1}^{k} g(x_i;\alpha,s)=\prod_{i=1}^{k}\binom{n}{x_i}\frac{\Gamma(s)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(\alpha+x_i)\Gamma(n+\beta-x_i)}{\Gamma(n+s)}$$

Hence the log-likelihood is given by

$$\ln L(\alpha,s)=k\left[\ln\Gamma(s)-\ln\Gamma(\alpha)-\ln\Gamma(s-\alpha)-\ln\Gamma(n+s)\right]+$$
$$\sum_{i=1}^{k}\left[\ln\Gamma(\alpha+x_i)+\ln\Gamma(n+s-\alpha-x_i)\right]+\text{constant}$$

Differentiate ln $L(\alpha, s)$ with respect to α and s, we have

$$\frac{d}{d\alpha}\ln L(\alpha,s)=k\left[-\psi(\alpha)+\psi(s-\alpha)\right]+\sum_{i=1}^{k}\psi(\alpha+x_i)-\sum_{i=1}^{k}\psi(n+s-\alpha-x_i) \text{ and}$$
$$\frac{d}{ds}\ln L(\alpha,s)=k\left[\psi(s)-\psi(s-\alpha)-\psi(n+s)\right]+\sum_{i=1}^{k}\psi(n+s-\alpha-x_i)$$

where $\psi(z)=\frac{d}{dz}\ln\Gamma(z)$.

The maximum likelihood estimates of α and *s* can then be obtained by solving the equations $\frac{\mathrm{d}}{\mathrm{d}\alpha}\ln L(\alpha,s)=0$ and $\frac{\mathrm{d}}{\mathrm{d}s}\ln L(\alpha,s)=0$

Knowing the values of x_i for i = 1, 2,… k, we can use Mathematica software to find the estimates of α and s.

For the example given in the last section about the choices of taxi services, we can use Mathematica to show that the maximum likelihood estimates for α and s to be 3.936 and 6.357 respectively.

I will next show how Rungie (2000) obtained another possible measure of brand loyalty from the Dirichlet distribution, which takes into consideration the concept of repeat purchases in the measure of brand loyalty.

The structure of the Dirichlet Model is to present each decision maker as having one set of brand choice probabilities which is fixed over time. The actual choices made by any one decision maker are independent of the prior choices he/she has made. There is no purchase feedback. However, the presence of the Dirichlet distribution as a mixing distribution for the brand choice probabilities creates an apparent purchase feedback. For any one decision maker the probability of selecting brand A is p. Thus for the one decision maker the probability of selecting brand A on two successive independent choices is p^2.

Over the population, the average probability of selecting brand A on 2 successive occasions is therefore given by:

$$E(P_1^2)=\int_0^1 p_1^{\,2}\frac{\Gamma(s)}{\Gamma(\alpha)\Gamma(\beta)}p_1^{\,\alpha-1}(1-p_1)^{\beta-1}dp_1$$

$$=\frac{\Gamma(2+\alpha)\Gamma(\beta)}{\Gamma(1+s)}\frac{\Gamma(s)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1\frac{\Gamma(2+s)}{\Gamma(2+\alpha)\Gamma(\beta)}p_1^{\,1+\alpha}(1-p_1)^{\beta-1}dp_1$$

$$=\frac{\Gamma(2+\alpha)\Gamma(\beta)}{\Gamma(2+s)}\frac{\Gamma(s)}{\Gamma(\alpha)\Gamma(\beta)}=\frac{\Gamma(2+\alpha)\Gamma(s)}{\Gamma(2+s)\Gamma(\alpha)}=\frac{\alpha(\alpha+1)}{s(s+1)}$$

Thus, with the estimated value of α and s, we have $\mathrm{E}(P_1^2) \approx \dfrac{\Gamma(2+\hat{\alpha})\Gamma(\hat{s})}{\Gamma(2+\hat{s})\Gamma(\hat{\alpha})}$

where the estimates can be obtained from the methods of moments or from the maximum likelihood method.

The ratio of the average probability of selecting brand A for 2 successive occasions to the average probability of selecting brand A for once is given by:

$$\frac{\mathrm{E}(P_1^2)}{\mathrm{E}(P_1)} = \frac{\Gamma(2+\alpha)\Gamma(s)}{\Gamma(2+s)\Gamma(\alpha)} \div \frac{\Gamma(1+\alpha)\Gamma(s)}{\Gamma(1+s)\Gamma(\alpha)} = \frac{1+\alpha}{1+s}$$

Thus, with the estimated value of α and s, we have $\dfrac{\mathrm{E}(P_1^2)}{\mathrm{E}(P_1)} = \dfrac{1+\alpha}{1+s}$

This is the probability of not switching brand A, i.e. the probability of staying loyal.

This ratio is an interesting way to measure the sense of loyalty for brand A. If the ratio is close to 1, it means that the average probability of selecting brand A twice is close to the average probability of selecting brand A once. This would mean that the customers have high chance of not switching the brand A.

The brand loyalty measure $\dfrac{\mathrm{E}(P^2)}{\mathrm{E}(P)}$ can be written as $\dfrac{Var(P)+\mu^2}{\mu} = \phi(1-\mu)+\mu$,

where μ = E(*P*) is the market share. This shows that a brand loyalty measure based on repeat purchase probabilities is a function of the consistency ϕ and the market share μ. This brand loyalty measure is also greater than or equal to the market share μ. This fact is also followed from the fact that the measure is based on a weighted average of values of *P*.

Next, we try to consider the ratio of the average probability of selecting brand A for N+1 successive occasions to the average probability of selecting brand A for N successive occasions. This is given by:

$$\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})} = \frac{N+\alpha}{N+s}$$

There are two very interesting cases to consider for this ratio:

Case 1:

If *N* is large and *s* is small, the value of $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})} \approx 1$.

This means that there is a high level of loyalty for the frequent customer for the brand A.

Case 2:

If both *N* and *s* are large so that $\frac{N}{s} \approx 1$, then $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})} \approx \frac{1}{2}\left(\frac{\alpha}{s}+1\right)$ which is not too close to 1 unless the market share of brand A is 1!

Relationship between $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})}$ and $\frac{\alpha}{s}$

Since $\frac{N+\alpha}{N+s} - \frac{\alpha}{s} = \frac{Ns - N\alpha}{(N+1)s} = \frac{N(s-\alpha)}{(N+s)s}$ and s > α,

we note that $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})} > \frac{\alpha}{s}$ = market share of A.

So, if the market share of A is large, there is a greater tendency for the customers to show a high level of brand loyalty to A.

Relationship between $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^N)}$ and *s*

Suppose $\frac{\alpha_1}{s_1} = \frac{\alpha_2}{s_2}$ but $s_1 > s_2$ and $\alpha_1 > \alpha_2$,

Then, $\frac{N+\alpha_1}{N+s_1} - \frac{N+\alpha_2}{N+s_2} = \frac{N(\alpha_1 - \alpha_2) + N(s_2 - s_1)}{(N+s_1)(N+s_2)}$

Since $|\alpha_1 - \alpha_2| < |s_2 - s_1|$ and $s_1 > s_2$, $\alpha_1 > \alpha_2$

We have $\frac{N+\alpha_1}{N+s_1} - \frac{N+\alpha_2}{N+s_2} < 0$.

Therefore, the larger the value of *s*, the smaller the value of consistency and also the value of $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^N)}$.

<u>Special cases for s and α using the measure of brand loyalty</u>

Case 1: For two brands where $\frac{\alpha}{s} = \frac{1}{2}$ (or $\alpha = \beta$)

In this case, P(*R* = *r*) = $\frac{\Gamma(s)N!\Gamma(\alpha + r)\Gamma(\beta + N - r)}{\Gamma(s+N)\Gamma(\alpha)r!\Gamma(\beta)(N-r)!}$

Then, we would have:

P(*R* = *r*) = $\frac{\Gamma(s)N!\Gamma(\beta + r)\Gamma(\alpha + N - r)}{\Gamma(s+N)\Gamma(\alpha)(N-r)!\Gamma(\beta)(N-r)!}$

= P(*R* = *N* − *r*)

Thus, in this situation, the distribution of R is symmetrical about the median!

Case 2: For commonly used products, the value of *N* tends to be large

Case 2a) Value of *s* is large (probability of switching brands tends to be high), e.g. *N* = 20, s = 800 and α = 600

In this case, the ratio of the average probability of selecting brand A on 20 successive occasions and the average probability of selecting brand A on the prior 19 successive occasions is:

$$\frac{\mathrm{E}(P_1^{20})}{\mathrm{E}(P_1^{19})} = \frac{19+600}{19+800} = 0.756 .$$

Also, the ratio of the average probability of selecting brand A on 2 successive occasions and the average probability of selecting brand A on the first occasion is:

$$\frac{\mathrm{E}(P_1^{2})}{\mathrm{E}(P_1\ )} = \frac{1+600}{1+800} = 0.750 .$$

Note that these values are very close to the market share for brand A, given by $\frac{\alpha}{s}$ = 0.75.

Also, Var(P) = $\left(\frac{600}{800}\right)\times\left(\frac{200}{800}\right)\times\left(\frac{1}{1+800}\right) = 0.000234$.

This means that there is little difference in the purchase probabilities among decision makers!

One commonly known product which is of this kind is the purchase of petrol at kiosks. The location of the petrol kiosks or price at different kiosks may be of greater importance as compared to the brand itself. Thus, car owners may switch brands quite rapidly depending on convenience or price differences.

The probability distribution function of R, in this case is given by:

P(R = r) = $\frac{\Gamma(800)20!\Gamma(600+r)\Gamma(220-r)}{\Gamma(820)\Gamma(600)r!\Gamma(200)(20-r)!}$, r = 0,1,…,20

The distribution for R is given in Table 2 below:

Table 2 : Distribution for R for case 2a

R	<=6	7	8	9	10	11	12
P(R = r)	0.0002	0.0002	0.0009	0.0033	0.0105	0.0280	0.0616

13	14	15	16	17	18	19	20
0.1121	0.1668	0.1998	0.1883	0.1344	0.0684	0.0221	0.0034

From Table 2 above, a few interesting features can be noted:

a) Though the market share for the brand is 0.75, the probability of the brand being chosen for 18, 19 or 20 times out of 20 is relatively low. This can be attributed to the high possibility of brand switching among decision makers.

b) The probability of the brand being chosen for less than 7 occasions out of 20 is very low ($\approx$ 0.0002), which is very similar to that of a binomial distribution with n = 20 and p = 0.75. The switching of brands are done in a rather random manner!

Case 2b) Value of s is small (probability of switching brands tends to be low), e.g. N = 20, s = 4 and α = 3

In this case, the ratio of the average probability of selecting brand A on 20 successive occasions and the average probability of selecting brand A on the prior 19 successive occasions is:

$$\frac{\mathrm{E}(P_1^{20})}{\mathrm{E}(P_1^{19})} = \frac{19+3}{19+4} = 0.957 .$$

Also, the ratio of the average probability of selecting brand A on 2 successive occasions and the average probability of selecting brand A on the first occasion is:

$$\frac{E(P_1^2)}{E(P_1)} = \frac{1+3}{1+4} = 0.80 .$$

In this situation, these values are significantly higher than the market share of 0.75.

Also, Var(P) = $\left(\frac{3}{4}\right) \times \left(\frac{1}{4}\right) \times \left(\frac{1}{1+4}\right) = 0.0375$

This means that there is a significant difference in the purchase probabilities among decision makers as compared to the earlier case in 2a. Thus, the smaller the value of s, the greater is the difference in the purchase probabilities.

One commonly known product that has low value of s is that of the purchase of rice. Decision makers do not change the brands of rice purchased as their individual tastes do not vary rapidly over time.

The probability distribution function of R, in this case is given by:

P(R = r) = $\dfrac{\Gamma(4)20!\Gamma(3+r)\Gamma(21-r)}{\Gamma(24)\Gamma(3)r!\Gamma(1)(20-r)!}$, r = 0,1,…20

The distribution table for R is given below:

Table 3 : Distribution for R for case 2b

R	0	1	2	3	4	5	6	7
P(R = r)	0.0006	0.0017	0.0034	0.0057	0.0085	0.0119	0.0158	0.0203

Table 3 (Continued)

R	8	9	10	11	12	13	14	15
P(R = r)	0.0254	0.0311	0.0373	0.0440	0.0514	0.0593	0.0678	0.0768

Table 3 (Continued)

R	16	17	18	19	20
$P(R = r)$	0.0864	0.0966	0.1073	0.1186	0.1304

From the table above, a few interesting features can be noted:

a) Though the market share for the brand is 0.75, the probability of the brand being chosen for 18, 19 or 20 times out of 20 is exceptionally high. This can be attributed to the low possibility of brand switching among decision makers. When the brand had been chosen on the last occasion, there is a high possibility that it will be selected again on the next occasion. Thus, there is high possibility that the brand may be chosen on all the 20 occasions even though the market share is only 0.75.

b) The probability of the brand being chosen for less than 7 occasions out of 20 is quite significant (≈ 0.0478). This is approximately 240 times larger than the case when switching of brands is high. This means that brands are chosen in a certain noticeable pattern and clearly not random.

Case 3: For products where purchases are rare

There are many products or services in the market where purchases or utilization of the product/service is rare. In most cases, N may be just 1 or 2. This problem can be dealt with by either modeling the probability of selecting a particular brand using logistic regression. This will be discussed in greater detail in the next chapter.

2.7 Limitations of Dirichlet Distribution in modeling purchase probabilities

Yim and Kannan (1998) gave some limitations of Dirichlet Modeling (1-5) with regards to loyalty:

1. It was not known how many buyers purchase a particular brand exclusively and how many have divided loyalties?
2. It was not known why some buyers exhibited divided loyalties
3. It was not known whether behaviour was driven by loyalties to certain product attributes or whether it was an outcome of marketing mix actions?

4. Nothing can be deduced about how a firm can do to maintain an exclusive loyal buyer base?
5. It was not known what actions can be taken to build the firm's position among the divided loyals?

Also, in our measure of brand loyalty in Dirichlet Distribution, it was assumed that the each individual has a fixed probability of choosing a particular brand. This is certainly not a realistic assumption to be made in the real world where individual's probability of choosing a particular brand may change on subsequent purchases, depending on the level of satisfaction from the recent purchases.

The introduction of covariates in our modeling will work towards trying to deal with the limitations 2-5 given above. The Logistic Regression Modeling will be discussed in chapter 3.

2.8 Estimate of the market share, purchase consistency and brand loyalty in the non-parametric case

In the non-parametric case, we can estimate E(P) using $\overline{p} = \frac{1}{k}\sum_{i=1}^{k}\hat{p}_i$ where $\hat{p}_i = \frac{x_i}{n_i}$ is the estimate of the probability of buying brand A for the ith individual decision maker. Here, x_i gives the number of purchases made of brand A by ith individual while n gives the number of purchases made by i^{th} individual. Also, we can estimate Var(P) using $\frac{1}{k-1}\sum_{i=1}^{k}\left(\hat{p}_i - \overline{p}\right)^2$. In this way, the market share, purchase consistency and brand loyalty can be estimated in the non-parametric case.

Next, we consider repeat purchase probabilities. Consider the conditional probability of buying brand A twice in a row given that he/she had purchased brand A at the first instance. If we let p be the probability of buying brand A, then the conditional probability of a repeat purchase of A, given one purchase of A is just *p*. Therefore, some average of the values of *p* over the population of decision maker may be used as a measure of brand loyalty of individual decision maker. However E(P) = μ. Hence, this way of defining brand loyalty is not useful as it cannot be distinguished from the market share.

In the non-parametric case, we can still make use of the similar concept of brand loyalty by writing brand loyalty as $\frac{\mathrm{E}(P^2)}{\mathrm{E}(P)}$ = $\frac{Var(P)+\mu^2}{\mu} = \phi(1-\mu)+\mu$. The expression for purchase consistency (ϕ) and market share (μ) is written as before.

To obtain an estimate of the brand loyalty, we can first estimate E(P^2) as $\overline{P^2} = \frac{1}{k}\sum_{i=1}^{k}\hat{p}_i^2$ where $\hat{p}_i^2 = \frac{y_i}{n_i}$ is the estimate of the probability that the i^{th} individual purchases brand A twice in a row. Here, n_i and y_i give the number of pairs of purchases made and the number of consecutive pairs of brand A being purchased by i^{th} individual respectively. E(P) is estimated in the same way as before. The brand loyalty is then estimated as $\frac{\overline{P^2}}{\bar{p}}$.

CHAPTER 3: LOGISTIC REGRESSION MODELING

This chapter discusses a model that characterizes a choice from discrete alternatives by a decision maker as a function of attributes associated with each alternative as well as the characteristics of the individual. This is termed as a multinomial logit model. For this chapter, we will be looking only at the 2-brand product category and in that case, the multinomial logit model reduces to the logistic regression model.

This method assumes that the probability of selecting a particular brand is dependent on a group of independent variables which measure the characteristics of the decision maker. These variables are called explanatory variables.

For each decision maker, these variables are observed once being $x_1, x_2, ..., x_l$. These explanatory variables are related to the probability of selecting a particular brand through a function, called the "link" function. In logistic regression, the "link" function is "logit".

So the model can be written as:

$$\log \frac{\hat{p}}{1-\hat{p}} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_l x_l$$

or $$\hat{p} = \frac{1}{1 + \exp\left\{-(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_l x_l)\right\}}$$

Suppose that the two brands are A and B and that x_i refers to gender (x_i = 0 for female and x_i = 1 for male.) Then, we can interpret the coefficients of the logistic regression model by considering for $i \neq 0$,

$e^{\beta_i} = \dfrac{\text{odds of selecting brand A among males}}{\text{odds of selecting brand A among females}}$ while the rest of the explanatory variables remain the same.

Also, e^{β_0} is the odds of selecting brand A when all the explanatory variables give a value zero.

Note that this is not possible in many cases as some explanatory variables cannot be zero. (e.g. if x_i refers to age of a person, it does not make much sense in general to discuss the odds of selecting brand A for a new born baby.)

The observation for the dependent variable is a single discrete choice r_n between the two brands and it takes value 0 or 1.

In this case, we assume that R_n has a Bernoulli distribution with parameter p_n.

