An Investigation of Purchasing Behaviour Models

by

YUK MING CHOW

A thesis
presented to the University of Manitoba
in partial fulfillment of the
requirements for the degree of
Master of Science
in
Actuarial and Management Sciences

Winnipeg, Manitoba

ISBN 0-315-34032-0

AN INVESTIGATION OF PURCHASING BEHAVIOUR MODELS

BY

YUK MING CHOW

A thesis submitted to the Faculty of Graduate Studies of the University of Manitoba in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

I hereby declare that I am the sole author of this thesis.

I authorize the University of Manitoba to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Chow, Yuk Ming

I further authorize the University of Manitoba to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Chow, Yuk Ming

The University of Manitoba requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

TO MY GOD

ABSTRACT

In this thesis, consumer behaviour is considered from the point of view of repeat buying. We shall study consumer behaviour models allowing prediction, for example, of proportions and expected frequencies of repeat purchases in a future time interval.

Empirical evidence indicates that characteristic consumer behaviour models involve unimodular, positively skewed distributions which slowly taper off under stationary conditions. Six such characteristic distribution models, the negative binomial distribution, logarithmic series distribution, lognormal distribution, condensed negative binomial distribution and modified negative binomial distribution, will be fitted to four sets of data and compared by the chi-square procedure. In addition we attempt to develop a new model which can better explain the "probabilistic mechanics" of consumer behaviour. To this end we introduce and compare two urn models which will explicitly recognize the tendency that the larger the number of consecutive days on which a consumer buys a given kind of frequently-bought items, the lower the probability that he will buy the item the next day; the larger the number of consecutive days on which a consumer does not buy the item, the higher the probability that he will buy the item the next day.

These two urn models lead to two different generalized hypergeometric distributions, each of which should then be mixed with a suitable parameter distribution to obtain compound hypergeometric distributions as complete consumer behaviour models. In this direction, further research is to be done. However, we have obtained approximations of these compound distributions in the form of type B3 generalized hypergeometric (classified by Shimizu), type IA generalized hypergeometric (classified by Kemp and Kemp), beta-binomial, digamma and trigamma distributions as new consumer behaviour models.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor, Professor S. K. Bhatt, the committee members, Professor E. S. W. Shiu, Professor H. J. Boom, Professor E. S. Rosembloom and the external examiner Professor M. Samanta, for their assistance and interest in examining my thesis. Their valuable and constructive criticism is much appreciated.

My most sincere thanks are due to Professor S. K. Bhatt for introducing me to the field of this work, and his continuing guidance and encouragement throughout the year of my study, as well as Professor H. J. Boom for his patience in improving the readability of the thesis.

I would also like to thank Professor E. S. W. Shiu for helping me in proving some of the theorems, and Joe Hong for typing of the appendix.

CONTENTS

Chapter I

INTRODUCTION

During the last three decades, about a dozen stochastic models have been proposed to estimate the number of times an individual consumer purchases a particular frequently-bought item within a given period of time. In this thesis, some of these models have been compared with respect to their suitability, to properly describe and predict consumer behaviour, concentrating on the one brand, one product theory. They are the Negative Binomial Distribution (NBD), Logarithmic Series Distribution (LSD), Lognormal Distribution, Condensed Negative Binomial Distribution (CNBD), Modified Negative Binomial Distribution (MNBD) and Beta Binomial Distribution (BBD). The BBD is a new consumer behaviour model which will be introduced in chapter 3.

For each of the models, the goodness of fit to four sets of data will be tested by the chi-square procedure. For the estimation of the parameters, we shall prefer maximum likelihood estimation (MLE) because it gives a best asymptotically normal (BAN) estimator when the sample size is large, and because the chi-square test requires the estimates of parameters to be BAN. For some of the above models, however, the maximum likelihood (ML) equations shall be cumbersome to solve. In these cases, other estimation methods will be used.

Of the six consumer behaviour models, only the parameters of LSD, lognormal distribution, NBD, and BBD will be estimated by MLE; the parameters of the MNBD will be estimated by the method of moments and those of the CNBD by the "mean and zeros" method. Some of the simple estimation methods will also be used for the purpose of comparison with MLE. For example, the "mean and zeros", a traditional estimation method for fitting the parameters of the NBD to purchase data, will be used to compare with MLE.

The most important contribution of this thesis will be chapter 5 in which another new consumer behaviour model, the Compound Hypergeometric Distributions, will be introduced. One important weakness of the NBD and BBD is that they assume that each purchase is independent of previous purchases and this is believed to be the main reason why the NBD and BBD do not properly fit consumer behaviour with respect to heavily-bought goods. However, compound hypergeometric distributions can recognize this kind of dependence and will give a better description based on two urn models, in which the probability of a purchase will depend on previous purchases. It is to be hoped that the compound hypergeometric distributions will therefore have a better fit than the NBD and BBD and thus avoid the "variance discrepancy" phenomenon which is an important drawback of the NBD.

1.1 DEFINITIONS AND BASIC CONCEPTS

In this section, we shall introduce and define a few terms which will be used in developing various ideas in the thesis. Most of the terms in consumer behaviour literature are taken from Ehrenberg [19].

Best asymptotically normal (BAN) estimator

An asymptotically normal estimator that is consistent and asymptotically efficient.

Consumer behaviour

A process by which consumers decide whether, what, where, when, how and from whom, to buy goods and services.

Consumer behaviour model

A model which estimates the number of times an individual consumer purchases a particular frequently-bought item within a given period of time. For convenience, we shall use the term "Consumer behaviour model" interchangeably with "Purchasing behaviour model".