The likelihood function for $(R_1, R_2, \ldots R_k)$ is $\prod_{n=1}^{k} p^{r_n}(1-p)^{1-r_n}$

The log likelihood for $(R_1, R_2, \ldots R_k)$ is:

$$l = \sum_{n=1}^{k} l_n = \sum_{n=1}^{k} r_n \ln(p) + (1-r_n)\ln(1-p)$$

Now, $E(R_n) = p$ and $Var(R_n) = p\,(1-p)$

Let $\log \dfrac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_l x_l$

The score vector has element $U_j = \sum_{n=1}^{k} (r_n - p)x_j$

The information matrix J will have elements J_{js} where

$$J_{js} = E(U_j U_s) = E\left\{\sum_{n=1}^{k}(R_n - p)x_j \sum_{l=1}^{k}(R_l - p)x_s\right\}$$

Suppose a sample of 200 people is taken and also x_1 refers to course of study (x_1 = 0 means Professional degree and x_1 = 1 means non-Professional degree), x_2 refers to academic results (x_2 = 0 means top 30% of university cohort, x_2 = 1 means the bottom 70% of university cohort), x_3 refers to gender (x_3 = 0 means males, x_3 = 1 means females) and r_n refers to university chosen (r_n = 1 means NUS and r_n = 0 means NTU.) The data collected is summarized below:

Table 4: Observations on University Data

x_1	x_2	x_3	Number of observations with r_n =1	Number of observations with r_n =0
0	0	0	20	10
0	0	1	15	6
0	1	0	10	15
0	1	1	6	10
1	0	0	30	15
1	0	1	20	10
1	1	0	5	11
1	1	1	7	10
Total			113	87

Using Minitab, we obtain the logistic regression model in this example as:

$$\ln\left(\frac{\hat{p}}{1-\hat{p}}\right) = 0.729 - 1.226x_2$$

The other two variables are not significant in the model and are removed from the model.

We can observe that the proportion of top 30% students (i.e. x_2 is 0) choosing NUS is 67.5%, while the proportion of the rest of the students (i.e. x_2 is 1) choosing NUS is 37.8%. If the data collected is a good representative sample, then, NUS seems to be a more popular choice among the top students.

Robert East, Patricia Harris, Wendy Lomax and Gill Willson (1997) defined the loyalty to a particular brand as high if the proportion of purchasing the brand is 0.81 or more and defined it as low if the proportion of purchasing the brand is below 0.81.

CHAPTER 4: PROBABILITY CHOICE MODELS AND MULTINOMIAL LOGIT REGRESSION FOR MORE THAN TWO BRANDS PROBLEM WITH MEASURES OF BRAND LOYALTY

4.1 How Beta and Dirichlet distribution can be extended to more than two brands product category

The two brands example given in the last chapter is now extended to a more general case where there can be more than two brands in the product category. Suppose that there are J brands in the product category and let the individual brand be j, j = 1, 2, 3, …, J. Over the population of decision makers, there are J random variables P_j representing the probability of choosing j^{th} brand among all the brands in the same product category. (note that $\sum_{j=1}^{J} p_j = 1$).

The J random variables P_1, P_2, …, P_J form a vector $\boldsymbol{P}$ with a probability density function f($\boldsymbol{p}$). A standard multivariate distribution for the variable $\boldsymbol{P}$ is the Dirichlet distribution. The probability density function has been given in Section 2.2 and the mean and variance of P_j for j = 1, 2, 3, …, J can be written as:

$$E(P_j) = \frac{\alpha_i}{s} = \text{share (j)}$$

$$\text{Var}(P_j) = \left(\frac{\alpha_j}{s}\right)\left(1-\frac{\alpha_j}{s}\right)\left(\frac{1}{1+s}\right) = \text{(share (j))(1– share (j))(consistency of brand choice)}$$

Suppose that there is a sample of k decision makers. The brand choice probability given by $p_{j,n}$ is the probability of decision maker n choosing brand j, where $\sum_{j=1}^{J} p_{j,n} = 1$. Each decision maker is observed making N independent choices. From these choices over the population of decision makers, brand j is selected R_j times where $0 \le R_j \le N$ and $\sum_{j=1}^{J} R_j = N$. The J random variables R_1,

$R_2, \ldots, R_J$ form a vector $\boldsymbol{R}$ with a probability distribution function h($\boldsymbol{R}$). The distribution of $\boldsymbol{R}$ is based on J+1 parameters, consisting of (1) the number of trials, N and (2) the probability of success for each brand p_j, j = 1, 2,…, J.

A standard multivariate distribution for the vector R is the Dirichlet Multinomial distribution. The probability distribution function for R is given in Section 5.4. The mean of R_j, variance of R_j and covariance of R_i & R_j for i = 1, 2,…, J and j = 1, 2,…, J, $i \neq j$, can be written as :

$$E(R_j) = \frac{\alpha_j N}{s} = N(\text{share } (j))$$

$$\mathrm{Var}(R_j) = N\left(\frac{\alpha_j}{s}\right)\left(1-\frac{\alpha_j}{s}\right)\left(\frac{N+s}{1+s}\right) = N(\text{share } (j))\ (1 - \text{share } (j))\left(\frac{N+s}{1+s}\right)$$

$$\mathrm{Cov}\ (R_i, R_j) = -\frac{N\alpha_i\alpha_j(s+N)}{s^2(s+1)}$$

Let's consider a simple example to illustrate this case. In the current market for the sale of handphones, the main brands are "NOKIA", "SAMSUNG", "SONY ERICSSON", "SIEMENS" and "MOTOROLA". Suppose that the market share of these brands are estimated to be respectively 0.35, 0.30, 0.15. 0.10, 0.10 and that s can be estimated to be 10.

The values of α_i, i = 1, 2, … , 5 are :

$\alpha_1 = 0.35 \times 10 = 3.5$, $\alpha_2 = 0.30 \times 10 = 3.0$, $\alpha_3 = 0.15 \times 10 = 1.5$,

$\alpha_4 = \alpha_5 = 0.10 \times 10 = 1.0$

and the distribution of $\boldsymbol{P} = (P_1, P_2, P_3, P_4, P_5)$ is given by

$$f(p_1,p_2,p_3,p_4,p_5) = \frac{\Gamma(10)}{\prod_{j=1}^{5}\Gamma(\alpha_j)}\prod_{j=1}^{5} p_j^{\alpha_j-1}$$

$$= \frac{9!}{\Gamma(3.5)\Gamma(3)\Gamma(1.5)\Gamma(1)\Gamma(1)} p_1^{2.5}p_2^{2}p_3^{0.5} = \frac{193536}{\pi} p_1^{2.5}p_2^{2}p_3^{0.5}$$

Also, E(P_1) = 0.35, E(P_2) = 0.30, E(P_3) = 0.15, E(P_4) = E(P_5) = 0.10

and Var(P_1) = $(0.35)\times(0.65)\times\left(\frac{1}{1+10}\right) = 0.0207$,

Var(P_2) = $(0.30)\times(0.70)\times\left(\frac{1}{1+10}\right) = 0.0191$,

Var(P_3) = $(0.15)\times(0.85)\times\left(\frac{1}{1+10}\right) = 0.0116$,

Var(P_4) = Var(P_5) = $(0.10)\times(0.90)\times\left(\frac{1}{1+10}\right) = 0.0082$.

Thus, the variation of P_j is the greatest among decision makers choosing "NOKIA" and least among decision makers choosing "SIEMENS" and "MOTOROLA". Suppose that a sample of 20 purchasers is taken and each decision maker had been observed to have made 3 purchases.

The distribution of $\boldsymbol{R}$ = (R_1,R_2,R_3,R_4,R_5) is given by:

$$P(R_1 = r_1, R_2 = r_2, R_3 = r_3, R_4 = r_4, R_5 = r_5)$$

$$= \frac{\Gamma(10)3!\Gamma(3.5+r_1)\Gamma(3+r_2)\Gamma(1.5+r_3)\Gamma(1+r_4)\Gamma(1+r_5)}{\Gamma(13)\Gamma(3.5)\Gamma(3)\Gamma(1.5)\Gamma(1)\Gamma(1)r_1!r_2!r_3!r_4!r_5!}$$

$$= \frac{2\Gamma(3.5+r_1)\Gamma(3+r_2)\Gamma(1.5+r_3)}{825\pi\, r_1!r_2!r_3!}$$

Also,

E(R_1) $= 3 \times 0.35 = 1.05$, E(R_2) $= 3 \times 0.30 = 0.90$, E(R_3) $= 3 \times 0.15 = 0.45$,

E(R_4) $= \mathrm{E}(R_5) = 3 \times 0.10 = 0.3$

and

Var(R_1) = $3 \times (0.35) \times (0.65) \times \left(\frac{3+10}{1+10}\right) = 0.8066$,

Var(R_2) = $3 \times (0.30) \times (0.70) \times \left(\frac{3+10}{1+10}\right) = 0.7445$,

Var(R_3) = $3 \times (0.15) \times (0.85) \times \left(\frac{3+10}{1+10}\right) = 0.4520$,

Var(R_4) = Var(R_5) = $3 \times (0.10) \times (0.90) \times \left(\frac{3+10}{1+10}\right) = 0.3191$

We will now try to estimate the parameters of the two distributions so that statistical inferences can be made about the purchases.

4.2 Estimation of the parameters α_j and s for j=1,2,…,J

Method of Moments

From the equations of expectation and variance of R_j,

E(R_j) $= \frac{\alpha_j N}{s}$ and

Var(R_j) = $N\left(\frac{\alpha_j}{s}\right)\left(1-\frac{\alpha_j}{s}\right)\left(\frac{N+s}{1+s}\right)$

We have j =1,2,…,J:

Var (R_j) = $N\frac{\mathrm{E}(R_j)}{N}\left(1-\frac{\mathrm{E}(R_j)}{N}\right)\left(\frac{N+s}{1+s}\right)$

$$\mathrm{Var}[R_j] = N\frac{\alpha_j}{s}\left(1-\frac{\alpha_j}{s}\right)\frac{N+s}{1+s}$$

Hence we consider taking average of Var[R_j] over j , we have

$$\frac{1}{J}\sum_{j=1}^{J}\mathrm{Var}[R_j] = \frac{N+s}{1+s}\left[\frac{1}{JN}\sum_{j=1}^{J}\mathrm{E}(R_j)\left(N-\mathrm{E}(R_j)\right)\right]$$

Substitute E(R_j) and Var(R_j) by the corresponding sample estimates $\hat{\mu}_j$ and $\hat{\sigma}_j^2$, where $\hat{\mu}_j = \frac{1}{k}\sum_{l=1}^{k} r_{j,l}$ and $\hat{\sigma}_j^2 = \frac{1}{k}\sum_{l=1}^{k} r_{j,l}^2 - \left(\frac{1}{k}\sum_{j=1}^{k} r_{j,l}\right)^2$ with $r_{j,l}$ is the number of purchases made by l^{th} decision maker on brand j.

Thus, we have:

$$\hat{s} = \frac{\sum_{j=1}^{J}\left(\hat{\mu}_j(N-\hat{\mu}_j)-\hat{\sigma}_j^2\right)}{\sum_{i=1}^{J}\left(\hat{\sigma}_j^2 - \frac{1}{N}\hat{\mu}_j(N-\hat{\mu}_j)\right)} \text{ and } \hat{\alpha}_j = \frac{\hat{\mu}_j\hat{s}}{N} \text{ for } j = 1, 2, \ldots, \mathrm{J}.$$

Let's extend our earlier example given on the bus services taken by 10 commuters to include "Mass Rapid Transit (MRT)" and "small scale private buses". Suppose also, that the 10 commuters are asked to record the choices that are being made on these four forms of public transport that they will use in the next 50 occasions. The results are shown below:

Table 5: Observations on Public Transport Choices

Individual	1	2	3	4	5	6	7	8	9	10
SBS (r_1)	22	25	10	40	15	15	12	35	25	5
TIBS (r_2)	11	5	10	10	30	0	10	5	0	25
MRT (r_3)	15	12	30	0	5	20	22	8	25	10
Small scale private buses	2	8	0	0	0	15	6	2	0	10

The analysis of the results is shown below:

E(R_1) $\approx \bar{r}_1 = 20.4$, Var (R_1) ≈ 11.14^2 , E(R_2) $\approx \bar{r}_2 = 10.6$, Var (R_2) ≈ 9.85^2

E(R_3) $\approx \bar{r}_3 = 14.7$, Var (R_3) ≈ 9.46^2 , E(R_4) $\approx \bar{r}_4 = 4.3$, Var (R_4) ≈ 5.25^2

$\hat{s}$ = 4.611, $\hat{\alpha}_1$ = 1.881, $\hat{\alpha}_2$ = 0.977, $\hat{\alpha}_3$ = 1.356, $\hat{\alpha}_4$ = 0.397.

So, E(P_1) = 0.41 = share (SBS) , E(P_2) = 0.21 = share (TIBS) ,

E(P_3) = 0.29 = share (MRT) and E(P_4) = 0.09 = share (small private buses)

4.3 Maximum Likelihood estimate of the parameters in "more than two brand category"

From the distribution of ***R*** of a sample of *k* individuals and *J* brands, we have:

$$f(r_1,\cdots,r_J) = \frac{n!}{r_1!\cdots r_J!}\frac{\Gamma(s)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_J)}\frac{\Gamma(\alpha_1+r_1)\cdots\Gamma(\alpha_J+r_J)}{\Gamma\left(s+\sum_{j=1}^{J}r_j\right)}.$$

The log-likelihood function of α_1, …, α_{J-1} and s is given by

$$\ln L(\alpha_1,\alpha_2,\cdots,\alpha_{J-1},s) = k\left[\ln\Gamma(s) - \sum_{j=1}^{J-1}\ln\Gamma(\alpha_j) - \ln\Gamma\left(s-\sum_{j=1}^{J-1}\alpha_j\right)\right] +$$
$$\sum_{l=1}^{k}\left[\sum_{j=1}^{J-1}\ln\Gamma(\alpha_j+r_{j,l}) + \ln\Gamma\left(s-\sum_{j=1}^{J-1}\alpha_j + r_{J,l}\right) - \ln\Gamma\left(s+n\right)\right] + \text{constant}$$

Differentiate ln $L(\alpha_1,\ldots,\alpha_{J-1},s)$ with respect to α_j, j = 1, …, J -1, and s, we have

$$\frac{d}{d\alpha_j}\ln L(\alpha_1,\cdots,\alpha_{J-1},s) = k\left[-\psi(\alpha_j)+\psi(s-\sum_{j=1}^{J-1}\alpha_j)\right] + \sum_{l=1}^{k}\left[\psi(\alpha_j+r_{j,l})+\psi(s-\sum_{j=1}^{J-1}\alpha_j+r_{J,l})\right]$$

and

$$\frac{d}{ds}\ln L(\alpha_1,\cdots,\alpha_{J-1},s) = k\left[\psi(s)-\psi(s-\sum_{j=1}^{J-1}\alpha_j)\right] + \sum_{l=1}^{k}\psi\left(s-\sum_{j=1}^{J-1}\alpha_j + r_{J,l}\right) - \psi(s+n)$$

where $\psi(z) = \dfrac{d}{dz}\ln\Gamma(z)$.

The maximum likelihood estimates of $\alpha_1,\ldots,\alpha_{J-1}$, and s can then be obtained by solving the equations $\frac{d}{d\alpha_j}\ln L(\alpha_1,\cdots\alpha_{J-1},s)=0$, j = 1, …., J – 1 and $\frac{d}{ds}\ln L(\alpha_1,\cdots\alpha_{J-1},s)=0$.

4.4 Measure of Brand Loyalty based on attitudes of consumers (from the probability perspective) for more than two-brand problem

For any one decision maker, the probability of selecting brand j is P_j and the probability of selecting brand j on all of N successive independent choices is p_j^N

In a similar way as before, it can be shown that

$$E(P_j^N)=\frac{\Gamma(s)\Gamma(\alpha_j+N)}{\Gamma(\alpha_j)\Gamma(s+N)}$$

$$=\frac{(\alpha_j+N-1)(\alpha_j+N-2)...(\alpha_j)}{(s+N-1)(s+N-2)...(s)} \text{ for } j=1,\ldots,J$$

Thus, with the estimated value of α and s, we have

$\mathrm{E}(P_j^N)\approx\frac{\Gamma(N+\hat{\alpha}_j)\Gamma(\hat{s})}{\Gamma(N+\hat{s})\Gamma(\hat{\alpha}_j)}$ where the estimates can be obtained from the methods of moments or from the maximum likelihood method. You may note that this average probability of choosing brand j for N successive occasions is very similar to the case for two-brand problem. It depends on α_i and s and is independent of other α_i

Next we try to consider the ratio of the average probability of selecting brand j for (N+1) successive occasions to the average probability of selecting brand j for N successive occasions as a brand loyalty measure. This is given by:

$$\frac{\mathrm{E}(P_{\mathrm{j}}^{N+1})}{\mathrm{E}(P_{\mathrm{j}}^{N})} = \frac{N+\alpha_{\mathrm{j}}}{N+s}$$

Thus, with the estimated value of α_j and s, we have $\frac{\mathrm{E}(P_{\mathrm{j}}^{N+1})}{\mathrm{E}(P_{\mathrm{j}}^{N})} \approx \frac{N+\hat{\alpha}_{\mathrm{j}}}{N+\hat{s}}$

Similar conclusion can be made when we look at the relationship between $\frac{\mathrm{E}(P_{\mathrm{j}}^{N+1})}{\mathrm{E}(P_{\mathrm{j}}^{N})}$ and $\frac{\alpha_{\mathrm{j}}}{s}$.

Hence if the market share of j is large, then the expected proportion of buying Brand j N +1 times is close to that of the expected proportion of buying Brand j N times.

Also, by observing the relationship between $\frac{\mathrm{E}(P_1^{N+1})}{\mathrm{E}(P_1^{N})}$ and *s*, we can make analogous conclusion to the two-brand problem - the larger the value of *s*, the smaller the value of consistency and also the value of $\frac{\mathrm{E}(P_{\mathrm{j}}^{N+1})}{\mathrm{E}(P_{\mathrm{j}}^{N})}$.

4.5 Multinomial Logit Model

We have seen in Chapter 3 how logistic regression model can be used to model the probability of choosing a particular brand for the situation when the number of purchases is rare, in some cases, occurring only once in the lifetime. We will now extend this idea to the situation when there are more than two brands in the same product category.

Suppose that there are *l* identical explanatory variables for the brands. We use $\beta_{1,j}, \beta_{2,j}, \ldots, \beta_{1,j}$ to denote the values of the explanatory variables $x_1, x_2, \ldots x_l$, choosing brand j, $1 \leq \mathrm{j} \leq \mathrm{J}$.

The model linking the probability of selecting a particular brand j, with that of the explanatory variables for each decision maker is given as:

$$p_j = \frac{\exp(\beta_{0,j} + \beta_{1,j}x_1 + \cdots + \beta_{l,j}x_l)}{\sum_{h=1}^{J}\exp(\beta_{0,h} + \beta_{1,h}x_1 + \cdots + \beta_{l,h}x_l)} \text{ for } j = 1, 2, \ldots, J.$$

As $\sum_{\mathrm{j=1}}^{\mathrm{J}} p_{\mathrm{j}} = 1$ is a constraint that has to be satisfied, one of the parameters among $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_\mathbf{J}$ is redundant in this representation.

We can thus re-normalize the above model by choosing one of the parameters to be zero. (Suppose without loss of generality, we choose $\boldsymbol{\beta}_\mathbf{J}$ to be zero.) This leads to a more simplified version as shown below:

$$\hat{p}_j = \frac{\exp(\hat{\beta}_{0,j} + \hat{\beta}_{1,j}x_1 + \hat{\beta}_{2,j}x_2 + \ldots + \hat{\beta}_{l,j}x_l)}{1 + \sum_{i=1}^{\mathrm{J}-1}\exp(\hat{\beta}_{0,i} + \hat{\beta}_{1,i}x_1 + \hat{\beta}_{2,i}x_2 + \ldots + \hat{\beta}_{l,i}x_l)} \quad \text{---- (1)}, \quad 1 \le \mathrm{j} \le \mathrm{J} - 1$$

and

$$\hat{p}_J = \frac{1}{1 + \sum_{i=1}^{\mathrm{J}-1}\exp(\hat{\beta}_{0,i} + \hat{\beta}_{1,i}x_1 + \hat{\beta}_{2,i}x_2 + \ldots + \hat{\beta}_{l,i}x_l)} \quad \text{---- (2)}$$

Suppose that the J brands are A, B,…,J and that x_i refers to gender (x_i = 0 for female and x_i = 1 for male). Then, we can interpret the coefficients of the logistic regression model by considering for $i \neq 0$,

$$e^{\beta_{i,j}} = \frac{\text{odds of selecting brand j against J among males}}{\text{odds of selecting brand j against J among females}}$$ while the rest of the explanatory variables remain the same

and

$e^{\beta_{0,j}}$ is the odds of selecting brand j against brand J when all the explanatory variables give a value zero.

Two interesting observations can be made from this model:

a) For two-brand category, of which one of the brands is brand j, the model reduces to :

$$\hat{p}_j = \frac{1}{1+\exp\{-(\hat{\beta}_{0,j} + \hat{\beta}_{1,j}x_1 + \hat{\beta}_{2,j}x_2 + ... + \hat{\beta}_{l,j}x_l)\}}$$

b) We can make use of the ratio of (1) and (2) to obtain :

$$\log\left(\frac{\hat{p}_{\mathrm{j}}}{\hat{p}_{\mathrm{J}}}\right) = \hat{\beta}_{0,j} + \hat{\beta}_{1,j}x_1 + \hat{\beta}_{2,j}x_2 + ... + \hat{\beta}_{l,j}x_l$$

I will now make use of likelihood theory to find the maximum likelihood estimator for the parameters $\underline{\beta}_j = (\beta_{0,j}, \beta_{1,j}, \cdots, \beta_{l,j})'$, $1 \le \mathrm{j} \le \mathrm{J}-1$

Let $\underline{x} = (x_1, \ldots, x_l)'$ be the vector of the covariate and $\underline{\beta}_j = (\beta_{0,j}, \beta_{1,j}, \cdots, \beta_{l,j})'$. Then

$$p_j = \frac{\exp(\underline{\beta}_j' \underline{x})}{\sum_{h=1}^{J} \underline{\beta}_h' \underline{x}}$$

Let $Q_h = \sum_{j=1}^{J} r_{h,j} \ln p_j$. Then $\ln L = \sum_{h=1}^{k} Q_h$.

Let us drop the subscript h in the subsequent discussion and consider $\frac{dQ}{d\beta_{i,j}}$. Let

P be the diagonal matrix with the diagonal $\mathbf{p} = (p_1, p_2, ..., p_J)'$, $r = (r_1, r_2, ..., r_J)'$, and $\underline{\beta}_j' = (\beta_{0,j}, \beta_{1,j}, \cdots, \beta_{l,j})$.

$$\frac{\partial Q}{\partial \beta_{i,j}} = \frac{\partial \sum_{j=1}^{J} r_j \ln p_j}{\partial \beta_{i,j}} = \frac{\partial r_1 \ln p_1}{\partial \beta_{i,j}} + \frac{\partial r_2 \ln p_2}{\partial \beta_{i,j}} + \cdots + \frac{\partial r_J \ln p_J}{\partial \beta_{i,j}}$$

Let us consider $\frac{\partial r_m \ln p_m}{\partial \beta_{ij}}$.

For $m \neq j$.

$$\frac{\partial r_m \ln p_m}{\partial \beta_{i,j}} = \frac{\partial r_m \ln p_m}{\partial p_m}\frac{\partial p_m}{\partial \beta_{i,j}}$$

$$= \frac{r_m}{p_m}\frac{\partial}{\partial \beta_{i,j}}\left(\frac{\exp\left(\sum_{i=1}^{l}\beta_{i,m}x_i\right)}{\sum_{n=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,n}x_i\right)}\right)$$

$$= \frac{r_m}{p_m}\frac{\exp\left(\sum_{i=1}^{l}\beta_{i,m}x_i\right)}{\left(\sum_{n=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)\right)^2}\left(-x_i \exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)\right)$$

$$= -\frac{r_m}{p_m}p_m x_i p_j = -r_m x_i p_j$$

For $m = j$.

$$\frac{\partial r_j \ln p_j}{\partial \beta_{i,j}} = \frac{\partial r_j \ln p_j}{\partial p_j}\frac{\partial p_j}{\partial \beta_{i,j}}$$

$$= \frac{r_j}{p_j}\frac{\partial}{\partial \beta_{i,j}}\left(\frac{\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)}{\sum_{n=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,n}x_i\right)}\right)$$

$$= \frac{r_j}{p_j}\left(\frac{x_i \exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)}{\left(\sum_{n=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)\right)} + \frac{\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)}{\left(\sum_{n=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)\right)^2}\left(-x_i \exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)\right)\right)$$

$$= \frac{r_j}{p_j}\left(x_i p_j - x_i p_j^2\right) = r_j x_i \left(1 - p_j\right)$$

Therefore

$$\frac{\partial Q}{\partial \beta_{i,j}} = \left(\sum_{m=1}^{J} -r_m x_i p_j\right) + r_j x_i = -x_i p_j \left(\sum_{m=1}^{J} r_m\right) + x_i r_j = x_i \left(r_j - np_j\right)$$

$$= x_i\left(r_j - n\frac{\exp\left(\sum_{i=1}^{l}\beta_{i,j}x_i\right)}{\sum_{m=1}^{J}\exp\left(\sum_{i=1}^{l}\beta_{i,m}x_i\right)}\right)$$

Let

$$\underline{\alpha}' = \left(\alpha_1, \alpha_2, \cdots, \alpha_J, \alpha_{J+1}, \cdots\cdots, \alpha_{(l+1)J}\right)$$
$$= \left(\beta_{0,1}, \beta_{1,1}, \cdots, \beta_{l,1}, \beta_{0,2}, \beta_{1,2}, \cdots, \beta_{l,2}, \cdots\cdots, \beta_{0,J}, \beta_{1,J}, \cdots, \beta_{l,J}\right)$$

where $\alpha_i = \beta_{\mathrm{int}[(i-1)/J],\mathrm{mod}[i/J]}$.

Score vector and the information matrix are given by $U = \frac{\partial \ln L}{\partial \underline{\alpha}}$ and $J = E\left[\frac{\partial \ln L}{\partial \underline{\alpha}} \frac{\partial \ln L}{\partial \underline{\alpha}'}\right]$ respectively.

By method of scoring, we have the (m+1)th approximation of $\underline{\alpha}$, $\underline{\alpha}^{(m+1)}$ given by: $\underline{\alpha}^{(m+1)} = \underline{\alpha}^{(m)} + \left[\mathbf{J}^{(m)}\right]^{-1} \mathbf{U}^{(m)}$ where $\mathbf{J}^{(m)}$ and $\mathbf{U}^{(m)}$ are the mth approximation of **J** and **U** respectively obtained by evaluating at $\underline{\alpha} = \underline{\alpha}^{(m)}$.

CHAPTER 5: FEEDBACK CONSIDERATION IN TWO BRANDS AND MULTIPLE BRANDS PRODUCT CATEGORY AND CORRELATED SERIAL PURCHASES USING MARKOV CHAIN

What happens if we do not assume independence for the probability between choices? That is, the probability in choosing brand A on subsequent choices are not independent of prior choices. Feedback from the purchase of say, on the 1st occasion may influence his/her subsequent choice of purchase. To further illustrate this, we consider the simple situation as follows in Section 5.1. An extension to the case of more than two brands is illustrated in Section 5.2.

5.1 Feedback consideration in two brands product category

Let us consider the case where there are only two brands within a particular product category, say brand A and brand B. Suppose each decision maker among a sample of size *k*, exercises only two choices within that product category and we define:

r_A : the number of decision makers choosing brand A on their first purchase from the product category.

r_{AA} : the number of decision makers choosing brand A on both purchases from the product category.

r_{AB} : the number of decision makers choosing brand B on their second purchase and brand A on their first purchase from the product category.

r_{BA} : the number of decision makers choosing brand A on their second purchase and brand B on their first purchase from the product category.

r_B : the number of decision makers choosing brand B on their first purchase from the product category.

A_i : the event that brand A is chosen on the i^{th} occasion.

B_i : the event that brand B is chosen on the i^{th} occasion.

Then, we have the following results:

(i) $P(A_1A_2) \approx \frac{r_{AA}}{k}$

(ii) $P(A_1B_2) \approx \frac{r_{AB}}{k}$

(iii) $P(B_1A_2) \approx \frac{r_{BA}}{k}$

(iv) $P(B_1B_2) \approx \frac{r_{BB}}{k}$

Note that the sum of these probabilities equals one.

Here, if the sample is unbiased, $\frac{r_{AA}}{k}$ gives a reasonable estimate for the probability of choosing brand A twice. Compare with the case where choices are independent of each other, a value of $\left(\frac{r_A}{k}\right)^2$ gives the estimated probability of choosing brand A twice. If the value for $\frac{r_{AA}}{k} > \left(\frac{r_A}{k}\right)^2$, then it would mean that brand A's loyalty is high and vice versa.

Also, $\frac{r_{AB}}{k}$ gives quite a reasonable estimate for finding the probability of switching to brand B after choosing brand A on the first instance while $\frac{r_{BA}}{k}$ gives the estimate of switching to brand A after choosing brand B on the first instance.

Thus, $\frac{r_{AB} + r_{BA}}{k}$ gives the probability of a decision maker switching brand after the first instance.

Let's now extend this to the case when there are more than two purchases.

Using the similar notations as before, we have the following results:

(i) $P(A_1A_2A_3) \approx \frac{r_{AAA}}{k}$

(ii) $P(A_1A_2B_3) \approx \frac{r_{AA} - r_{AAA}}{k}$

(iii) $P(A_1B_2A_3) \approx \frac{r_{ABA}}{k} = \frac{r_A - r_{AA} - r_{ABB}}{k}$

(iv) $P(A_1B_2B_3) \approx \frac{r_{ABB}}{k} = \frac{r_{AB}}{k} - \frac{r_{ABA}}{k}$

(v) $P(B_1A_2A_3) \approx \frac{r_{BAA}}{k}$

(vi) $P(B_1A_2B_3) \approx \frac{r_B - r_{BB} - r_{BAA}}{k} = \frac{k - r_A - r_{BB} - r_{BAA}}{k}$

(vii) $P(B_1B_2A_3) \approx \frac{r_{BB} - r_{BBB}}{k}$

(viii) $P(B_1B_2B_3) \approx \frac{r_{BBB}}{k}$

If we were to find the sum of these probabilities, we will get it to be one.

A number of interesting deductions can be made from the above probabilities:

(i) $\frac{r_{AAA}}{k}$ and $\frac{r_{BBB}}{k}$ gives a measure of the probability of the decision maker choosing brand A and B for all three occasions. Compare with the case where choices are independent of each other, a value of $\left(\frac{r_A}{k}\right)^3$ and $\left(\frac{r_B}{k}\right)^3$ give the estimated probability of choosing brand A and B thrice respectively. If $\frac{r_{AAA}}{k} > \left(\frac{r_A}{k}\right)^3$, it would mean that there is greater level of brand loyalty for A and vice versa.

(ii) $\frac{k - r_A - r_{BB} - r_{BAA}}{k}$ and $\frac{r_A - r_{AA} - r_{ABB}}{k}$ gives a measure of the probability of the decision maker switching brands each time they purchase from the product category. A high value for these two quantities may mean that the decision maker may have chosen the brands out of convenience or at random. There is little separating the utility for each of the two brands. Compare with the case where choices are independent of each other, a value of $\left(\frac{r_A}{k}\right)^2\left(\frac{r_B}{k}\right)$ and $\left(\frac{r_A}{k}\right)\left(\frac{r_B}{k}\right)^2$ give the estimated probability of switching brands each time they purchase from the product category.

If the value of *k* chosen is relatively small, the above approach to estimating the probabilities may be less than satisfactory as small samples may be inevitably biased.

Also, if the number of brands under consideration increases, without any change in the value of *k* will result in inaccurate estimation of the probabilities as the number of decision makers making a unique set of choice will decrease, thus, also leading to the biased-ness in the computation.

This process can be rather tedious when the number of brands available increases or when the number of purchases made increases. Suppose that there are five possible brands and each decision maker is observed to make twenty purchases over time. This would lead to the consideration of 9.5 $\times 10^{13}$ possible combinations!

5.2 Feedback consideration in more than two brands product category

In the last section, we drop the assumption of independence for the probability between choices in the two-brand category? In this section, we will extend our analysis to the situation when there are more than two brands in the category. The approach will be similar to the earlier case, except that the mathematics behind the analysis will become far more complicated!

In this consideration, we assume that there are J brands within a particular product category, say brand A, B, …,J. Suppose each decision maker among a sample of size k, exercises only two choices within that product category and we use similar notations as that given in Section 5.1 as given below:

r_i : the number of decision makers choosing brand I on their first purchase from the product category.

r_{ii} : the number of decision makers choosing brand I on both purchases from the product category.

r_{ij} : the number of decision makers choosing brand I on their first purchase and brand J on their second purchase from the product category.

I_j : the event that brand I is chosen on the j^{th} occasion.

Then, we have the following results:

(i) $P(I_1 I_2) \approx \frac{r_{ii}}{k}$

(ii) $P(I_1M_2) \approx \frac{r_{im}}{k}$

Here, if the sample is randomly selected, $\sum_{i=1}^{J} \frac{r_{ii}}{k}$ gives a reasonable estimate for the probability of choosing the same brand twice. If this probability is high for all i =1, 2,…,J, as compared to the case when the choices are independent, then the brand loyalty is high across all brands and decision makers do not frequently change the brand after using it once. If there are some values of i where the value $\frac{r_{ii}}{k}$ are low and some values where the value $\frac{r_{ii}}{k}$ are high as compared to the corresponding values when the choices are independent, we would say that some brands may have been better than others, thus, being able to instil a higher level of loyalty among its customers than others.

Also, $\sum_{I=1}^{J}\sum_{\substack{M=1\\ M\neq I}}^{J} P(I_1M_2) = \sum_{i=1}^{J}\sum_{\substack{m=1\\ m\neq i}}^{J} \frac{r_i}{k}\frac{r_{im}}{r_i} = \sum_{i=1}^{J}\sum_{\substack{m=1\\ m\neq i}}^{J} \frac{r_{im}}{k}$

gives an estimate for finding the probability of switching brands for two consecutive purchases. A high value for each of the quantities $\sum_{\substack{m=1\\ m\neq i}}^{J} \frac{r_{im}}{k}$, i = 1, 2, …, J as compared to the independent case indicates that decision makers purchase a product in the same product category at random and there is little brand loyalty among the customers.

If purchases are been made independent of one another, $\sum_{i=1}^{J}\left(\frac{r_i}{k}\right)^2$ will provide an estimate for the probability of choosing the same brand twice if there is no brand loyalty. This value can then be compared with $\sum_{i=1}^{J} \frac{r_{ii}}{k}$ to find out whether there is noticeable difference between the two values, which in turn will allow us to find out whether there is any preference between the brands.

Lets us extend our discussion to the case when the decision maker exercises three choices instead of two as given above:

Using similar notations as before, we have the following results:

(i) $P(I_1I_2I_3) \approx \frac{r_{iii}}{k}$

(ii) $P(I_1I_2M_3) \approx \frac{r_{iim}}{k}$

(iii) $P(I_1M_2I_3) \approx \frac{r_{imi}}{k}$

(iv) $P(I_1M_2M_3) \approx \frac{r_{imm}}{k}$

(v) $P(I_1M_2N_3) \approx \frac{r_{imn}}{k}$

A number of interesting deductions can be made from the above probabilities:

(i) $\sum_{i=1}^{J} \frac{r_{iii}}{k}$ gives a measure of the probability of the decision maker choosing the same brand for all three occasions. A high value for $\frac{r_{iii}}{k}$, i = 1, 2,…,J as compared to the independent case would mean that there is a high opportunity cost for switching of brands. Advertising in this case may not help in encouraging decision makers to switch brands.