Efficiency

Let T_1 and T_2 be estimators for the parameter P. If $E(T_1-P)^2 < E(T_2-P)^2$, then T_1 is said to be more efficient than T_2.

Filler trips

Trips to the store in which a consumer fails to purchase a particular item.

Frequently-bought items

According to Ehrenberg [17], these items include bread, breakfast cereal, canned vegetables, cat and dog foods, cocoa, coffee, detergents, food drinks, household and toilet soaps, jams and marmalade, sausages, shampoos, soft drinks, soap, etc. The category of frequently-bought items may be subdivided in heavily-bought items and lightly-bought items.

Household

An individual living alone or a group of people living together and sharing a common dwelling.

Individual, mixing and compound distributions

Let $c(x) = \int_{-\infty}^{\infty} i(x;u)m(u)du$ where c, i and m are probability density functions. For a given u, $i(x;u)$ is an individual distribution; $m(u)$ is the mixing distribution and $c(x)$ is the resulting compound distribution.

Lapsed buyers

Buyers who do not buy a particular item in the current time-period but did buy it in a previous time-period.

Never-buyers

Buyers who never buy the particular item in their life.

New buyers

Buyers who have not bought a particular item in a previous time-period but buy it in the current time-period.

Repeat buyers

Buyers who have bought a particular item in a previous time-period and also buy it in the current time-period.

Repeat buying

Ehrenberg [19] uses the term repeat buying in any situation in which a person buys the item in question more than once.

Stationary condition

Ehrenberg [19] uses this term as referring to "[a] situation where there is no short-term change in the aggregate sales or penetration level of the brand or item in question", that is, "the sum total of all the varying and dynamic marketing inputs - advertising, pricing, distribution, etc - has had no overall effect on the sales of the item in question during the relevant time-period."

Stochastic process

From the non-mathematical point of view, it is a process running along in time and bound by some probabilistic laws.

Variance discrepancy

This is the phenomenon that in certain models of consumer behaviour with respect to certain heavily-bought items, the expected variance may be systematically higher than the observed variance.

In the following section, we shall present some basic symbols and well-known probability distributions in statistics which will be used throughout the thesis.

1.2 BASIC SYMBOLS AND PROBABILITY DISTRIBUTIONS

1. Beta function

$\beta(a,b)$ denotes the beta function, defined by

$$\beta(a,b) = \int_0^1 p^{a-1} (1-p)^{b-1} dp \quad \text{for } a > 0 \text{ and } b > 0.$$

2. Gamma function

$\Gamma(r)$ denotes the gamma function, defined by

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx \quad \text{for } r > 0.$$

3. $\binom{a}{b}$

$\binom{a}{b}$ is defined as $\dfrac{\Gamma(a+1)}{\Gamma(b+1)\Gamma(a-b+1)}$.

4. Beta distribution

A random variable P is defined to have a beta distribution if its density function is

$$f(p) = \frac{1}{\beta(a,b)} p^{a-1} (1-p)^{b-1} \qquad \text{for } 0 < p < 1\ ; \quad a > 0\ ,\ b > 0.$$

Properties:

$E[P] = a/(a+b)$,

$$\mathrm{Var}[P] = \frac{ab}{(a+b+1)(a+b)^2},$$

$$E[P - E(P)]^j = \frac{\beta(j+a,b)}{\beta(a,b)}.$$

5. <u>Binomial distribution</u>

A random variable X is defined to have a binomial distribution if its density function is

$$f(x) = \binom{n}{x} p^{x} (1-p)^{n-x} \quad \text{for } x = 0,1,2,\ldots..n \; ; \; 0 \leqslant p \leqslant 1 \; , \; n = 0,1,2,\ldots.$$

Properties:

$E[X] = np,$

$Var(X) = npq,$

$E[\exp(tX)] = (1 - p + pe^{t})^{n}$.

6. <u>Gamma distribution</u>

A random variable U is defined to have a gamma distribution if its density function is

$$f(u) = \frac{e^{-u/a} u^{k-1}}{a^{k} \Gamma(k)} \quad \text{for } 0 < u < \infty \; ; \; a > 0 \; , \; k > 0.$$

Properties:

$E(U) = ak,$

$Var(U) = a^2k,$

$E[\exp(tU)] = \{1/(1-at)\}^{k}$ for $at < 1$.

7. <u>Hypergeometric distribution</u>

A random variable X is defined to have a hypergeometric distribution if its density function is

$$f(x) = \frac{\binom{r}{x}\binom{s}{n-x}}{\binom{r+s}{n}} \qquad \text{for } \max(0,n-s) \leq x \leq \min(r,n)\ ; \ r,s,n \text{ positive integers},\ n \leq r+s.$$

Properties:

$E(X) = nr/(r+s)$,

$Var(X) = n[r/(r+s)][(s/(r+s)][(r+s-n)/(r+s-1)]$,

$$E[X(X-1)\ldots.(X-j+1)] = \frac{j!\binom{r}{j}\binom{n}{j}}{\binom{r+s}{j}}.$$

8. Lognormal distribution

A random variable X is defined to have a lognormal distribution if its density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\, x\, \sigma} \exp\{-(\ln(x) - \mu)^2/2\sigma^2\} \quad \text{for } 0 < x < \infty;\ -\infty < \mu < \infty,\ \sigma > 0.$$

Properties:

$E[X] = \exp\{\mu + (1/2)\sigma^2\}$,

$Var(X) = \exp\{2\mu + 2\sigma^2\} - \exp\{2\mu + \sigma^2\}$,

$E[X^j] = \exp\{\mu j + (1/2)j^2\sigma^2\}$.