(ii) $\sum_{i=1}^{J} \sum_{\substack{m=1 \\ m \neq i}}^{J} \sum_{\substack{n=1 \\ n \neq m}}^{J} \frac{r_{imn}}{k}$ gives a measure of the probability of the decision maker switching brands each time they purchase from the product category. A high value for this sum may mean that the decision maker may have chosen the brands out of convenience or at random. There is little separating the utility for each of the brands in the same product category.

If purchases are made independent of each other, then we have:

P(having the same brand for each of the three purchases) = $\sum_{i=1}^{J}\left(\frac{r_i}{k}\right)^3$

P(having the same brand for two out of the three purchases) = $3\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\left(\frac{r_i}{k}\right)^2\left(\frac{r_m}{k}\right)$

P(having different brands for all three purchases) = $\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\sum_{\substack{n=1\\n\neq m}}^{J}\left(\frac{r_i}{k}\right)\left(\frac{r_m}{k}\right)\left(\frac{r_n}{k}\right)$

The above three values of probability can then be compared to $\sum_{i=1}^{J}\frac{r_{i^3}}{k}$,

$\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\frac{r_{iim}}{k}+\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\frac{r_{imi}}{k}+\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\frac{r_{mii}}{k}$ and $\sum_{i=1}^{J}\sum_{\substack{m=1\\m\neq i}}^{J}\sum_{\substack{n=1\\n\neq m}}^{J}\frac{r_{imn}}{k}$ respectively to identify any noticeable difference (if any) between the actual observations and expected observations if there are no brand preferences.

What happens if we consider the general case when the decision maker exercises p choices instead of the case for two or three choices as described above?

Using similar notations as before, we have the following useful results:

(i) $P(I_1 I_2 \ldots I_p) = P(I_1)P(I_2 I_3 \ldots I_p \mid I_1)$

$= P(I_1)\, P(I_2 \mid I_1)P(I_3 \mid I_1 I_2)\ldots P(I_p \mid I_1 I_2 \ldots I_{p-1}) \approx \frac{r_{i^p}}{k}$

(ii) P(one of the choices made among p choices is brand M and the rest are brand I)

=

$\sum_{j=1}^{p}$P(brand M is selected on the j^{th} occasion and brand I is selected on the rest of the occasions)

$= \frac{r_m}{k}\frac{r_{mi}}{r_m}\frac{r_{mi^2}}{r_{mi}}\cdots\frac{r_{mi^{p-1}}}{r_{mi^{p-2}}}+\frac{r_i}{k}\frac{r_{im}}{r_i}\frac{r_{imi}}{r_{im}}\cdots\frac{r_{imi^{p-2}}}{r_{imi^{p-3}}}+\ldots+\frac{r_i}{k}\frac{r_{i^2}}{r_i}\frac{r_{i^3}}{r_{i^2}}\cdots\frac{r_{i^{p-1}m}}{r_{i^{p-1}}}$

$$= \frac{r_{mi^{p-1}}}{k} + \frac{r_{imi^{p-2}}}{k} + \frac{r_{i^2mi^{p-3}}}{k} + \ldots \frac{r_{i^{p-1}m}}{k}$$

(iii) P(only two brands are considered on all the *p* occasions)

= P(brand M is selected on *s* occasions while brand I is selected on *p*−*s* occasions)

=

$$\frac{r_{m^s i^{p-s}}}{k} + \left(\frac{r_{m^{s-1} imi^{p-s-1}}}{k} + \frac{r_{m^{s-1} i^2 mi^{p-s-2}}}{k} + \ldots \frac{r_{m^{s-1} i^{p-s} m}}{k} \right) + \left(\frac{r_{m^{s-2} im^2 i^{p-s-1}}}{k} + \frac{r_{m^{s-2} imimi^{p-s-2}}}{k} \ldots + \frac{r_{m^{s-2} i^{p-s} m^2}}{k} \right) +$$

$$\ldots + \left(\frac{r_{mim^{s-1} i^{p-s-1}}}{k} + \frac{r_{mi^2 m^{s-1} i^{p-s-2}}}{k} + \ldots + \frac{r_{mi^{p-s} m^{s-1}}}{k} \right) + \frac{r_{i^{p-s} m^s}}{k}$$

Note that there are altogether $\frac{p!}{s!(p-s)!}$ terms in the above probability.

(iv) P(brand I is selected on *p* – *s* occasions while brands other than brand I is selected on *s* occasions)

=

$$\left(\frac{r_{\bar{i}^s i^{p-s}}}{k} \right) + \left(\frac{r_{\bar{i}^{s-1} i\bar{i} i^{p-s-1}}}{k} + \frac{r_{\bar{i}^{s-1} i^2 \bar{i} i^{p-s-2}}}{k} + \ldots + \frac{r_{\bar{i}^{s-1} i^{p-s} \bar{i}}}{k} \right) + \left(\frac{r_{\bar{i}^{s-2} i\bar{i}^2 i^{p-s-1}}}{k} + \frac{r_{\bar{i}^{s-2} i\bar{i} i\bar{i} i^{p-s-2}}}{k} + \ldots + \frac{r_{\bar{i}^{s-2} i^{p-s} \bar{i}^2}}{k} \right) +$$

$$\ldots + \left(\frac{r_{\bar{i} i\bar{i}^{s-1} i^{p-s-1}}}{k} + \frac{r_{\bar{i} i^2 \bar{i}^{s-1} i^{p-s-2}}}{k} + \ldots + \frac{r_{\bar{i} i^{p-s} \bar{i}^{s-1}}}{k} \right) + \frac{r_{i^{p-s} \bar{i}^s}}{k}$$

We conclude this section by giving a numerical example for the three-brand problem (J = 3). Let's name the brands I, M and N. Suppose also that each decision maker performs three choices (*p* = 3).

We summarize the choices made by a sample of 500 decision makers using the following table:

Table 6: Observations based on a sample of 500 decision makers choosing among three brands

	I_2			M_2			N_2		
	I_3	M_3	N_3	I_3	M_3	N_3	I_3	M_3	N_3
I_1	90	20	40	25	10	5	30	15	15
M_1	10	15	5	20	35	30	15	10	10
N_1	10	10	10	10	10	10	10	10	20

So, we obtain the following results:

P(choosing brand I for all three purchases) = $\frac{9}{50}$

P(choosing brand M for all three purchases) = $\frac{7}{100}$

P(choosing brand N for all three purchases) = $\frac{1}{25}$

and P(switching brands each time he makes a purchase)

$$= \left(\frac{30}{500}\right)+\left(\frac{45}{500}\right)+\left(\frac{20}{500}\right)+\left(\frac{25}{500}\right)+\left(\frac{20}{500}\right)+\left(\frac{20}{500}\right)=\frac{8}{25}$$

<u>Comments on the brand loyalty for Brand I, M and N</u>

It is noted that $P(I_1I_2I_3) = 0.18$, $P(I_1) = 0.5$, $P(I_2) = 0.42$ and $P(I_3) = 0.44$. Hence, $P(I_1)P(I_2)P(I_3) = 0.0924$ if the choices are chosen independently. Since 0.18 is almost double the probability of having brand I in all 3 purchases if the purchases are independent, therefore it seems that there exist some degrees of brand loyalty in Brand I.

It is also noted that $P(M_1M_2M_3) = 0.18$, $P(M_1) = 0.294$, $P(M_2) = 0.31$ and $P(M_3) = 0.27$. Hence, $P(M_1)P(M_2)P(M_3) = 0.0246$ if the choices are chosen independently. Since 0.07 is almost twice the probability of having brand M in all 3 purchases if the purchases are independent, therefore it seems that there exist some degrees of brand loyalty in Brand M.

It is also noted that $P(N_1N_2N_3) = 0.04$, $P(N_1) = 0.200$, $P(N_2) = 0.27$ and $P(N_3) = 0.29$. Hence, $P(N_1)P(N_2)P(N_3) = 0.0157$ if the choices are chosen independently. Since 0.04 is more than twice the probability of having brand N in all 3 purchases if the purchases are independent, therefore it seems that there exist some degrees of brand loyalty in Brand N.

5.3 Correlated Serial Purchases using Markov Chain

A Markov chain (MC) is a probabilistic technique used to represent correlations between successive observations of a random variable (Berchtold 2001). This sequence analysis technique is a form of time-series modeling, and was introduced at the beginning of the 20th century by Andrej Andreevic Markov. It is used in many disciplines, including meteorology, geography, biology, chemistry, physics, social sciences and music. In marketing, it has already been successfully applied in modeling purchases of financial services (Prinzie and Van den Poel, 2006), predicting website purchases using clickstream data (Montgomery, 2004) or predicting software performance (Bai et al., 2005, Durand and Gaudoin, 2005). One clear advantage of using Markov chain to model the probability of purchase is that it takes into consideration that purchases are not independent of each other.

Consider a two-state Markov Chain, with transition probability

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array}$$

with $q \neq 1-p$ and p ~ Beta(α_1, β_1) and q ~ Beta(α_2, β_2), p and q are independent.

The stationary distribution can be obtained by solving the equation

$$\begin{pmatrix} \pi_A & \pi_B \end{pmatrix} \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} = \begin{pmatrix} \pi_A & \pi_B \end{pmatrix}, \text{ where } \pi_A + \pi_B = 1$$

At the stationary state, the market share of A, $\mu_A = \pi_A = E\left(\frac{Q}{P+Q}\right)$ and that for B, $\mu_B = \pi_B = E\left(\frac{P}{P+Q}\right)$.

Remark:

The market share for A to some extent depends on the brand loyalty measure, *Q*, of B. However the link between the market share of A and the brand loyalty of the other brand may be reduced in the case of three or more brands. (See the discussion on the three brand case on p66 to p68.)

Using the Taylor's expansion, we know that

$$
\begin{aligned}
g(x,y) = g(\mu_x,\mu_y) &+ (x-\mu_x)\left.\frac{\partial g(x,y)}{\partial x}\right|_{(x,y)=(\mu_x,\mu_y)} + (y-\mu_y)\left.\frac{\partial g(x,y)}{\partial y}\right|_{(x,y)=(\mu_x,\mu_y)} \\
&+ \frac{(x-\mu_x)^2}{2}\left.\frac{\partial^2 g(x,y)}{\partial x^2}\right|_{(x,y)=(\mu_x,\mu_y)} + \frac{(y-\mu_y)^2}{2}\left.\frac{\partial^2 g(x,y)}{\partial y^2}\right|_{(x,y)=(\mu_x,\mu_y)} \\
&+ \frac{2(x-\mu_x)(y-\mu_y)}{2}\left.\frac{\partial g(x,y)}{\partial x}\right|_{(x,y)=(\mu_x,\mu_y)}\left.\frac{\partial g(x,y)}{\partial y}\right|_{(x,y)=(\mu_x,\mu_y)} + R
\end{aligned}
$$

Take $g(p,q) = \frac{q}{q+p}$ and take expectation,

$$
\begin{aligned}
E\left[\frac{Q}{Q+P}\right] &\approx \frac{E[Q]}{E[Q]+E[P]} + \frac{\mathrm{Var}[Q]E[P]}{\left(E[P]+E[Q]\right)^3} - \frac{\mathrm{Var}[P]E[Q]}{\left(E[P]+E[Q]\right)^3} \\
&\approx \frac{\frac{\alpha_2}{(\alpha_2+\beta_2)}}{\frac{\alpha_2}{(\alpha_2+\beta_2)}+\frac{\alpha_1}{(\alpha_1+\beta_1)}} + \frac{\frac{\alpha_1}{(\alpha_1+\beta_1)}\frac{\alpha_2}{(\alpha_2+\beta_2)}\frac{\beta_2}{(\alpha_2+\beta_2)}\frac{1}{1+(\alpha_2+\beta_2)}}{\left(\frac{\alpha_2}{(\alpha_2+\beta_2)}+\frac{\alpha_1}{(\alpha_1+\beta_1)}\right)^3} \\
&\quad - \frac{\frac{\alpha_2}{(\alpha_2+\beta_2)}\frac{\alpha_1}{(\alpha_1+\beta_1)}\frac{\beta_1}{(\alpha_1+\beta_1)}\frac{1}{1+(\alpha_1+\beta_1)}}{\left(\frac{\alpha_2}{(\alpha_2+\beta_2)}+\frac{\alpha_1}{(\alpha_1+\beta_1)}\right)^3}
\end{aligned}
$$

The brand loyalty measures for A and B may then be defined as 1 – E[probability of switching brand from A] = $1-E(P)$ and 1 – E[probability of switching brand from B] = $1-E(Q)$ respectively. In this case, the market share for A to some extent depends on the brand loyalty measure, *Q*, of B.

Let *W* = P (Buying A) = $\frac{q}{q+p}$.

Purchase consistency over the population of decision makers can then be defined as Var(*W*) = $Var\left(\frac{Q}{Q+P}\right)$

$$\approx Var(Q)\frac{E^2(P)}{\left(E(P)+E(Q)\right)^4}+Var(P)\frac{E^2(Q)}{\left(E(P)+E(Q)\right)^4}$$

$$=\frac{Var(Q)E^2(P)+Var(P)E^2(Q)}{\left(E(P)+E(Q)\right)^4}$$

$$=\frac{\frac{\alpha_2}{(\alpha_2+\beta_2)}\frac{\beta_2}{(\alpha_2+\beta_2)}\frac{1}{1+(\alpha_2+\beta_2)}\frac{\alpha_1^2}{(\alpha_1+\beta_1)^2}+\frac{\alpha_1}{(\alpha_1+\beta_1)}\frac{\beta_1}{(\alpha_1+\beta_1)}\frac{1}{1+(\alpha_1+\beta_1)}\frac{\alpha_2^2}{(\alpha_2+\beta_2)^2}}{\left(\frac{\alpha_1}{(\alpha_1+\beta_1)}+\frac{\alpha_2}{(\alpha_2+\beta_2)}\right)^4}$$

The values of brand loyalty defined in this way are not the same for different brands, which is a much better approach as compared to the use of purchase consistency discussed earlier in Section 2.5.

The parameters can be estimated using method of moments or maximum likelihood method as discussed in Section 2.6.

In the distribution for P and Q are not assumed to follow the beta distribution (non-parametric case), we can still estimate the brand loyalty, market share and purchase consistency in the following way:

We estimate the value of P_j *for* j^{th} individual with $\frac{n_{AB}}{n_A}$ where n_{AB} is the number of pairs of purchases where brand A is purchased on the i^{th} occasion and brand B is purchased on the (i+1)th occasion by the j^{th} individual and n_A is the number of times (excluding the last purchase) where A is being purchased by the j^{th} individual. For example if an individual purchases ABBBAABABAABAAB, then n_{AB} = 5 and n_A = 8 and hence, $\hat{p}_j \approx \frac{5}{8}$. The estimate for E(P) is obtained by taking the average of $\hat{p}_j$ over all the individuals. In the same way, E(Q) can be estimated.

Var (P) is then calculated by $\frac{1}{k-1}\sum_{j=1}^{k}\left(\hat{p}_j - E(P)\right)^2$ with E(P) estimated from above. Var(Q) is calculated in the same way as well.

Suppose that P(B|A) = P(B|B) (i.e. the i.i.d. case), then the transitional matrix will be given by:

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} \end{array}$$

Suppose also that we want to test the hypothesis

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix} \end{array} \text{against} \begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} A \\ B \end{array} & \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \end{array}.$$

In other words, we are testing $P = 1 - Q$.

Let's refer to the example on p24 regarding the 30 rides for 20 customers. Let assume that on top of the information about the number of rides out of 30 rides using Comfort, we also now the sequence of rides of these 30 customers. They are given in the following table.

Table 7: Dataset on choice of taxi company

Individual 1	0 1 1 1 0 0 1 0 1 0 0 0 1 0 0 1 1 1 0 0 1 1 0 1 1 1 1 1 1 1
Individual 2	1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 0 0 1 1 1 1 0 0 1 1 1
Individual 3	0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 0 0 0 1 0
Individual 4	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 0 0 0
Individual 5	1 1 1 1 1 0 0 0 0 0 0 0 1 0 1 1 1 1 1 0 1 1 0 1 0 0 0 1 1 1
Individual 6	1 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 1 1 1 0 0 1 0 1 0 1 0 0 0 0
Individual 7	1 1 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 1 1 0 1 0 0 0 0 0 1 1 0 1
Individual 8	1 1 1 0 0 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1
Individual 9	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 1 1 1 1 0 1 1 1 1
Individual 10	1 1 0 0 0 1 1 0 1 0 1 1 1 1 0 1 1 1 1 1 1 1 0 0 0 1 1 0 1 1
Individual 11	0 1 0 0 1 1 1 0 1 0 1 0 0 1 0 1 0 0 0 0 1 1 0 1 1 0 0 1 1 1
Individual 12	1 0 1 0 1 1 0 1 1 1 1 1 1 0 0 1 1 1 0 1 1 0 0 0 1 0 1 0 0 1
Individual 13	0 1 0 1 1 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 0 1 0 1 1 1 0 0 0
Individual 14	0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 1 0 1 0 0 0 0 1 1 1 0 0 1 0 1
Individual 15	0 0 0 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 1 1 1 1 0 1 1 0 1 0 1 0
Individual 16	1 1 1 0 1 1 1 1 0 1 1 1 1 1 1 0 1 1 1 1 0 1 1 0 1 1 1 1 1 1
Individual 17	1 1 1 1 0 1 0 1 0 1 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 0 0 1 0 0
Individual 18	1 1 1 1 1 0 0 1 0 1 1 1 1 1 0 0 1 0 1 0 1 0 0 0 0 1 0 1 1 1
Individual 19	1 1 1 1 0 1 1 1 1 1 1 1 0 0 1 1 1 1 1 1 1 1 0 1 1 1 0 1 1 1
Individual 20	1 1 1 1 0 1 0 1 1 1 1 0 0 1 1 1 1 1 1 1 0 1 1 0 1 0 1 1 1 1

where 1 stands for a ride on a Comfort taxi and 0 stands for a ride on a TIBS taxi.

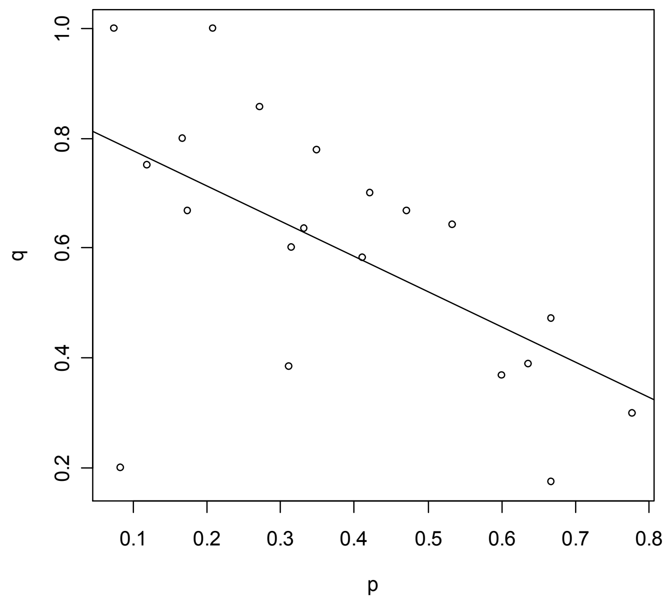
Let us look at the Individual 1's ride pattern. There are altogether 6 pairs of (1, 0) out of 17 pairs of (1, *). Hence *P* for Individual 1 is estimated by $\frac{6}{18}$ = 0.3333. Similarly there are 7 pairs of (0, 1) out of 11 pairs of (0, *). Hence *Q* for Individual 1 is estimated by $\frac{7}{11}$ = 0.6364. Therefore for Individual 1, estimate of (*P*, *Q*) is given by (0.333, 0.636).

The following are the 20 observed values of (*P*, *Q*) from the above data set

Table 8: Table showing the estimates of (P,Q) for the 20 individuals

Individual	Estimate of (P,Q)
1	(0.333,0.636)
2	(0.174,0.667)
3	(0.667,0.174)
4	(0.0833,0.2)
5	(0.312,0.385)
6	(0.778,0.3)
7	(0.636,0.389)
8	(0.12,0.75)
9	(0.0741,1)
10	(0.316,0.6)
11	(0.533,0.643)
12	(0.471,0.667)
13	(0.35,0.778)
14	(0.6,0.368)
15	(0.667,0.471)
16	(0.208,1)
17	(0.421,0.7)
18	(0.412,0.583)
19	(0.167,0.8)
20	(0.273,0.857)

The plot of the 20 points is given below:



We want to test H_0: $\beta = -1$ against H_1: $\beta \neq -1$. The test is equivalent to test H_0: $P = 1 - Q$ against H_1: $P \neq 1 - Q$.

From the data, we have $\hat{\beta} = -0.6447$ and s.e.($\hat{\beta}$) = 0.2181. Hence the test statistic is given by $t = \dfrac{\hat{\beta} - (-1)}{s.e.(\hat{\beta})} = \dfrac{-0.6447 + 1}{0.2181} = 1.629$. Since the observed p-value = 0.06, we do not reject H_0 and conclude that there is no significant evidence to show that the i.i.d model does not fit the data. Of course, we made the normality assumption for the above test.

Let us extend our discussion to a more than three-brand situation. Consider the following transition matrix with the probability of switching brand evenly divided among the other brands:

$$\begin{array}{cc} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{c} A \\[2ex] B \\[2ex] C \end{array} & \begin{pmatrix} a & \frac{1-a}{2} & \frac{1-a}{2} \\ \frac{1-b}{2} & b & \frac{1-b}{2} \\ \frac{1-c}{2} & \frac{1-c}{2} & c \end{pmatrix} \end{array}$$

Then the stationary distribution is given by

$$\pi_A = \frac{(1-b)(1-c)}{(1-a)(1-b)+(1-a)(1-c)+(1-b)(1-c)},$$

$$\pi_B = \frac{(1-a)(1-c)}{(1-a)(1-b)+(1-a)(1-c)+(1-b)(1-c)} \text{ and}$$

$$\pi_C = \frac{(1-a)(1-b)}{(1-a)(1-b)+(1-a)(1-c)+(1-b)(1-c)}.$$

Following the same idea as the two-brand problem, the brand loyalty is defined for A as E[*A*], where *A* is the transition probability from *A* to *A*.

Remarks:

1. The market share will be the same if all three brands have the same brand loyalty.
2. If $\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} 0.9 & 0.05 & 0.05 \end{pmatrix}$, then $\begin{pmatrix} \pi_A & \pi_B & \pi_C \end{pmatrix} = \begin{pmatrix} 0.826 & 0.087 & 0.087 \end{pmatrix}$. If $\begin{pmatrix} a & b & c \end{pmatrix} = \begin{pmatrix} 0.9 & 0.5 & 0.5 \end{pmatrix}$, then $\begin{pmatrix} \pi_A & \pi_B & \pi_C \end{pmatrix} = \begin{pmatrix} 0.7142 & 0.1429 & 0.1429 \end{pmatrix}$.

   Hence the brand with higher brand loyalty will have a larger market share.

Next we consider the case where customers of brand A have high brand loyalty, while the customers of the other two brands have a very large probability of switching between the two brands. The following transition matrix reflects such scenario.

$$\begin{array}{c} \\ A \\ B \\ C \end{array} \begin{array}{c} \begin{array}{ccc} A & B & C \end{array} \\ \begin{pmatrix} a & \frac{1-a}{2} & \frac{1-a}{2} \\ 1-b-d & b & d \\ 1-b-d & d & b \end{pmatrix} \end{array}$$

Here we assume that the brand loyalty for both Brands B and C are the same. We also assume that *d* > *b* to reflect the large probability of switching among the two brands.

The stationary distribution of the above transition matrix is given by

$$\pi_A = \frac{b+d-1}{a+b+d-2}$$

$$\pi_B = \frac{a-1}{2(a+b+d-2)} \text{ and}$$

$$\pi_C = \frac{a-1}{2(a+b+d-2)}$$

Remarks:

1. If (a, b, d) = (0.9, 0.05, 0.9), then (π_A, π_B, π_C) = (0.333, 0.333, 0.333). Though Brand A has a high brand loyalty, it does not have a very large market share.

This is due to the fact that the customers of the other two brands keep switching among the other two brands and have a very low probability of switching to Brand A.

2. If (a, b, d) = (0.9, 0.4, 0.55), then (π_A, π_B, π_C) = (0.333, 0.333, 0.333). In this case, the probability of switching among Brands B and C is not very large, while the probability of the customers of Brands B and C switching to Brand A is very small (0.05). Hence even though Brand A has high brand loyalty and it does not have a big market share.
3. If (a, b, d) = (0.9, 0.4, 0.45), then (π_A, π_B, π_C) = (0.6, 0.2, 0.2). In this case, the probability of customers of Brands B or C switching to Brand A is moderate (0.15). The moderate probability of switching from other brands to Brand A and the high brand loyalty jointly help Brand A to have a bigger market share. We can see that the market share of a brand does not totally depend on its brand loyalty.

CHAPTER 6 NEW APPROACHES TO THE MEASURING OF BRAND LOYALTY FOR SEMI-PARAMETRIC AND NON-PARAMETRIC MODELS

In this section, I will be looking at some innovative ways of measuring brand loyalty if we drop the assumptions for the distribution of purchase probabilities or if we drop the assumption that purchase decisions for each individual are independent.

6.1 Measure of Brand Loyalty using Brand-oriented behavioural measures (C) based on behavioural perspective (from the perspective of recent purchases)

6.1.1 Two brands product category

This method of measuring brand loyalty is a modified version of what is common in the marketing literature, which measures brand loyalty using the number of strings of 3 consecutive purchases. In this method introduced by me, I only consider recent purchases made by consumers as of great importance to the brand managers to measuring brand loyalty at current situation.

Let me approach this method with a simple example. Suppose there are only two brands in the market of a particular product. The market share for Brand A is 0.7 and that of brand B is 0.3. We carry out a sample of say 30 consumers and study their choices made over 6 purchases. We consider in this case, that by recent purchases, it would mean to be the last three purchases.

Then, under the assumption of independence of the purchases,

P(selecting brand A for the last three occasions), $p_A{}^R = (0.7)^3 = 0.343$ and

P(selecting brand B for the last three occasions), $p_B{}^R = (0.3)^3 = 0.027$

We may then proceed to test the hypothesis that:

H_0 : $p_A{}^R$ = 0.343

against

H_1 : $p_A{}^R \neq$ 0.343 at 5% level of significance.

Test statistic $Z = \dfrac{\hat{P}_A^R - 0.343}{\sqrt{\dfrac{0.343(1-0.343)}{30}}} \sim N(0,1)$ under H_0

Suppose there are 10 consumers out of 30 who bought Brand A consecutively for the last three purchases and 5 consumers out of 30 who bought Brand B consecutively for the last three purchases, then, the value of Z is given by:

$$z = \frac{\frac{10}{30} - 0.343}{\sqrt{\frac{0.343(1-0.343)}{30}}} = -0.112$$

Since z = – 0.112 > –1.96, we do not reject H_0 and conclude that there is no brand loyalty for Brand A at 5% level of significance.

Lets us look at Brand B:

H'_0 : $p_B{}^R$ = 0.027

against

H'_1 : $p_B{}^R \neq$ 0.027 at 5% level of significance.

X ~ Bin (30,0.027)

Since P(X $\geq$ 5) = 0.0012 < 0.025, we reject H'_0 and conclude that there is a significant brand loyalty for Brand B at 5% level of significance.

Of course, this method may have its criticism as it completely ignores the information that is obtained about the purchase pattern other than the recent three purchases. However, this method does have its merits especially in situations where the time lag between each purchase is significantly long and individuals may not be able to remember the purchase patterns other than the three most recent purchases.

6.1.2 More than two brands product category

Let us extend the idea given in Section 6.1.1 to the situation when there are more than two brands in the same product category. In a similar way, we can obtain a measure of brand loyalty if we define brand loyalty to a particular brand I to mean choosing brand I successively for the more recent purchases.

As an example, suppose there are 6 choices being made from a group of 3 brands (A, B and C) by an individual. A sample of 100 decision makers is taken to study the choices made. A table showing the choices made by the first 20 decision makers are given below:

Table 9: Observations based on the first 20 decision makers choosing among three brands

Individual	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1st purchase	B	C	A	B	C	C	A	C	A	C	B	A	C	B	C	B	B	B	C	B
2nd purchase	B	B	A	C	B	B	C	B	A	C	A	A	B	C	A	C	C	C	C	B
3rd purchase	C	B	A	A	A	B	C	B	A	A	B	A	B	B	A	C	A	C	B	B
4th purchase	A	B	A	B	A	B	C	C	A	C	C	A	A	B	A	B	A	C	C	B
5th purchase	A	C	A	C	A	C	C	C	A	B	C	A	B	C	A	B	A	B	C	B
6th purchase	A	C	A	C	A	C	C	B	A	B	A	A	B	C	A	B	A	B	C	B

In this example, there are altogether 600 purchases being made on the product and they are equally distributed across the three brands (equal market share for each of the three brands).

However, If we define brand loyalty to a particular brand (say A) to mean choosing brand A for all the last 3 choices, we obtain an estimate for the brand loyalty to be as follows:

P(choosing brand A for all the last 3 choices) = $\frac{7}{20}$ while that for brand B and C to be

P(choosing brand B for all the last 3 choices) = P(choosing brand C for all the last 3 choices) = $\frac{1}{10}$.

We may next proceed to test the hypothesis whether there is any difference between the brand loyalty for A against that for brand B.

To test H_0: P(choosing brand A for all the last 3 choices) = P(choosing brand B for all the last 3 choices)

against

H_1: P(choosing brand A for all the last 3 choices) > P(choosing brand B for all the last 3 choices) at 5% level of significance

Test statistic Z = $\dfrac{\hat{P}_1 - \hat{P}_2}{\sqrt{\dfrac{\frac{1}{27}\left(\frac{26}{27}\right)}{100} + \dfrac{\frac{1}{27}\left(\frac{26}{27}\right)}{100} + 2\dfrac{\frac{1}{27}\left(\frac{1}{27}\right)}{100}}}$ ~ N(0,1) under H_0.

Value of test statistic, z = $\dfrac{0.35 - 0.10}{\sqrt{\dfrac{54}{72900}}} = 9.19$

Since P(Z>9.19) < 0.05, we reject H_0 and conclude that the brand loyalty for A is greater than that for B at 5% level of significance.

6.2 Measure of Brand Loyalty using Brand-oriented behavioural measures (C) based on behavioural perspective (from the perspective of recent purchases but with weights being attached to the order of purchase)

We can refine the method given in 8.1 by assigning different weights according to the order of purchase, by having greater weights for the more recent purchases. In this method, we look through the last three purchase patterns of the sample of consumers and assign ω_1, ω_2 and ω_3 to the most recent purchase, the 2nd most recent purchase and the 3rd most recent purchase respectively. Of course, we choose these weights in such a manner where $\omega_1 > \omega_2 > \omega_3$.

If there are for example, three brands in the product category, i.e. either brand A, brand B or brand C is to be chosen, then we give a score to each of these brands based on the sum of the weights attached to the order of purchase. We will also deduct the score if the brand in question is not chosen at any point of purchase to penalize the brand for the consumer's disloyalty to the brand.

Lets consider a simple example as shown below:

Table 10: Observations based on 12 individuals in a three-brand problem

Individual	1	2	3	4	5	6	7	8	9	10	11	12
1st purchase	A	B		A	A	B	A	C	C			
2nd purchase	A	B		A	B	A	B	B	A	A		
3rd purchase	A	B	A	B	A	A	C	A	B	B	B	

The score obtained for brand A, B and C contributed by each of the twelve individuals is as follows:

Table 11: Score obtained for Brand A, B and C Based on Purchase History

Individual	1	2	3	4
Brand A	$\omega_1 + \omega_2 + \omega_3$	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	ω_1	$-0.5\ \omega_1 + \omega_2 + \omega_3$
Brand B	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	$\omega_1 + \omega_2 + \omega_3$	$-0.5\ \omega_1$	$\omega_1 - 0.5\ \omega_2 - 0.5\ \omega_3$
Brand C	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	$-0.5\ \omega_1$	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$

Table 11 (Continued)

Individual	5	6	7
Brand A	$\omega_1 - 0.5\ \omega_2 + \omega_3$	$\omega_1 + \omega_2 - 0.5\ \omega_3$	$-0.5\ \omega_1 - 0.5\ \omega_2 + \omega_3$
Brand B	$-0.5\ \omega_1 + \omega_2 - 0.5\omega_3$	$-0.5\ \omega_1 - 0.5\ \omega_2 + \omega_3$	$-0.5\ \omega_1 + \omega_2 - 0.5\ \omega_3$
Brand C	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	$-0.5\ (\omega_1 + \omega_2 + \omega_3)$	$\omega_1 - 0.5\ \omega_2 - 0.5\ \omega_3$

Table 11 (Continued)

Individual	8	9	10
Brand A	$\omega_1 - 0.5\ \omega_2 - 0.5\ \omega_3$	$-0.5\ \omega_1 + \omega_2 - 0.5\ \omega_3$	$-0.5\ \omega_1 + \omega_2$
Brand B	$-0.5\ \omega_1 + \omega_2 - 0.5\ \omega_3$	$\omega_1 - 0.5\ \omega_2 - 0.5\ \omega_3$	$\omega_1 - 0.5\ \omega_2$
Brand C	$-0.5\ \omega_1 - 0.5\ \omega_2 + \omega_3$	$-0.5\ \omega_1 - 0.5\ \omega_2 + \omega_3$	$-0.5\ \omega_1 - 0.5\ \omega_2$

Table 11 (Continued)

Individual	11	12
Brand A	$-0.5\ \omega_1$	
Brand B	ω_1	
Brand C	$-0.5\ \omega_1$	

Thus, we have the following analysis:

average score for brand A = $\frac{1}{12}\left(2\omega_1 + 3\omega_2 + 2\omega_3\right)$

average score for brand B = $\frac{1}{12}\left(2\omega_1 + \frac{3}{2}\omega_2 - \omega_3\right)$

average score for brand C = $\frac{1}{12}\left(-4\omega_1 - \frac{9}{2}\omega_2 - \omega_3\right)$

Suppose we assign the values of ω_1, ω_2 and ω_3 as follows:

$\omega_1 = 0.5$

$\omega_2 = 0.3$

$\omega_3 = 0.1$

Then,

average score for brand A = 2.1

average score for brand B = 1.35

average score for brand C = –3.45

In conclusion, brand A is the preferred brand among the consumers and its measure of brand loyalty is the highest. On the other hand, brand C has the lowest measure for brand loyalty. However, as the weights are chosen arbitrarily, it may lead to different conclusion if another set of weights are used. Thus, there should be some good reasons for the experts to choose a particular set of weights over another to ensure that there is greater agreement on the brand with the highest loyalty score.

6.3 New Measure of Brand Loyalty using a mixture of Brand-oriented attitude measures (A) and Brand-oriented behavioural measures (C) based on both behavioural and attitudinal perspective)

Two stages will be employed in this measure:

A randomly selected group of *N* consumers who purchase a product within a particular product category, comprising of *k* brands, is considered in the study.

Stage 1:

Get them to write down the last three purchases (if any) in the product category. We use the same example as the one given in 8.2 to illustrate this new approach. In that study, there are 12 consumers (*N* = 12) in the study and there are 3 brands (A, B, C and *k* = 3).

Referring to Table 8, we can then give a score according to both the position and their choice of purchase.

In computing the brand loyalty for A, we give s_3 points if A is chosen on 3rd purchase, s_2 points if A is chosen on 2nd purchase and s_1 point if A is chosen on 1st purchase.

If A is not chosen on 3rd purchase, $\frac{1}{2}s_3$ points are deducted (as it can be brand B or C that is chosen), if A is not chosen on 2nd purchase, $\frac{1}{2}s_2$ points are deducted (as it can be brand B or C) and if A is not chosen on 1st purchase, $\frac{1}{2}s_1$ points are deducted (as it can be brand B or C). No points will be awarded if there is no purchase pattern history.