9. Normal distribution

A random variable X is defined to have a normal distribution if its density function is

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\{-[1/2][(x-\mu)/\sigma]^2\} \quad \text{for } -\infty < x < \infty;\ -\infty < \mu < \infty,\ \sigma > 0.$$

Properties:

$E[X] = \mu,$

$\mathrm{Var}(X) = \sigma^2$

$E[\exp(tX)] = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

10. Poisson distribution

A random variable X is defined to have a Poisson distribution if its density function is

$$f(x) = \frac{e^{-\mu}\mu^{x}}{x!} \qquad \text{for } x=0,1,2,3,\ldots\ ;\ \mu > 0.$$

Properties:

$E[X] = \mu,$

$\mathrm{Var}(X) = \mu,$

$E[\exp(tX)] = e^{\mu(e^{t}-1)}.$

11. Negative binomial distribution

A random variable X is defined to have a negative binomial distribution if its density function is

$$f(x) = \binom{r+x-1}{x} p^{r} (1-p)^{x} \qquad \text{for } x = 0,1,2,\ldots\ ;\ r > 0,\ 0 < p \leq 1.$$

Properties:

$E[X] = r(1-p)/p,$

$\mathrm{Var}(X) = r(1-p)/p^2,$

$$E[\exp(tX)] = \left(\frac{p}{1-(1-p)e^{t}}\right)^{r}.$$

1.3 REVIEW OF LITERATURE

The general pattern of repeat purchasing behaviour under the stationary condition, based on empirical evidence, points towards theoretical distributions which are unimodular and positively skewed, so it was natural to try the Poisson distribution to represent these data. However, this distribution did not show a good fit because it tapers off very fast to the right (dies down). The NBD, introduced in 1959 by Ehrenberg [17], is considered to be the earliest model describing the repeat-buying of a customer with great success. The NBD is more skewed than the Poisson distribution and tapers off more slowly to the right. It has been extensively applied and developed in empirical commercial studies of purchasing patterns.

Ehrenberg's early NBD model focuses on the number of units purchased of a given brand-package-size as the variable to be analysed. This usually works well for a single pack-size of any particular brand. But aggregation of different pack-size poses a major problem [19]. Also, for some commodities such as gasoline, of which many customers always buy the same amount, the NBD model is clearly not suitable. In 1969, after Grahn [24] introduced the purchase occasion as the unit of analysis in the purchase frequency model, this was accepted as a standard unit. From then on, the purchase occasion has been used as the analysis unit of the NBD model.

A practical limitation to the direct applicability of the NBD model to the case of certain heavily-bought items is the so called "variance discrepancy" [19]. Heavily-bought items include margarine, detergents, weekly-bought items such as bread, etc. For weekly-bought items, the observed distribution contains a shelf-like discontinuity because most people buy at most once a week but very few buy more than once a week. This causes a discontinuity at the point where the number of purchases equals the number of weeks in the analysis-period.

There are many ways of generating a negative binomial distribution [2]. However, Ehrenberg's NBD model is based on a mixed Poisson-Gamma assumption. That is, purchases of individual consumers follow Poisson distributions for which the mean rate of purchasing across consumers follows a gamma distribution. On the basis of empirical evidence, it is generally agreed that purchases of a light buyer fit the Poisson part of the NBD but this is not so in the case of heavier buyers [16]. Moreover, the Poisson part of the NBD, since it allows a consumer to buy immediately, is also criticized [10]. If purchases of a given consumer follow a Poisson distribution, then the inter-purchase times of that consumer will be exponentially distributed. But the mode of this exponential distribution is attained at zero. This result of our assumption creates a practical difficulty: if a consumer has purchased an item on a certain occasion, then our model requires a high probability that he or she will make the next purchase immediately (indicating a zero inter-purchase time). Nevertheless, the assumption that the mean rate of purchasing across consumers follows a gamma distribution is supported by Goodhardt and Chatfield [23], based on a theorem of Mosimann [54] together with some empirical results.

Another drawback of the NBD is due to the never-buyers. The gamma assumption of the NBD implies that every consumer in the population is a potential buyer, that is, sooner or latter he or she will buy. Morrison [51,53] has shown that the NBD model will be unsuitable if the proportion of never-buyers is large. Moreover, it will lead to badly biased second period predictions, particularly for the proportion of non-buyers in the second period. Assuming that there is a certain number of never-buyers, Morrison [51] has developed a new model by modifying the original NBD. But this new model has been criticized by Ehrenberg [18], who states that the estimating procedure described by Morrison will require the existence of never-buyers, even when they do not exist. Also, Ehrenberg [18] claims that the NBD repeat-buying theory is not sensitive to the distinction between very infrequent buyers and never-buyers.

The question in modelling is not only that a distribution does fit a set of data but that it should also reflect the various features in the data through a physical interpretation of its parameters. Therefore, efforts have been made to search for distributions that fit a given set of data more closely than those already available. Also, if possible, the parameters of the chosen distribution should relate to the characteristics of the given data.

Following the NBD model, a number of new consumer behaviour models have been developed which try to avoid the shortcomings of the NBD.