Based on this point system, we compute the total score for each individual as given in the table below:

Table 12: Score for Brand A, B and C based on Stage 1, with weights attached

Individual	1	2	3	4
Score for brand A	$s_1+s_2+s_3$	$-\frac{1}{2}(s_1+s_2+s_3)$	s_3	$s_1+s_2-\frac{1}{2}s_3$
Score for brand B	$-\frac{1}{2}(s_1+s_2+s_3)$	$s_1+s_2+s_3$	$-\frac{1}{2}s_3$	$-\frac{1}{2}s_1-\frac{1}{2}s_2+s_3$
Score for brand C	$-\frac{1}{2}(s_1+s_2+s_3)$	$-\frac{1}{2}(s_1+s_2+s_3)$	$-\frac{1}{2}s_3$	$-\frac{1}{2}(s_1+s_2+s_3)$

Table 12: (Continued)

Individual	5	6	7	8
Score for brand A	$s_1-\frac{1}{2}s_2+s_3$	$-\frac{1}{2}s_1+s_2+s_3$	$s_1-\frac{1}{2}s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_1-\frac{1}{2}s_2+s_3$
Score for brand B	$-\frac{1}{2}s_1+s_2-\frac{1}{2}s_3$	$s_1-\frac{1}{2}s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_1+s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_1+s_2-\frac{1}{2}s_3$
Score for brand C	$-\frac{1}{2}(s_1+s_2+s_3)$	$-\frac{1}{2}(s_1+s_2+s_3)$	$-\frac{1}{2}s_1-\frac{1}{2}s_2+s_3$	$s_1-\frac{1}{2}s_2-\frac{1}{2}s_3$

Table 12: (Continued)

Individual	9	10	11	12
Score for brand A	$-\frac{1}{2}s_1+s_2-\frac{1}{2}s_3$	$s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_3$	0
Score for brand B	$-\frac{1}{2}s_1-\frac{1}{2}s_2+s_3$	$-\frac{1}{2}s_2+s_3$	s_3	0
Score for brand C	$s_1-\frac{1}{2}s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_2-\frac{1}{2}s_3$	$-\frac{1}{2}s_3$	0

Let's suppose that we assign '3' to s_3, '2' to s_2 and '1' to s_1

Based on this point system, we compute the score for each individual as given in the table below:

Table 13: Score for Brand A, B and C based on Stage 1

Individual	1	2	3	4	5	6	7	8	9	10	11	12
Score for brand A	6.0	–3.0	3.0	1.5	3.0	4.5	–1.5	1.5	0	0.5	–1.5	0
Score for brand B	–3.0	6.0	–1.5	1.5	0	–1.5	0	0	1.5	2.0	3.0	0
Score for brand C	–3.0	–3.0	–1.5	–3.0	–3	–3.0	1.5	–1.5	–1.5	–2.5	–1.5	0

From the table above, one can see that individual 1 is the most loyal while individual 2 is the least loyal for brand A. For brand B, individual 2 is the most loyal customer while individual 1 is the least loyal customer. For brand C, individual 7 is the most loyal while individual 1, 2, 4, 5 and 6 are the least loyal customer.

The average score of the group of 12 individuals for brand A, B and C are respectively 1.17, 0.67 and –1.83. The score for the group ranges from –3 to 6. A positive value for the average score of a particular brand indicates that the sample chosen is brand loyal to the brand while a negative value for the average score of a particular brand indicates that the sample chosen is not brand loyal to the brand in question.

We can also group brand A and B as the two most purchased brands based on the group's historical data of purchase and conclude that the sample chosen is loyal to both brand A and B.

Stage 2

This stage looks at the choice set under consideration as well as the product features that make them attractive to the individuals at the next purchase. The consumers are first asked to name the top three product features that are important to them for making a decision as to which brand they will buy. Some of these features include price, availability and durability. It would also be useful to list down in a form of a table, all the features that are mentioned by the individuals and indicate how many of these individuals had chosen each of the features as the most important, 2nd most important and the 3rd most important for their consideration. This table will be useful for brand managers when they decide on the feature that they can improve upon in their product so as to increase the consumers' loyalty to their brand.

Table 14 below shows the current ranking of each brand based on the top three product features made by the 12 individuals.

Table 14: Current Ranking of the Brands based on the Product Features

Individual	1	2	3	4	5	6
Choice set	{A}	{B}	{A, B}	{A, B, C}	{A, B }	{ B, C }
Brand(s) with most important feature	A	B	A	A, B	B	B, C
Brand(s) with 2nd most important feature	A	B	A, B	B	A	B
Brand(s) with 3rd most important feature	A	B	B	A, C	A	C

Table 14: (Continued)

Individual	7	8	9	10	11	12
Choice set	{A, B, C}	{A, B }	{ B, C}	{A, C}	{ B }	{A, B, C}
Brand(s) with most important feature	C	B	B, C	A	B	A, B, C
Brand(s) with 2nd most important feature	A, C	A	B, C	A	B	A, B
Brand(s) with 3rd most important feature	B, C	A, B	B	C	B	B, C

We will then construct a scoring system for each brand and each individual, depending on the size of the choice set as well as whether the brand has the feature(s) which is/are highly regarded by the individual.

We will give p_1 points under the choice set for brand A if only brand A is considered for the next purchase, p_2 points under the choice set for brand A if two brands are considered for the next purchase, of which A is one of them and 0 point under the choice set for brand A if all three brands are considered for the next purchase.

Under the feature category,

(i) we will award q_1 points for brand A if only brand A has the feature which is most important to the individual, $\frac{1}{2}q_1$ points for brand A if both brand A and another brand has the feature which is most important to the individual and 0 point for brand A if all three brands are considered.

(ii) we will award q_2 points for brand A if only brand A has the feature which is 2^{nd} most important to the individual, $\frac{1}{2}q_2$ points for brand A if both brand A and another brand has the feature which is 2^{nd} most important to the individual and 0 point for brand A if all three brands are considered.

(iii) we will award q_3 point for brand A if only brand A has the feature which is 3^{rd} most important to the individual, $\frac{1}{2}q_3$ points for brand A if both brand A and another brand has the feature which is 3^{rd} most important to the individual and 0 point for brand A if all three brands are considered.

Likewise, we will deduct points if the brand is not selected just like before.

The table showing the scores for all the three brands based on the choice set is given in the two tables below:

Table 15: Scoring for the Brands for all the Individuals Based on Choice Set

Individual	1	2	3	4	5	6
Choice set	$\{A\}$	$\{B\}$	$\{A, B\}$	$\{A, B, C\}$	$\{A, B\}$	$\{B, C\}$
Scoring for Brand A based on choice set	p_1	$-\frac{1}{2}p_1$	p_2	0	p_2	$-2\,p_2$
Scoring for Brand B based on choice set	$-\frac{1}{2}p_1$	p_1	p_2	0	p_2	p_2
Scoring for Brand C based on choice set	$-\frac{1}{2}p_1$	$-\frac{1}{2}p_1$	$-2\,p_2$	0	$-2\,p_2$	p_2

Table 15: (Continued)

Individual	7	8	9	10	11	12
Choice set	$\{A, B, C\}$	$\{A, B\}$	$\{B, C\}$	$\{A, C\}$	$\{B\}$	$\{A, B, C\}$
Scoring for Brand A based on choice set	0	p_2	$-2\,p_2$	p_2	$-\frac{1}{2}p_1$	0
Scoring for Brand B based on choice set	0	p_2	p_2	$-2\,p_2$	p_1	0
Scoring for Brand C based on choice set	0	$-2\,p_2$	p_2	p_2	$-\frac{1}{2}p_1$	0

The next table below show the scoring for the various brands based on their features:

Table 16: Scoring for the Brands for all the Individuals Based on Features, with weights attached

Individual	1	2	3
Scoring for Brand A based on features	$q_1 + q_2 + q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$q_1 + \frac{1}{2}q_2 - \frac{1}{2}q_3$
Scoring for Brand B based on features	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$q_1 + q_2 + q_3$	$-\frac{1}{2}q_1 + \frac{1}{2}q_2 + q_3$
Scoring for Brand C based on features	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$-\frac{1}{2}q_1 - q_2 - \frac{1}{2}q_3$

Table 16: (Continued)

Individual	4	5	6
Scoring for Brand A based on features	$\frac{1}{2}q_1 - \frac{1}{2}q_2 + \frac{1}{2}q_3$	$-\frac{1}{2}q_1 + q_2 + q_3$	$-q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$
Scoring for Brand B based on features	$\frac{1}{2}q_1 + q_2 - q_3$	$q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$\frac{1}{2}q_1 + q_2 - \frac{1}{2}q_3$
Scoring for Brand C based on features	$-q_1 - \frac{1}{2}q_2 + \frac{1}{2}q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$\frac{1}{2}q_1 - \frac{1}{2}q_2 + q_3$

Table 16: (Continued)

Individual	7	8	9
Scoring for Brand A based on features	$-\frac{1}{2}q_1 + \frac{1}{2}q_2 - q_3$	$-\frac{1}{2}q_1 + q_2 + \frac{1}{2}q_3$	$-q_1 - q_2 - \frac{1}{2}q_3$
Scoring for Brand B based on features	$-\frac{1}{2}q_1 - q_2 + \frac{1}{2}q_3$	$q_1 - \frac{1}{2}q_2 + \frac{1}{2}q_3$	$\frac{1}{2}q_1 + \frac{1}{2}q_2 + q_3$
Scoring for Brand C based on features	$q_1 + \frac{1}{2}q_2 + \frac{1}{2}q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - q_3$	$\frac{1}{2}q_1 + \frac{1}{2}q_2 - \frac{1}{2}q_3$

Table 16 (Continued)

Individual	10	11	12
Scoring for Brand A based on features	$q_1 + q_2 - \frac{1}{2}q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$\frac{1}{2}q_2 - q_3$
Scoring for Brand B based on features	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$q_1 + q_2 + q_3$	$\frac{1}{2}q_2 + \frac{1}{2}q_3$
Scoring for Brand C based on features	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 + q_3$	$-\frac{1}{2}q_1 - \frac{1}{2}q_2 - \frac{1}{2}q_3$	$-q_2 + \frac{1}{2}q_3$

As an illustration, we assign numerical values to the points given above. For example, we give 3 points under the choice set for brand A if only brand A is considered for the next purchase, 1.5 points under the choice set for brand A if two brands are considered for the next purchase, of which A is one of them and 0 point under the choice set for brand A if all three brands are considered for the next purchase.

Under the feature category,
(i) we will award 3 points for brand A if only brand A has the feature which is most important to the individual, 1.5 points for brand A if both brand A and another brand has the feature which is most important to the individual and 0 point for brand A if all three brands are considered.

(ii) we will award 2 points for brand A if only brand A has the feature which is 2nd most important to the individual, 1 point for brand A if both brand A and another brand has the feature which is 2nd most important to the individual and 0 point for brand A if all three brands are considered.

(iii) we will award 1 point for brand A if only brand A has the feature which is 3rd most important to the individual, 0.5 point for brand A if both brand A and another

brand has the feature which is 3rd most important to the individual and 0 point for brand A if all three brands are considered.
Likewise, we will deduct points if the brand is not selected just like before.

The table showing the scores for all the three brands based on the choice set is given in the two tables below:

Table 17: Scoring for the Brands for all the Individuals Based on Features

Individual	1	2	3	4	5	6
Choice set	{A}	{B}	{A, B}	{A, B, C}	{A, B }	{ B, C }
Scoring for Brand A based on choice set	3.0	–1.5	1.5	0	1.5	–3.0
Scoring for Brand B based on choice set	–1.5	3.0	1.5	0	1.5	1.5
Scoring for Brand C based on choice set	–1.5	–1.5	–3.0	0	–3.0	1.5

Table 17: (Continued)

Individual	7	8	9	10	11	12
Choice set	{A, B, C}	{A, B }	{ B, C}	{A, C}	{ B }	{A, B, C}
Scoring for Brand A based on choice set	0	1.5	–3.0	1.5	–1.5	0
Scoring for Brand B based on choice set	0	1.5	1.5	–3.0	3.0	0
Scoring for Brand C based on choice set	0	–3.0	1.5	1.5	–1.5	0

Table 18 below show the scoring for the various brands based on their features:

Table 18: Scoring for the Brands for all the Individuals Based on Features

Individual	1	2	3	4	5
Scoring for Brand A based on features	= (3.0 + 2.0 + 1.0) = 6.0	–3.0	3.5	1.0	1.5
Scoring for Brand B based on features	–3.0	6.0	0.5	2.5	1.5
Scoring for Brand C based on features	–3.0	–3.0	–4.0	–3.5	–3.0

Table 18: (Continued)

Individual	6	7	8	9	10	11	12
Scoring for Brand A based on features	–4.5	–1.5	1.0	–5.5	4.5	–3.0	0
Scoring for Brand B based on features	3.0	–3.0	2.5	3.5	–3.0	6.0	1.5
Scoring for Brand C based on features	1.5	4.5	–3.5	2.0	–1.5	–3.0	–1.5

The scoring table under stage 2 of our measure of brand loyalty is given by:

Scoring for brand = (2/3) $\times$ (scoring based on choice set + scoring based on features)

This is ensure that the scoring for the brand values have a maximum of 6, just as in stage 1. The scoring for brand under stage 2 is given in the table as shown below:

Table 19: Scoring for the Brands for all the Individuals Based on Choice Set and Features

Individual	1	2	3	4	5	6
Scoring for Brand A based on choice set and features	6.0	–3.0	3.33	0.67	2.0	–5.0
Scoring for Brand B based on choice set and features	–3.0	6.0	1.33	1.67	2.0	3.0
Scoring for Brand C based on choice set and features	–3.0	–3.0	–4.67	–2.33	–4.0	2.0

Table 19: (Continued)

Individual	7	8	9	10	11	12
Scoring for Brand A based on choice set and features	–1.0	1.67	–5.67	4.0	–3.0	0
Scoring for Brand B based on choice set and features	–2.0	2.67	3.33	–4.0	6.0	1.0
Scoring for Brand C based on choice set and features	3.0	–4.33	2.33	0	–3.0	–1.0

From the table above, individual 1 is the most loyal customer for Brand A, while individual 2 is the most loyal customer for Brand B and individual 7 is the most loyal customer for Brand C.

This method of measure also allows the brand manager to find out whether the consumer will switch his/her brand when a particular feature of the brand changes. For example, if most people rated price as the most important and attractive feature of brand A, an increase in the price of the brand will affect the sales significantly as the consumers will switch to other brands which are considerably cheaper.

Another usefulness of this method of measure for brand loyalty is that the brand manager would be able to calculate the average brand loyalty measure for the group of 12 individuals. This is helpful as it indicates to the brand manager whether the consumers buying his brand are loyal to him or not. The table below shows the average score of brand loyalty for each of the three brands.

Table 20: Average Brand Loyalty Score for Stage 1 and 2

Brand	A	B	C
Average brand loyalty score under stage 1	1.17	0.67	–1.83
Average brand loyalty score under stage 2	0	1.5	–1.5
Average brand loyalty score under both stages	0.585	1.085	–1.665

It can be deduced from Table 18 above that though brand A scores well in terms of the behavioural aspect (based on historical purchases), it did not perform as well when compared with brand B under stage 2 which focuses on attitudinal aspect of brand loyalty. As a whole, brand B has the highest measure of brand loyalty while brand C has the lowest measure of brand loyalty.

CHAPTER 7 APPLICATION TO THE CHOICES FOR POST-SECONDARY EDUCATION

7.1 Background of Case Study

In the Singapore educational setting, teenagers of age 14-16 have to decide for themselves, probably with some guidance from their parents, teachers and peers as to the next significant step to take after they have completed their ‘O’ Levels. The options available for them, though not many, will to a great extent affect their entire educational horizon in the next 5-7 years of their life, if not more.

Some of the possible options in their consideration include:

a) Enrolling into a Junior College (Brand A);
b) Enrolling into a Polytechnic (Brand B);
c) Enrolling into an Institute of Technical education (Brand C);
d) Overseas education (Brand D);
e) Entering into the job market (Brand E)

In this thesis, I will first be doing a survey of a group of students of age 14-16 to find out their considerations that are important for them when choosing the next possible path after their `O’ Levels. I will be using three types of modeling tools that have been discussed earlier to understand better the working mechanism of these tools when applied to a real life situation. They are namely:

a) Using Non-parametric statistical methods (given in Section 8.4) to help us in identifying the loyalty of individuals to the brands in question;
b) Using Binary Logistic Regression by considering the choice (only Brand A and Brand B are considered here due to the small sample size) made from the individual to be determined by a set of factors;
c) Using Dirichlet Multinomial Distribution modeling by considering past choices made in a family as coming from one individual.

7.2 Details of sampling

A sample of 100 pupils of ages 14-16 years old and coming from different streams (Special, Express and Normal Academic) is being randomly selected from different schools in Singapore. The pupils are individually surveyed on their possible choices of Post-Secondary Schools and the factors affecting their choices. A copy of the survey form is given in Appendix A and the survey results is being summarized in Appendix B.

7.3 Non-parametric statistical methods of measuring loyalty

I will show how the brand loyalty measure as described in Section 8.4 can be applied to the given dataset.