To answer the problem of the never-buyers, in 1966, Ehrenberg and Goodhardt [7] had developed the LSD as a model of purchasing behaviour. The LSD has a simpler form than the NBD. It is a special case of the truncated NBD, i.e., it is insensitive to the number of non-buyers during the analysis period. The original purpose of proposing the LSD was to bring the "variance discrepancy" under descriptive control. However, these efforts have not been successful and the LSD suffers from the same problem as the NBD. This is why Ehrenberg claims that the NBD repeat-buying theory is not sensitive to the distinction between very infrequent buyers and never-buyers as mentioned earlier.

The CNBD , which focuses on the problem of the Poisson assumption of the NBD, was developed by Chatfield and Goodhardt [10] in 1973. The empirical findings by Herniter [30] and Lawrence [47] reveal that the coefficient of variation (i.e. variance/mean) of the inter-purchase times for an individual consumer lies between 0.7 and 0.8. Also, Herniter suggests that an Erlang-2 distribution, which has a coefficient of variation equal to 0.71 ($=1/\sqrt{2}$), may be more appropriate for a distribution of mean interpurchase times. This is supported by the fact that the coefficient of variation of an Erlang-p distribution is $1/\sqrt{p}$. Also, the Erlang-2 distribution now helps to ease the earlier difficulty with the Poisson part of the NBD. Whereas the result of the Poisson assumption in the NBD was that the ensuing (exponential) interpurchase time distribution had its mode at zero, the Erlang-2 has its mode away from zero interpurchase time. The purchase rate distribution corresponding to the Erlang-2 interpurchase distribution will be called the condensed Poisson distribution.

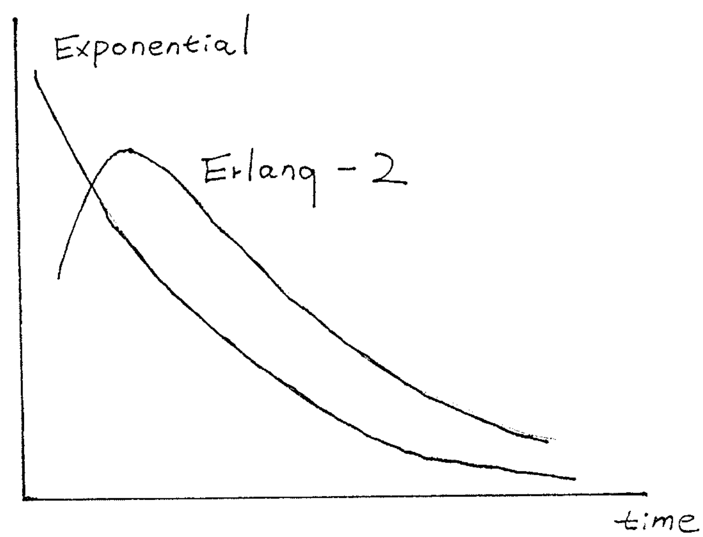


fig 1

Thus, if the inter-purchase times of a given consumer have an Erlang-2 distribution, the number of purchases by a given consumer will follow a condensed Poisson distribution. The CNBD, a (condensed Poisson)-gamma distribution, has substituted the Poisson part of NBD by the condensed Poisson distribution.

It has been showed that the Poisson and condensed Poisson distributions can be significantly distinguished in the case of the heavier buyers [10]. Apparently, this should imply that the CNBD could be a better fit than the NBD in the case of the heavily-bought items in which the NBD suffers from "variance discrepancy". But the result obtained is completely opposite. The CNBD has a slightly better fit than the NBD in the case of lightly-bought items but a worse fit than the NBD in the case of heavily-bought items. Chatfield and Goodhardt [10] explain that this unexpected result is due to the fact that the heavy buyers form only a small part of the buying population for most brands.

In spite of strong evidence that the mean rate of purchasing across consumers follows a gamma distribution [23], Lawrence [47] replaced the gamma distribution with the lognormal in 1980. He fitted a (left-truncated) lognormal distribution successfully to purchases of toothpaste data for a group of households and also for their subgroups.

The use of the lognormal distribution is theoretically based on the assumption that the average interpurchase times for a product or brand are lognormally distributed in the market population. Morrison [52] argues that his modelling effort is not complete in that no individual distribution is specified. Lawrence [48] later replies that his personal belief is that the interpurchase times of a given consumer and mean interpurchase times across consumers both have lognormal distributions, but in his opinion the resulting compound distribution denoting the probability of exactly r purchases by an individual consumer will not have a convenient form in lognormal terms.

In 1980, Frisbie [21] extended the application of the NBD to grocery store trips. However, Bhatt and Vedanand [4] found that the NBD was not a good fit to the data considered by Frisbie and therefore developed a new model, the MNBD, in 1982. Bhatt and Vedanand showed that their MNBD allows a better fit to the data considered by Frishie than the NBD.

Since the analysis unit of the NBD model had been changed to the purchase occasion, but a marketing manager is more interested in the number of units of a particular item a customer will buy, the Generalized Negative Binomial Distribution (GNBD), which is an extended model of the NBD, had been developed by Paull in 1978. This model, using the regression approach, again focused on the number of units of a particular item purchased.

Whereas the NBD deals with a one brand theory, a number of multi-brand theories (i.e., the probability of purchasing different brands of the same product and their relationship) have also been studied and researched [3,11,22,30,31,32,31,49]. Recently, a comprehensive model using the NBD-Dirichlet distribution, combining the NBD and the multi-brand theory, has been developed with great success [22].

Chapter II

CONSUMER BEHAVIOUR MODELS

2.1 A TYPICAL PURCHASING BEHAVIOUR MODEL

Consumer behaviour is a branch of human behaviour. It is the process by which consumers decide whether, what, where, when, how and from whom to buy goods and services [68]. The process is complex. In this thesis, we consider only consumer behaviour in a specific sense, that is repeat-buying [19].

There are a number of variables affecting repeat-buying. They are needs and attitudes before purchasing, the experience of previous usage, the preferences and loyalties with respect to different items and brands, and external influences such as advertising and promotion, retail availability, and differences in product- formulation, packaging and pricing. It is found that many aspects of consumer behaviour can be predicted by using only a few of the many variables involved.

Empirically it appears that most people tend to buy frequently-bought or non-durable items fairly regularly. It then follows that the common pattern of the repeat-buying behaviour of a frequently-bought or undurable item can be described by some stochastic models. Also, as mentioned in the last chapter, the general pattern of repeat purchasing behaviour, based or some empirical evidences, points towards theoretical distributions which are unimodular and positively skewed, under the stationary condition. The scheme of this thesis is to investigate some purchasing behaviour models which have these properties. These models are particularly defined to estimate the number of purchases (purchase occasions) of a particular item by an individual consumer within a given time-period.

A purchasing behaviour model describes how rather than why people buy. For example, the number of people who buy a particular item is related to how often they buy it. A purchasing behaviour model, with a correct theoretical explanation about the building of the model, can also allow us to predict the purchase behaviour in a future period accurately. It is essentially a stochastic model which, if it describes the purchasing behaviour correctly, can allow us to predict the proportions and expected frequencies of purchases of the new buyers, of the lapsed buyers and of the repeat buyers in the future time-period, as long as the stochastic process is in a stationary condition.

As a marketing manager is always interested in knowing how people buy now and how they will buy in future, a suitable consumer behaviour model should allow accurate analysis and decision making.

Now we are in a position to discuss the theory behind various consumer behaviour models. In the following sections we are considering mainly the NBD, LSD, Lognormal, CNBD and MNBD.

2.2 THE NEGATIVE BINOMIAL DISTRIBUTION AS A CONSUMER BEHAVIOUR MODEL

The NBD model was introduced by Ehrenberg [17,19]. It is the earliest model to describe the repeat-buying with great success. In the following sections, the model is discussed in detail with respect to its construction, fitting the parameters and predicting.

2.2.1 Building the Model

The NBD model was derived as follows :

Assumptions:

1. The probability that a consumer with average rate of purchases μ makes R purchases has a Poisson distribution given by

$$P[\ R=r\,|\,\mu\ ] = \frac{e^{-\mu}\mu^{r}}{r!} \quad \text{for } r = 0,1,2,\ldots.;\ \mu > 0.$$

2. The long run average rates of purchases (μ's) among consumers vary; they are considered to have a gamma distribution with two parameters k and a and with density function

$$f(\mu) = \frac{e^{-\mu/a}\,\mu^{k-1}}{a^{k}\,\Gamma(k)} \qquad \text{for } 0 < \mu < \infty \,;\; a > 0 \,,\; k > 0.$$

Hence,

$$P[R=r] = \int_0^\infty P[R=r\,|\,\mu]\,f(\mu)\,d\mu$$

$$= \int_0^\infty \frac{e^{-\frac{\mu}{a}}\,\mu^{k-1}}{a^{k}\,\Gamma(k)}\;\frac{e^{-\mu}\,\mu^{r}}{r\,!}\;d\mu$$

$$= \frac{1}{a^{k}}\;\frac{\Gamma(k+r)}{\Gamma(k)\,r\,!}\left[\frac{1}{1+\frac{1}{a}}\right]^{k+r}$$

$$\int_0^\infty \frac{e^{-\mu(1+\frac{1}{a})}}{\left(\frac{1}{1+\frac{1}{a}}\right)^{k+r}}\;\frac{\mu^{r+k-1}}{\Gamma(r+k)}\;d\mu$$

$$= \left(\frac{a}{1+a}\right)^{k+r}\frac{1}{a^{k}}\;\frac{\Gamma(k+r)}{\Gamma(k)\,r\,!}$$

$$= \frac{\Gamma(k+r)}{\Gamma(r+1)\,\Gamma(k)}\left(\frac{a}{1+a}\right)^{r}(1+a)^{-k}$$

$$= \binom{k+r-1}{r}\left(\frac{a}{1+a}\right)^{r}\left(\frac{1}{1+a}\right)^{k}$$

The density function given above describes the NBD. Derived in this way the NBD is sometimes called a Compound Poisson process.

The mean and variance of the NBD are :

$E[R] = ak,$

$Var(R) = ak(1+a).$

The NBD can also be expressed in terms of the mean $m = ak$:

$$P[R=r] = \frac{(1+m/k)^{-k}\,\Gamma(k+r)}{\Gamma(r+1)\Gamma(k)}\left(\frac{m}{m+k}\right)^{r} \quad \text{for } r = 0,1,2,\ldots\ ; \quad m > 0\ ,\ k > 0.$$

2.2.2 Predictions for Future Period

One of the advantages of NBD is that predictions can be calculated easily. Many of them can be derived from the Bayesian approach. We use the following notation for periods of length T:

m_T = the mean purchasing rate per consumer in a period of length T units.

b_T = the proportion of buyers in the population during the current period of length T.

b_R = the proportion of repeat buyers in the population.

b_N = the proportion of new buyers in the population.

b_L = the proportion of lapsed buyers in the population.

m_N = the average purchasing rate by new buyers on a 'per consumer' basis.

m_R = the average purchasing rate by repeat buyers on a 'per consumer' basis.

m_L = the average purchasing rate by lapsed buyers on a 'per consumer' basis.

w_N = the average purchasing rate per new buyer.

w_R = the average purchasing rate per repeat buyer.

w_L = the average purchasing rate per lapsed buyer.