7.3.1 Stage 1

A table showing the past purchases (past purchases here refers to choices made by the individual's siblings/cousins) is being constructed. A portion of such a table is shown below:

Table 21: Observations Based on First 16 Individuals in a Five-brand Problem

Individual	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
3rd purchase	B	A	A	C	C	B	A	B		B	D	A		A	A	B
2nd purchase			A	B	C	D	A			A	A			D	C	
1st purchase			D		C											

Weights are then being allocated depending on whether the brand is chosen during their 1st purchase, 2nd purchase or 3rd purchase. $\omega_1, \omega_2, \omega_3$ are being allocated to the brand according to whether it is their 3rd purchase, 2nd purchase or 1st purchase. A table showing some of these results is given below:

Table 22: Score obtained for Brand A, B, C, D and E Based on Purchase History

Score	1	2	3	4	5
Brand A	$-0.25\omega_1$	ω_1	$\omega_1+\omega_2-0.25\omega_3$	$-0.25(\omega_1+\omega_2)$	$-0.25(\omega_1+\omega_2+\omega_3)$
Brand B	ω_1	$-0.25\omega_1$	$-0.25(\omega_1+\omega_2+\omega_3)$	$-0.25\omega_1+\omega_2$	$-0.25(\omega_1+\omega_2+\omega_3)$
Brand C	$-0.25\omega_1$	$-0.25\omega_1$	$-0.25(\omega_1+\omega_2+\omega_3)$	$\omega_1-0.25\omega_2$	$\omega_1+\omega_2+\omega_3$
Brand D	$-0.25\omega_1$	$-0.25\omega_1$	$-0.25(\omega_1+\omega_2)+\omega_3$	$-0.25(\omega_1+\omega_2)$	$-0.25(\omega_1+\omega_2+\omega_3)$
Brand E	$-0.25\omega_1$	$-0.25\omega_1$	$-0.25(\omega_1+\omega_2+\omega_3)$	$-0.25(\omega_1+\omega_2)$	$-0.25(\omega_1+\omega_2+\omega_3)$

The average scores for each of the five brands are then computed for the whole sample and the results are summarized as follows:

Brand A = $\frac{1}{100}\left(23\frac{1}{4}\omega_1 + 24\frac{3}{4}\omega_2 + 7\frac{3}{4}\omega_3\right)$; Brand B = $\frac{1}{100}\left(25\frac{3}{4}\omega_1 + 14\frac{3}{4}\omega_2 + 2\frac{3}{4}\omega_3\right)$

Brand C = $\frac{1}{100}\left(-13\omega_1 - 11\frac{1}{2}\omega_2 + \frac{1}{4}\omega_3\right)$; Brand D = $\frac{1}{100}\left(-14\frac{1}{4}\omega_1 - 9\omega_2 - 3\frac{1}{2}\omega_3\right)$

Brand E = $\frac{1}{100}\left(-21\frac{3}{4}\omega_1 - 19\omega_2 - 7\frac{1}{4}\omega_3\right)$

Suppose, we chose $\omega_1 = 3$, $\omega_2 = 2$, $\omega_3 = 1$
Then, average scores for brand A, B, C, D and E under Stage 1 will be 1.270, 1.095, −0.618, −0.643 and −1.105

Stage 2

In this stage, the results of the individual's choice set are first summarized in the form of a table. In this stage, a value of p_1 is awarded if a particular brand is the only brand considered by the individual, a value of p_2 is awarded if two brands are considered by the individual and a value of p_3 is awarded if three brands are considered by the individual. Table 23 shows some of the results obtained:

Table 23: Scoring Based on Choice Set

Individual	1	2	3	4	5	6
Choice set	$\{A,B\}$	$\{A\}$	$\{A,B,D\}$	$\{A,B\}$	$\{B,C\}$	$\{B,D\}$
Scoring for Brand A based on choice set	p_2	p_1	p_3	p_2	$-\frac{2}{3}p_2$	$-\frac{2}{3}p_2$
Scoring for Brand B based on choice set	p_2	$-\frac{1}{4}p_1$	p_3	p_2	p_2	p_2
Scoring for Brand C based on choice set	$-\frac{2}{3}p_2$	$-\frac{1}{4}p_1$	$-\frac{3}{2}p_3$	$-\frac{2}{3}p_2$	p_2	$-\frac{2}{3}p_2$
Scoring for Brand D based on choice set	$-\frac{2}{3}p_2$	$-\frac{1}{4}p_1$	p_3	$-\frac{2}{3}p_2$	$-\frac{2}{3}p_2$	p_2
Scoring for Brand E based on choice set	$-\frac{2}{3}p_2$	$-\frac{1}{4}p_1$	$-\frac{3}{2}p_3$	$-\frac{2}{3}p_2$	$-\frac{2}{3}p_2$	$-\frac{2}{3}p_2$

The average scores for all the five brands based on choice set are then computed and given as follows:

For Brand A, average score based on choice set = $\frac{1}{100}\left(9p_1 + 16\frac{2}{3}p_2 + 8\frac{1}{2}p_3\right)$

For Brand B, average score based on choice set = $\frac{1}{100}(9p_1 + 50p_2 + 11p_3)$

For Brand C, average score based on choice set = $-\frac{1}{100}\left(8\frac{1}{2}p_1 + 11\frac{2}{3}p_2 + 14p_3\right)$

For Brand D, average score based on choice set = $-\frac{1}{100}\left(4\frac{3}{4}p_1 + 18\frac{1}{3}p_2 - 8\frac{1}{2}p_3\right)$

For Brand E, average score based on choice set = $-\frac{1}{100}\left(4\frac{3}{4}p_1 + 36\frac{2}{3}p_2 + 14p_3\right)$

If we let p_1 = 3, p_2 = 1.5, p_3 = 1, the average scores for all the five brands based on choice set will be 0.605, 1.130, −0.570, − 0.332 and − 0.833

For the next part of Stage 2 where I look at the factors determining their choice set, pupils were asked about the three main reasons (in rank order) that may influence their choice of Post-Secondary education. For each of the options considered in the choice set, pupils were asked what the reasons were for considering the options. Values of q_1, q_2, q_3 are being allocated to the brands as explained in Section 6.4. A table showing some of the data is given below:

Table 24: Scoring Based on Features

Individual	1	2	3
Scoring for Brand A based on features	$-\frac{1}{4}q_1 + q_2 + \frac{1}{2}q_3$	$q_1 + q_2 + q_3$	$\frac{1}{3}q_1 - \frac{1}{3}q_2 + q_3$
Scoring for Brand B based on features	$q_1 - \frac{1}{4}q_2 + \frac{1}{2}q_3$	$-\frac{1}{4}(q_1 + q_2 + q_3)$	$\frac{1}{3}q_1 + \frac{1}{2}q_2 - \frac{1}{4}q_3$
Scoring for Brand C based on features	$-\frac{1}{4}(q_1 + q_2) - \frac{1}{3}q_3$	$-\frac{1}{4}(q_1 + q_2 + q_3)$	$-\frac{1}{2}q_1 - \frac{1}{3}q_2 - \frac{1}{4}q_3$
Scoring for Brand D based on features	$-\frac{1}{4}(q_1 + q_2) - \frac{1}{3}q_3$	$-\frac{1}{4}(q_1 + q_2 + q_3)$	$\frac{1}{3}q_1 + \frac{1}{2}q_2 - \frac{1}{4}q_3$
Scoring for Brand E based on features	$-\frac{1}{4}(q_1 + q_2) - \frac{1}{3}q_3$	$-\frac{1}{4}(q_1 + q_2 + q_3)$	$-\frac{1}{2}q_1 - \frac{1}{3}q_2 - \frac{1}{4}q_3$

Table 24: (Continued)

Individual	4	5
Scoring for Brand A based on features	$q_1+\frac{1}{2}q_2-\frac{1}{4}q_3$	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$
Scoring for Brand B based on features	$-\frac{1}{4}q_1+\frac{1}{2}q_2+q_3$	$q_1+\frac{1}{2}q_2-\frac{1}{4}q_3$
Scoring for Brand C based on features	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$	$-\frac{1}{4}q_1+\frac{1}{2}q_2+q_3$
Scoring for Brand D based on features	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$
Scoring for Brand E based on features	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$	$-\frac{1}{4}q_1-\frac{1}{3}q_2-\frac{1}{4}q_3$

The average scores based on the features are then computed and given as follows:

For Brand A, average score based on features = $\frac{1}{100}\left(28\frac{5}{12}q_1+16\frac{1}{2}q_2+10\frac{3}{4}q_3\right)$

For Brand B, average score based on features = $\frac{1}{100}\left(27\frac{1}{6}q_1+24q_2+33\frac{1}{4}q_3\right)$

For Brand C, average score based on features = $-\frac{1}{100}\left(17\frac{5}{6}q_1+13\frac{1}{2}q_2+15\frac{1}{2}q_3\right)$

For Brand D, average score based on features = $-\frac{1}{100}\left(14\frac{1}{2}q_1+3\frac{11}{12}q_2+7\frac{1}{6}q_3\right)$

For Brand E, average score based on features = $-\frac{1}{100}\left(23\frac{1}{4}q_1+23\frac{1}{12}q_2+21\frac{1}{3}q_3\right)$

If we let q_1 = 3, q_2 = 2, q_3 = 1, then the average scores for the five brands are 1.290, 1.628, −0.960, −0.585 and −1.373.

The summarized table for the average brand loyalty score under the two stages is given below:

Table 25: Average Brand Loyalty Score Under the Two Stages

Brand	A	B	C	D	E
Average brand loyalty score under stage 1	1.270	1.095	−0.618	−0.643	−1.015
Average brand loyalty score under stage 2	1.263	1.839	−1.020	−0.611	−1.471

When the method of bootstrapping is being used 500 times, the results are summarized as follows:

Table 26: Summary Statistics for Brand Loyalty Score Under Two Stages

Brand	A	B	C	D	E
Average brand loyalty score under stage 1	1.270	1.041	−0.642	−0.616	−1.111
Standard Deviation for the brand loyalty score under stage 1	0.237	0.228	0.125	0.134	0.054
Average brand loyalty score under stage 2	1.282	1.832	−1.024	−0.586	−1.473
Standard Deviation for the brand loyalty score under stage 2	0.273	0.232	0.156	0.185	0.116

So, the corresponding 90% confidence intervals for the brand loyalty score under stage 1 and 2 are given as follows:

Table 27: Confidence Interval for Brand Loyalty Score Under Two Stages

Brand	A	B	C	D	E
90% Confidence Interval for the brand loyalty score under stage 1	(0.873, 1.675)	(0.648, 1.385)	(−0.830, −0.428)	(−0.833, −0.408)	(−1.193, −1.105)
90% Confidence Interval for the brand loyalty score under stage 2	(0.847, 1.748)	(1.432, 2.196)	(−1.286, −0.773)	(−0.866, −0.265)	(−1.641, −1.276)

The confidence intervals between Brands A and B under stage 1 does overlap, showing that the difference may be due to sampling error. However, it is interesting to note that the effect size difference between Brand A and B is close to 1 which is an indication that Brand A is having a higher brand loyalty score under stage 1. Also, the confidence intervals between Brand A and B under stage 2 does overlap, showing that the difference may be due to sampling error, it is also interesting to note that the effect size difference between Brand A and B is close to 2!

Conclusion

The three tables shown above indicate that Brand A and B are the more popular brand among all the five brands if we are to consider past purchases as an indication. This is also true, if we examine the average brand loyalty score under stage 2. It can be noted that Brand A and B are also considered more often as a possible option by many pupils of this group of sample. Though the confidence intervals for brand loyalty of A and B overlap under both stages, it can be observed that the average brand loyalty score for A under stage 1 is higher than that under stage 2 while the average brand loyalty score for A under stage 2 is lower than that under stage 2. Three possible reasons can be attributed to this:

(a) Polytechnic education is getting more popular among would-be school leavers and it indicates that Polytechnic places will become more competitive in the near future.

(b) Though Polytechnic is considered by many before their 'O' Levels, it is not chosen as often as that of JC when the pupil's 'O' Levels results are obtained as it is still viewed as the alternative option if the individual could not enter into JC.

(c) It is possible that there is no difference in the brand loyalty between A and B and the difference in the average brand loyalty of A and B is due to chance.

7.4 Binary Logistic Regression to predict the choice to be made

The individuals in the sample were being asked on some characteristics such as the Stream that they were in, the housing type that they lived in, the preference for studying in a competitive environment among others. At the same time, each individual's score under both stages is computed and total score tallied to give the best predicted choice of study for the individual. A table showing some of the individual scores is given below:

Table 28: Brand Loyalty Score for the Brands under the Two Stages

Individual	1	2	3	4	5	6	7	8	9	10
Brand A	−0.75	3.00	4.75	−1.25	−1.50	−1.25	5.00	−0.75	0.00	1.25
	2.00	3.00	1.00	2.00	−1.333	−1.333	2.00	1.00	−1.333	3.00
	1.75	6.00	1.333	3.75	−1.667	−1.667	3.50	2.25	−1.75	6.00
	0.875	**4.500**	**3.153**	1.292	−1.750	−1.625	**4.333**	0.708	−1.028	**3.625**
Brand B	3.00	−0.75	−1.50	1.25	−1.50	2.50	−1.25	3.00	0.00	2.50
	2.00	−0.75	1.00	2.00	2.00	2.00	2.00	1.00	2.00	−0.75
	3.00	−1.50	1.75	1.25	3.75	3.75	1.00	−0.25	2.00	−1.5
	3.167	−1.125	−0.167	**1.708**	1.167	**3.167**	0.375	**1.750**	1.333	0.500
Brand C	−0.75	−0.75	−1.50	2.50	6.00	−1.25	−1.25	−0.75	0.00	−1.25
	−1.333	−0.75	−1.50	−1.333	2.00	−1.333	−1.333	−1.50	2.00	−0.75
	−1.583	−1.50	−2.417	−1.667	1.25	−1.667	−1.50	−1.50	3.25	−1.5
	−1.347	−1.125	−2.056	0.250	**4.083**	−1.625	−1.569	−1.375	**1.750**	−1.375
Brand D	−0.75	−0.75	−0.25	−1.25	−1.50	1.25	−1.25	−0.75	0.00	−1.25
	−1.333	−0.75	1.00	−1.333	−1.333	2.00	−1.333	1.00	−1.333	−0.75
	−1.583	1.50	1.75	−1.667	−1.667	1.25	−1.50	1.00	−1.75	−1.50
	−1.347	−0.125	0.792	−1.625	−1.750	1.708	−1.569	0.292	−1.028	−1.375
Brand E	−0.75	−0.75	−1.50	−1.25	−1.50	−1.25	−1.25	−0.75	0.00	−1.25
	−1.333	−0.75	−1.50	−1.333	−1.333	−1.333	−1.333	−1.50	−1.333	−0.75
	−1.583	−1.50	−2.417	−1.667	−1.667	−1.667	−1.50	−1.50	−1.75	−1.50
	−1.347	−1.125	−2.056	−1.625	−1.750	−1.625	−1.569	−1.375	−1.028	−1.375

The cell with the highest score for each column is in bold and it indicates the brand that the individual is expected to choose after his/her ‘O’ Levels.

A Binary Logistic Regression is then being done, by first ignoring Brand C-E (due to a lack of data being collected). Some of the cells for the characteristics are also combined. It was found that the variables that are significant (at 8% level of significance) in the analysis are

(i) whether the student prefers a competitive environment or not and

(ii) the stream that the student was currently in

The Binary Logistic Regression equation is given by:

$$P_A = \frac{e^{-3.62+1.80X_1+1.07X_2}}{1+e^{-3.62+1.80X_1+1.07X_2}}$$

where P_A represents the probability of choosing Brand A

X_1 represents the stream that the student was currently in (X_1 = 1 means Express/Special and X_1 = 2 represents Normal)

X_2 represents whether the student prefers a competitive environment or not (X_2 =1 means 'Yes' and X_2 = 2 means 'No')

7.5 Dirichlet Multinomial Distribution Modeling

Based on sub-set of the given data set (sample size = 76) and excluding Brand E which does not appear at all in the reduced data set, the estimated values for the parameters of the Dirichlet Multinomial Distribution are as follows:

$\hat{\alpha}_1 = 0.43$, $\hat{\alpha}_2 = 0.39$, $\hat{\alpha}_3 = 0.12$, $\hat{\alpha}_4 = 0.14$ and $\hat{s} = 1.08$

The distribution of P = (P_1, P_2, P_3, P_4), is thus given by:

$$f(p_1, p_2, p_3, p_4) = \frac{\Gamma(1.08)}{\Gamma(0.43)\Gamma(0.39)\Gamma(0.12)\Gamma(0.14)}\left(p_1^{0.43-1} p_2^{0.39-1} p_3^{0.12-1} p_4^{0.14-1}\right)$$

$$= 0.0039\left(p_1^{0.43-1} p_2^{0.39-1} p_3^{0.12-1} p_4^{0.14-1}\right)$$

The distribution of R is given by:

$$P(R_1 = r_1, R_2 = r_2, R_3 = r_3, R_4 = r_4) = \frac{\Gamma(\hat{s})N!}{\Gamma(\hat{s}+N)}\prod_{j=1}^{4}\frac{\Gamma(\hat{\alpha}_j + r_j)}{\Gamma(\hat{\alpha}_j)r_j!} ,$$

$$r_j = 0, 1, 2, 3 \text{ and } \sum_{j=1}^{4} r_j = 3 \quad \text{for all} \quad j = 1, 2, \ldots, J$$

7.6 Review of the Three Modeling Approaches to the Case Study

In our discussion of the first modeling approach, a detailed method was suggested to enable us to measure the brand loyalty towards each of the educational paths. However, one shortcoming of this approach is that the measure is not robust and the brand loyalty score for each brand may be influenced by the way the purchases and features are coded. It is thus important that one has a good reason for choosing a suitable set of scores for the purchases and features. In the second approach using Logistic Regression, it is highly dependent on the variables which are perceived by the experts to be important factors under consideration. It is difficult to be able to choose a small selected set of variables which can explain a large proportion of the variation in the brand choice among consumers. Finally, in the last approach using Dirichlet modeling, it is easy to use but it is based on the assumption that the probability of choosing a particular brand follows the Dirichlet distribution. However, in some situations, the probability of choosing a particular brand may not follow the distribution well and when that happens, it would be better to consider non-parametric methods of modeling.

BIBLIOGRAPHY

Aaker, D. A. (1991). *Managing Brand Equity*. New York: The Free Press.

Adams, J. C. (1878). "On the value of Euler's constant", *Proceedings of the Royal Society of London, 27*, 88-94.

Anderson, T. W. (1958). *An introduction to multivariate statistical analysis*. New York: Wiley.

Bai, C. (2005). Bayesian network based software reliability prediction with an operational profile. *Journal of Systems and Software*, 77(2), 103-112

Ben-Akiva, M. E. & Lerman, S. R. (1985). *Discrete Choice Analysis: theory and application to travel demand*. Cambridge, Mass: MIT Press.

Berchtold, A. (2001). Estimation in the mixture transition distribution model. *Journal of Time Series Analysis, 22 (4)*, 379-397.

Brasini, S., Freo, M., & Tassinari, G. (2003). *The effects of marketing activities on fast moving consumer good purchases: the case of yoghurt Italian market*. Proceedings of the Third Annual Meeting of ENBIS and ISIS3 (Barcelona, 21-22 august, 2003). A cura di Jeroen de Mast, Barcelona 2003, CD-ROMes: the case of yoghurt Italian market.