Under stationary condition, we will have

$$m_T = Tm$$

$$k_T = k$$

$$b_T = 1 - \left(\frac{k_T + m_T}{k_T}\right)^{-k_T}$$

in particular, $b_2 = 1 - (1+2a)^{-k}$

$$b_R = 1 - 2(1+a)^{-k} + (1+2a)^{-k}$$

$$b_L = b_N = (1+a)^{-k} - (1+2a)^{-k}$$

$$m_R = m\{1 - (1+a)^{-k-1}\}$$

$$m_N = m_L = \frac{m}{(1+a)^{k+1}}$$

$$w_R = m_R/b_R$$

$$w_N = w_L = m_N/b_N = m_L/b_L$$

2.2.3 Estimation of the Parameters

The NBD has two parameters, m and k, of which m can be easily estimated by the sample mean; this is the maximum-likelihood (ML) estimator for m, and it is unbiased. To estimate the parameter k, three methods are commonly available: 1. Method of moments, 2. Mean and zeros, 3. Maximum likelihood. The problem with the method of moments is that the estimator is of low efficiency whenever the distribution is reversed J-shaped [2,8,20], and the purchase data usually form a reversed J-shaped distribution. Therefore we will not use the method of moments for estimating k.

If the distribution is reversed J-shaped, the estimation of k by "mean and zeros" will be highly efficient (usually over 90 per cent [19]). The "mean and zeros" is a simple method. In this method, k is obtained by equating the observed proportion of zeros to the theoretical proportion of zeros, and solving the resulting equation, $b_o = (1+m/k)^{-k}$, for k. Here b_o is the observed proportion of zeros and m is estimated by the sample mean. The equation can be simply solved by iterations.

However, finding the maximum likelihood estimate is more cumbersome [29,56,69]. According to Paull [56], the ML estimates of the parameters m and k can be found by solving the following two equations:

$$\hat{m} = \sum_{r=0}^{\infty} n_r r/N$$

$$\ln\{(\hat{m}+\hat{k})/\hat{k}\} + \{\hat{k}-1\}^{-1} = \sum_{r=0}^{\infty} n_r \sum_{j=1}^{r+1} (\hat{k}+r-j)^{-1} /N$$

Here n_r represents the number of households making r purchases and $N = \sum_{r=0}^{\infty} n_r$.

These two equations can be solved by computer, using standard numerical methods.

2.3 THE LOGARITHMIC SERIES DISTRIBUTION AS A CONSUMER BEHAVIOUR MODEL

The LSD model was introduced by Chatfield, Ehrenberg and Goodhardt [2,19]. It is a simplified model of consumer behaviour derived from a truncated NBD. In this section, we will discuss the model in detail with respect to its construction and fitting.

2.3.1 Building the Model

The LSD model was derived as follows :

Assumptions:

1. There is some (unspecified) proportion of never buyers, i.e., people in the population who never buy the item.
2. Purchases of any one buyer in successive time periods follow a Poisson distribution with a long-time average purchase rate μ.
3. μ has a truncated gamma distribution with density function

$$f(\mu) = \frac{ce^{\frac{\mu}{a}}}{\mu} \qquad \delta \le u \le \infty,$$

where δ is a very small positive number, $a > 0$ and c is a constant chosen so that

$$\int_{\delta}^{\infty} \frac{ce^{\frac{\mu}{a}}}{\mu} d\mu = 1$$

Hence, the probability of an individual consumer making R purchases will have a special truncated NBD with probability function

$$P[\ R=r\ ] = \frac{c \int_{\delta}^{\infty} \frac{e^{-\mu}\mu^{r}}{r!} \left(\frac{e^{-\frac{\mu}{a}}}{\mu}\right) d\mu}{c \int_{\delta}^{\infty} \sum_{r=1}^{\infty} \frac{e^{-\mu}\mu^{r}}{r!} \left(\frac{e^{-\frac{\mu}{a}}}{\mu}\right) d\mu}$$

$$= \frac{\frac{1}{r!\,(1+\frac{1}{a})^{r}} \int_{\delta}^{\infty} e^{-(1+\frac{1}{a})\mu} \left\{(1+\frac{1}{a})\mu\right\}^{r-1} d\left\{(1+\frac{1}{a})\mu\right\}}{\int_{\delta}^{\infty} \sum_{r=1}^{\infty} \frac{e^{-\mu}\mu^{r}}{r!} \left(\frac{e^{-\frac{\mu}{a}}}{\mu}\right) d\mu}$$

Let $t = \int_{\delta}^{\infty} \sum_{r=1}^{\infty} \frac{e^{-\mu}\mu^{r}}{r!} \left(\frac{e^{-\frac{\mu}{a}}}{\mu}\right) d\mu$, then

$$P[R=r] \approx \frac{1}{r!\,(1+\frac{1}{a})^r}\,\Gamma(r)\Big/t$$

$$= \frac{1}{(1+\frac{1}{a})^r\, r}\Big/t \quad \text{, or, putting } q = \frac{a}{1+a},$$

$$P[R=r] = \frac{q^r}{r}\Big/t \,.$$

We have $\sum_{r=1}^{\infty} P[R=r] = 1 = \sum_{r=1}^{\infty} \frac{q^r}{r}\Big/t$

$$= -\frac{\ln(1-q)}{t}$$

$\therefore \quad t = -\ln(1-q)$.

Therefore, $P[R=r] = \dfrac{q^r/r}{-\ln(1-q)}$

$$= \frac{-q^r}{r\ln(1-q)}$$

for $r = 1, 2, 3, \ldots\ldots;$

$0 < q < 1$

This probability function describes the LSD. It will give the same result as the non-zero part of the NBD when the NBD parameter k equals 0. This is why the LSD is a limiting case of the NBD.

The mean and variance of the LSD are :

$$E[R] = -\frac{-q}{(1-q)\ln(1-q)},$$

$$Var(R) = -\frac{-q\{1+(q/\ln(1-q)\}}{(1-q)^2\ln(1-q)}.$$

2.3.2 Estimation of the Parameters

The LSD has only one parameter, q . Since the mean of the LSD, say w, equals $-q/\{(1-q)\ln(1-q)\}$, q can be estimated by substituting the sample mean for w. The sample mean is the ML estimator for w, so, due to the invariance property of maximum-likelihood estimators, the result is the ML estimator for q. The estimate of q can either be found by iterations or approximated by the formula $q \simeq (w-1.4)/(w-1.5)$. This algebraic approximation is acceptable within a range for w of about $2 < w < 20$ [19]. Another way of estimating q is by letting $a = q/(1-q)$, then $w = a/\ln(1+a)$. The latter equation still cannot be solved directly in terms of w, but values of a can be read from a table found in Ehrenberg's book "Repeat Buying" [11].

2.4 THE LOGNORMAL DISTRIBUTION AS A CONSUMER BEHAVIOUR MODEL

The Lognormal distribution was used by Lawrence [47] to represent repeat buying data. He replaced the gamma distribution, which is the traditional assumption for the distribution of mean purchase rates, by a lognormal distribution. In the following sections, the construction and fitting of the model will be described.

2.4.1 Building the Model

If $Y = \ln(X)$ is normally distributed with mean μ and variance σ^2, the distribution of X is known as the lognormal distribution. The distribution functions of X and Y are respectively denoted by

$$\Lambda(X|\mu,\sigma^2) = P[\ X < x\ ],$$
$$N(Y|\mu,\sigma^2) = P[\ Y < y\ ].$$

Now, the specific lognormal distribution model that arises from repeat buying data is based on the assumption:

"Average interpurchase times for a product or brand are lognormally distributed in the market population." [47]

If we start to count the occasions of purchase at a random point of time, then the expected number of purchases by a given consumer with mean interpurchase time s during the time interval of length t is given by $E[N] = t/s$ [13,14].

Under the lognormal assumption we have

$t/s \sim \Lambda(t/s \mid \log(t) - \mu, \sigma^2)$.

Lawrence applies the lognormal distribution to estimate the probability of an individual consumer making r purchases during a time interval of length t.

2.4.2 Estimation of the Parameters

ML estimators for the parameters μ and σ^2 of a normal distribution which is truncated at a known point are provided by Hald [27]. A simpler method of finding the ML estimators is provided by Cohen [12]. The procedure can be easily applied using a logarithmic transformation method.

Because the lognormal is a continuous distribution, we need to group the lognormal variables into intervals to make it possible to represent the number of purchases in discrete numbers. Lawrence [47] uses the group interval from ln(n) to ln(n+1) to represent the nth discrete frequency. However, when the data are grouped, complications arise. Nevertheless, for grouped data, an adjustment for the raw first and second moments is provided by Grundy [25] by means of a formula incorporating the population mean and variance. After estimating the mean and variance by Cohen's method [12], Grundy's adjustment for grouping can be applied. The results can then be reinserted into the Cohen estimation procedure.

The Cohen-Grundy [12,25,47] procedure of finding the maximum likelihood estimate for the truncated lognormal distribution when the data are grouped is given as follows:

After the log-transformation, the lognormal distribution becomes a normal distribution. Cohen [12] gives the ML estimators for the normal distribution which is truncated on the left as follows:

$$\hat{\sigma}^2 = s^2 + \hat{\theta}(\bar{x} - x_0)^2$$
$$\hat{\mu} = \bar{x} - \hat{\theta}(\bar{x} - x_0) \tag{2.1}$$

where $\hat{\mu}$ is the ML estimator for the mean of the full distribution.

$\hat{\sigma}^2$ is the ML estimator for the variance of the full distribution.

s^2 is the observed variance,

$\bar{x}$ is the observed mean,

x_0 is the truncation point,

$\hat{\theta}$ is a function of $s^2/(\bar{x} - x_0)^2$.

The value of $\hat{\theta}$ can be found from a table given in Cohen [12].

If the point of truncation is at ln(1) = 0, (2.1) becomes

$$\hat{\sigma}^2 = s^2 + \hat{\theta}(\bar{x})^2$$
$$\hat{\mu} = \bar{x} - \hat{\theta}(\bar{x}) \tag{2.2}$$

From a grouped truncated normal distribution, Grundy adjusted the first and second moment of the truncated sample as follows:

$$M_1 = [Q] - \frac{[H^2Q] - [H^2]\mu}{12\sigma^2}$$

$$M_2 = [Q^2] + \frac{[H^2]\sigma^2 + 2\mu[H^2Q] - 2[H^2Q^2]}{12\sigma^2} \tag{2.3}$$

where M_1 is the first adjusted moment of the truncated sample,

M_2 is the second adjusted moment of the truncated sample,

H_i is the length of the ith interval,

Q_i is the midpoint of the ith interval,

μ is the mean of the full distribution,

σ^2 is the variance of the full distribution.