Carlin, B. P., & Chib, S. (1995). Bayesian model choice via Markov chain Monte Carlo methods. *Journal of Royal Statistical Society, Series B, 57(3),* 473–484.

Cunningham, R. M. (1956). Brand loyalty – What, where, how much? *Harvard Business Review, 34*, 116-128.

Day, G. S. (1969), A two-dimensional concept of brand loyalty. *Journal of Advertising Research*, Vol 9 No. 3, 29-35.

Durand, J., & Gaudoin, O. (2005). Software reliability modeling and prediction with hidden Markov chains. *Statistical Modeling, 5(1)*, 75-93.

East, R., Harris, P., Lomax, W., & Wilson, G. (1997). First-store loyalty to US and British supermarkets. *Kingston Business School, Occasional Paper Series No 27*.

Efron, B & Tibshirani, R. J.(1993). *An introduction to the bootstrap*. Chapman and Hall.

Ehrenberg, A. S. C. (1972). *Repeat Buying, Theory and Applications*. Amsterdam: North-Holland Publishing Company.

Elrod, T. (1988) A management Science Assessment of Behavioural Measure of Brand Loyalty, *Advances in Consumer Research*, 15, 481-486.

Fisher, R. A. (1921). On the mathematical foundation of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, Vol. 222*, 309-368.

Goodhardt, G. J., Ehrenberg, A. S. C., & Chatfield, C. (1984). The Dirichlet: A comprehensive model of buying behaviour. *Journal of the Royal Statistical Society, Series A, 147(5)*, 621-655.

Guest, L. (1955). Brand loyalty – Twelve years later. *Journal of Applied Psychology, 39*, 405-408.

Hafstrom, J.L., Jung, S. C., & Young, S. C. (1992). Consumer decision-making styles: Comparison between United States and Korean Young Consumers. *The Journal of Consumer Affairs, 26(1),* 146-158.

Hammond, K., East, R., & Ehrenberg, A. (1996). Buying more and buying longer : concepts and measures of consumer loyalty. *London Business School Journal*.

Hawkes, P. (1994). Building brand loyalty and commitment. *The Journal of Brand Management, 1*, 337-347.

Hirschman, E. C. (1987). Marketing research: to serve what purpose? In R. W. Belk and K. Clarke (eds), *Marketing Theory, Proceedings of the AMA Winter Educators' Conference*. Chicago: American Marketing Association, pp. 204-208.

Jacoby, J. (1971). A model of multi-brand loyalty. *Journal of Advertising Research, 11(3)*, 822-829.

Jacoby, J. (1978). *Brand loyalty: measurement and management*. New York : Wiley.

Knox, S. & Walker, D. (2001). Measuring and managing brand loyalty. *Journal of Strategic Marketing, 9(2)*, 111-128.

Krishnamurthi, L., & Raj, S. P. (1991). An empirical analysis of the relationship between brand loyalty and consumer price elasticity. *Marketing Science. 10(2),* 172-183.

Mellens, M., & Dekimpe, M. G., & Steenkamp, J-B. E. M. (1996). A review of brand loyalty measures in marketing. *Tijdschrift voor Economie en Management, XLI (4)*, 507-533.

Montgomery, A. L., Li, S. H., & Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science, 23(4)*, 579-595.

Pessemier, E. A. (1960). An experimental method for estimating demand. *The Journal of Business, Vol. 33*, 373-383.

Prinzie, A., & Van den Poel, D. (2006). Investigating purchasing patterns for financial services using Markov, MTD and MTDg Models. *European Journal of Operational Research, 170(3),* 710-734.

Rossiter. J. R., & Percy, L. (1987). *Advertising and promotion management*. New York: McGraw-Hill.

Rungie, C. (2000). *Heterogeneity in Brand Choice*. Marketing Science Centre.

Yim, C.and Kannan, P. 1998. Consumer Behavioural Loyalty: A Segmentation Model and Analysis. *Journal of Business Research*. 44.

APPENDIX

APPENDIX A

Post-Secondary Education Survey

Section I – Your Choice of Post-Secondary Education

1 What are some possible choices that you will consider after finishing your GCE `O' Levels? (**You may tick more than one.**)

☐**1** Junior Colleges (**JC**) / Pre-U Centres

☐**2** Local Polytechnics

☐**3** Institute of Technical Education (**ITE**)

☐**4** Overseas Colleges / Polytechnics

☐**5** Starting work with no intention to study in the near future

☐**6** Others, please specify (____)

2. Indicate the **three** main reasons that may influence your choice of Post-Secondary Education as given in Q1. Please **rank** them with '1' being the most important reason, '2' being the second most important reason and '3' being the third most important reason.

☐**1** Academic results

☐**2** Financial reasons

☐**3** Popularity of school

☐**4** Interest in the courses offered

☐**5** Family influence

☐**6** Influence from friends

☐**7** Distance from home

☐**8** Lower stress level of school

☐**9** Others, please specify (____)

3. If you choose '**Junior Colleges / Pre-University Centres'** in Q1, what are some of the reasons for making the choice? (**You may tick more than one.**) If not, go to Q4.

☐**1** Academic results

☐**2** Financial reasons

☐**3** Popularity of school

- ☐4 Interest in the courses offered
- ☐5 Family influence
- ☐6 Influence from friends
- ☐7 Distance from home
- ☐8 Lower stress level of school
- ☐9 Others, please specify(____)

4. If you choose '**Polytechnics'** in Q1, what are some of the reasons for making the choice? (**You may tick more than one.**) If not, go to Q5.

- ☐1 Academic results
- ☐2 Financial reasons
- ☐3 Popularity of school
- ☐4 Interest in the courses offered
- ☐5 Family influence
- ☐6 Influence from friends
- ☐7 Distance from home
- ☐8 Lower stress level of school
- ☐9 Others, please specify (____)

5. If you choose '**Institute of Technical Education**' in Q1, what are some of the reasons for making the choice? (**You may tick more than one.**) If not, go to Q6.

- ☐1 Academic results
- ☐2 Financial reasons
- ☐3 Popularity of school
- ☐4 Interest in the courses offered
- ☐5 Family influence
- ☐6 Influence from friends
- ☐7 Distance from home
- ☐8 Lower stress level of school
- ☐9 Others, please specify (____)

6. If you choose '**Overseas Colleges / Polytechnics**' in Q1, what are some of the reasons for making the choice? (**You may tick more than one.**) If not, go to Q7.

- ☐1 Academic results
- ☐2 Financial reasons
- ☐3 Popularity of school
- ☐4 Interest in the courses offered
- ☐5 Family influence
- ☐6 Influence from friends
- ☐7 Distance from home
- ☐8 Lower stress level of school
- ☐9 Others, please specify (____)

7. If you choose to '**start work with no intention to study in the near future**' in Q1, what are some of the reasons for making the choice? (**You may tick more than one.**) If not, go to Q8.

- ☐1 Academic results
- ☐2 Financial reasons
- ☐3 Popularity of school
- ☐4 Interest in the courses offered
- ☐5 Family influence
- ☐6 Influence from friends
- ☐7 Distance from home
- ☐8 Lower stress level of school
- ☐9 Others, please specify (____)

Section II – Your Family's choices of Post Secondary Education

8. Do you have any older siblings/cousins who have **finished** their 'O' Levels? If yes, indicate in the boxes below the choice that they have made after their 'O' Levels.

Eldest sibling/ cousin	2nd oldest sibling cousin	3rd oldest sibling cousin	
☐	☐	☐	Junior Colleges (**JC**) / Pre-U Centres
☐	☐	☐	Local Polytechnics
☐	☐	☐	Institute of Technical Education (**ITE**)
☐	☐	☐	Overseas Colleges / Polytechnics
☐	☐	☐	Starting work with no intention to study in the near future

Section III – More about yourself / family

Tick the box that best describes yourself/family.

9a. I am currently in .

☐1 Special/Express stream
☐2 Normal Academic stream
☐3 Normal Technical stream

9b. How often do you discuss with my family/friends before making an important decision.

☐1 Always/Very Often
☐2 Sometimes
☐3 Rarely/Not at all

9c. I currently live in:

- ☐1 Landed Property/Private Condominium/Executive Condominium
- ☐2 HDB 5 room/Executive flat
- ☐3 HDB 4 room flat
- ☐4 HDB 2/3 room flat

9d. My family owns a car.

- ☐1 Yes
- ☐2 No

9e. I enjoy participating in co-curricular activities in school.

- ☐1 Agree
- ☐2 Disagree
- ☐3 Neither

9f. I enjoy studying in a competitive environment.

- ☐1 Agree
- ☐2 Disagree
- ☐3 Neither

9g. My gender is

- ☐1 Male
- ☐2 Female

9h. My ethnic group is

- ☐1 Chinese
- ☐2 Malay
- ☐3 Indian
- ☐4 Others, specify: ________________

APPENDIX B

S/No	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9a	Q9b	Q9c	Q9d	Q9e	Q9f	Q9g	Q9h
1	1,2	4,1,6	1,6	4,6				2	1	2	3	2	1	1	1	3
2	1	1,3,6	1,3,6					1	1	2	1	2	2	1	2	1
3	1,2,4	2,4,5	2,5	2,4		2,4		4,1,1	1	2	1	1	2	1	1	3
4	1,2	3,5,6	3,5	5,6				2,3	1	1	3	2	1	1	1	1
5	2,3	1,4,8		1,4	4,8			3,3,3	2	3	3	2	1	2	2	2
6	2,4	2,1,4		1,2		1,4		4,2	2	3	1	1	1	2	1	1
7	1,2	1,4,5	1,5	4				1,1	1	2	3	1	1	1	2	1
8	1,2,4	1,2,4	1	4		2		2	1	3	1	1	1	1	1	1
9	2,3	1,6,5		1,5	1,6				2	1	2	1	1	2	2	3
10	1	1,3,6	1,3,6					1,2	1	2	4	2	1	1	2	1
11	1,2	5,4,1	1,5	1,4				1,4	1	1	4	2	1	1	1	1
12	1,2	5,6,7	5,7	6,7				1	1	1	1	2	2	1	1	1
13	3,4	2,1,4			1	2,4			2	3	1	1	1	2	1	2
14	1,4	1,4,5	1,5			1,4		4,1	1	2	1	1	1	1	2	1
15	1,2	4,1,8	1	4,8				3,1	2	3	2	1	1	1	2	4
16	2	4,8,6		4,6,8				2	2	2	1	1	2	1	1	2
17	2,3,5	4,8,2		4	4,8		2	3	2	3	4	2	1	2	1	1
18	2	4,5,8		4,5,8				2,3,2	1	1	3	1	1	1	2	1
19	1,2	1,4,7	1,4	4,7				1,1	1	3	2	2	1	1	1	1
20	1	3,5,1	1,3,5					1	1	1	3	2	1	1	2	2
21	2,3	5,1,4		1,4,5	1,4			2,2	2	1	2	1	1	2	1	3
22	2,4	1,8,4		1,8		4,8		4,4	1	3	1	1	1	2	1	1
23	1,2	1,6,8	1,6	8				1,1	1	1	2	1	1	1	2	1
24	1,2,4	1,2,5	1,5	1		2			1	2	1	1	2	1	1	1
25	5	2,5					2,5	5	3	1	4	2	1	2	1	2
26	2	4,6,8		4,6,8				3,2,2	2	1	3	1	1	1	1	3
27	1,2	3,1,5	1,3	1,5				1,1	1	2	1	1	1	1	2	1
28	1	1,5,3	1,3,5					1,1	1	1	3	2	2	1	1	1
29	2,3	4,8		4	4,8			3,2	2	3	2	2	1	1	1	2
30	2,4	1,4,8		1,4		4,8		2	2	3	1	1	1	1	2	1
31	1	5,6,1	1,5,6					1,1	1	1	2	1	1	1	1	1
32	1,2	1,4,6	1,6	4,6				2,1	1	1	2	1	1	1	2	1
33	2	4,6		4,6				2,1,2	1	1	3	2	2	1	2	1
34	1,2	5,1,4	1,5	4				1,1	1	1	3	2	1	1	2	1
35	2,3	1,2		1,2	1			1,3	2	3	3	2	1	2	2	2
36	1	1,4,5	1,4,5					2,1	1	1	3	2	1	1	1	1
37	1,2,4	1,2,4	1	1,4		2,4		4,1	1	1	1	1	1	1	1	1
38	1,2	1,3,6	1,3	6				1,1,2	1	1	3	2	1	1	1	1
39	2	8,4,6		4,6,8				2	1	1	2	2	1	1	2	1
40	2,3	8,1,6		1	6,8			3,2,2	2	1	3	2	1	2	1	1
41	1,2	1,7,4	1	4,7				2,1	2	2	2	2	2	1	1	2
42	2	4,6,8		4,6,8				1,2,2	1	1	3	2	1	2	1	1
43	2,4	1,2,5		1,2		1,2,5		1,4	1	2	1	1	2	2	1	1
44	2,3	4,1,8		1,4	1,8			3,2,2	2	3	2	1	1	2	2	2
45	4	8.4,1				1,4,8		4,1	2	2	1	1	2	2	1	1
46	1,2	4,1,5	1,5	4,5				1	1	1	1	1	1	1	2	1
47	1	1,3,6	1,3,6						1	1	1	1	1	1	2	3
48	1	4,1,5	1,4,5					1,2	1	1	2	1	1	1	2	1
49	1,2	5,1,6	1,5	5,6				2,2	1	1	1	2	1	1	1	1
50	3,4	2,5,1			1,2	2,5			3	2	1	1	1	2	1	1

51	2	6,1,5		1,5,6				3,2	1	1	4	2	2	1	2	2
52	2,4	4,5,2		2,5		2,4,5		2,4	1	2	1	1	2	2	1	1
53	1,2	1,6	1,6	6				2,2	1	1	2	1	1	1	1	3
54	5	2,8					2,8		3	2	4	2	1	2	2	2
55	2,3	6,1,5		1,5,6	1,6			2	2	1	4	2	1	2	1	1
56	2	4		4				2,3	1	1	3	2	2	1	1	1
57	1	1,4	1,4					1,1	1	3	1	2	2	1	2	3
58	2,3	1,8,6		1,6	1,8			2,2,3	2	1	3	1	1	2	1	2
59	1,2	4,6,1	1,6	4,6				1,1,2	1	1	1	1	1	1	2	1
60	1,2,4	1,2,8	1	8		2,8		4,1,1	1	1	1	1	1	2	1	1
61	1,2	1,6,8	1,6	1,8				2,1,1	1	1	3	2	1	2	1	1
62	1	3,5,6	3,5,6					2,1,1	1	1	1	1	1	1	2	1
63	2	6,4,7		4,6,7				2	1	1	4	2	2	1	1	1
64	2,3	8,1,6		1	6,8			3,2,2	2	1	3	2	2	2	1	3
65	1,2,4	1,2,4	1	1,4		2,4		4,1	1	2	1	1	2	1	1	1
66	2,4	1,4,8		1,4		4,8		2	1	3	1	1	2	2	1	1
67	4	4,5				4,5		4,4	2	2	1	1	1	2	1	2
68	1,2	4,5,1	1,4,5	1,4				1,1	1	1	3	1	1	2	2	2
69	1,2	5,6,8	5,6	6,8				1,1,2	1	1	2	1	1	2	2	2
70	1	4,1,3	1,3,4					1,1	1	3	3	1	2	1	2	3
71	1	5,6	5,6					1,4,1	1	1	1	1	1	1	1	1
72	2	5,8,1		1,5,8				2,3,2	1	1	4	2	1	2	2	1
73	2	8,4		4,8				1,2,1	1	2	2	1	2	2	1	1
74	2,3	1,8,5		1,5	1,8				2	2	3	2	1	2	2	2
75	1,2	1,7,4	1	4,7				2,1	1	3	1	1	1	1	2	1
76	1,2,4	2,4,5	2,5	2,4		2,4		4,1,1	1	2	1	1	1	1	2	1
77	1,2	5,1,6	1,5	5,6				2,2	1	1	4	2	2	1	1	1
78	2,4	2,4		2		2,4			2	3	2	1	2	2	2	2
79	2,3	5,1,6		5,6	1			2,3	2	1	3	2	1	2	1	1
80	1,2	3,1,5	1,3	1,5				1,1	1	2	3	2	1	1	2	3
81	4	5,8,4				4,5,8		1,4	2	1	1	1	2	2	1	1
82	1,2	1,4,6	1,6	4,6				2,1	1	2	3	2	2	1	1	2
83	1,2	4,1,6	1,6	4,6				2	1	2	3	1	1	1	2	3
84	2	6,5,8		5,6,8				2,2	1	1	4	2	1	2	2	1
85	1,2	3,1,5	1,3	1,5				1,1	1	2	4	2	2	1	2	1
86	1,2,4	4,6,5	5	4,6		4		2,2,1	1	1	1	1	2	1	2	1
87	1,2	1,4,7	1,4	4,7				1,1	1	3	3	2	1	1	2	2
88	1	3,6,4	3,4,6					2,1,1	1	1	3	2	1	1	1	3
89	1,2	1,6,8	1,6	8				1,1	1	1	4	2	1	2	1	2
90	5	5,2					2,5	3	3	1	4	2	1	2	1	2
91	1,2,4	1,2,4	1,2,5	1,5	1		2	1,1,2	2	2	2	1	2	2	2	1
92	2,3	5,1,4		4,5	1			1,2,2	2	1	1	1	1	2	1	3
93	2	4,8,5		4,5,8				1,2,2	1	2	2	1	2	2	2	1
94	2,4	1,2,4		1,2		1,4		2,2	2	2	1	1	2	2	1	1
95	1,2	1,3,6	1,3	6				1,1,2	1	1	3	2	2	1	2	1
96	1	3,1,6	1,3,6					1,4	1	2	1	2	1	1	2	2
97	1,2,4	4,2,5	2,5	2		4		1,1,2	1	1	2	1	1	1	1	2
98	1,2	1,6	1,6	6				2,2	1	1	3	2	1	1	1	1
99	1,2	4,6,1	1,6	4,6				1,1,2	1	1	3	2	1	1	2	1
100	2	8,4,5		4,5,8				3,2,2	1	2	2	1	1	2	1	1