Bold square brackets are used to denote averages weighted with the frequencies F_i ; for example,

$$[H_i^2] = [H^2] = \sum_{i=1}^{\infty} F_i H_i^2 / \sum_{i=1}^{\infty} F_i$$

After Grundy's adjustment, we get

$$s^2 = M_2 - M_1^2$$
$$\bar{x} = M_1 \tag{2.4}$$

For the initial estimates $\hat{\sigma}^2$ and $\hat{\mu}$, we can substitute $s^2 = [Q^2] - [Q]^2$ and $\bar{x} = [Q]$ into (2.2) to obtain $\hat{\sigma}^2$ and $\hat{\mu}$, then substitute $\mu = \hat{\mu}$ and $\sigma^2 = \hat{\sigma}^2$ into (2.3) and can get s^2 and $\bar{x}$ from (2.4) after calculating (2.3). The current values of s^2 and $\bar{x}$ will be reinserted into (2.2). Six or seven iterations of the Cohen-Grundy cycle are usually sufficient to determine the ML estimate for μ and σ^2.

Calculating the probability of each interval and converting the variable into that of the lognormal distribution leads to the probability of r purchases being made by an individual consumer.

2.5 THE CONDENSED NEGATIVE BINOMIAL DISTRIBUTION AS A CONSUMER BEHAVIOUR MODEL

The CNBD was introduced by Chatfield and Goodhardt [10]. The model replaces the Poisson component of the NBD by a Condensed Poisson distribution. In the following sections, the construction and fitting of the model will be described.

2.5.1 Building the Model

The CNBD model was derived as follows:

Assumptions:

1. The inter-purchase time T of a particular brand-size by a given consumer has an Erlang 2 distribution, viz.

   $$I(t|\lambda) = \lambda^2 t\{\exp(-\lambda t)\} \quad \text{for } 0 < t < \infty ; \lambda > 0.$$

2. The average rate of purchases among consumers in the long run, μ, is assumed to have a gamma distribution with parameters k and a, and density function

   $$f(\mu) = \frac{e^{-\mu/a}\,\mu^{k-1}}{a^k\,\Gamma(k)} \quad \text{for } 0 < \mu < \infty ; a > 0 , k > 0.$$

Under Assumption 1, if we start to count the occasions of purchases at a random point of time, then R, the number of purchases by a given consumer, will follow a condensed Poisson distribution [26] with mean $\lambda/2$, namely:

$$P[\,R=r\,|\,\lambda\,] = P_P(0) + (1/2)P_P(1) \qquad \text{if } r=0,$$

$$P[\,R=r\,|\,\lambda\,] = (1/2)P_P(2r-1) + P_P(2r) + (1/2)P_P(2r+1) \qquad \text{if } r=1,2,\ldots$$

$$\text{where } P_P(r) = \frac{e^{-\lambda}\lambda^{r}}{r!}.$$

Due to assumption 2, $\mu = \lambda/2$. By a simple transformation, the distribution of λ can be expressed as

$$g(\lambda) = \frac{\lambda^{k-1} e^{-\lambda/2a}}{(2a)^{k}\,\Gamma(k)} \qquad \text{for } 0 < \lambda < \infty;\ a > 0,\ k > 0.$$

Hence, the probability that an individual consumer makes r purchases is given by

$$P[R=r] = \int_0^\infty g(\lambda)P[R=r\,|\,\lambda]\,d\lambda$$

Integrating gives

$$P[R=r] = P_N(0) + (1/2)P_N(1) \qquad \text{if } r=o,$$

$$P[R=r] = (1/2)P_N(2r-1) + P_N(2r) + (1/2)P_N(2r+1) \qquad \text{if } r=1,2,\ldots$$

$$\text{where } P_N(r) = \frac{(1+2a)^{-k}\,\Gamma(k+r)}{\Gamma(k)\,r!}\left(\frac{2a}{1+2a}\right)^{r}.$$

This distribution is called the CNBD with parameters a and k.

The mean and variance of the CNBD are :

$$E[R] = ak,$$

$$Var(R) = ak/2 + [1 - (1+4a)^{-k}]/8 + a^2k.$$

Also, the CNBD can be expressed in term of the mean m, that is equal to ak. It becomes

$$P_N(r) = (1+2m/k)^{-k} \frac{\Gamma(k+r)}{\Gamma(k)r!}\left(\frac{2m}{k+2m}\right)^r$$

2.5.2 Estimation of the Parameters

As in the case with the NBD, the parameters of CNBD can be estimated by "mean and zeros". In both cases the observed mean is the point estimate $\hat{m}$ of m, that is ak. k can be estimated by solving

$$(1 + 2\hat{m}/\hat{k})^{-\hat{k}} \{1 + \hat{m}\hat{k}/(2\hat{m}+\hat{k})\} = b_0$$

where b_0 is the observed proportion of zeros.

The solution for $\hat{k}$ is usually obtained by iterative methods.

2.6 THE MODIFIED NEGATIVE BINOMIAL DISTRIBUTION AS A CONSUMER BEHAVIOUR MODEL

This distribution was introduced by Bhatt and Vedanand [4]. It was successfully applied to a grocery filler trip model.

2.6.1 Building the Model

The MNBD was originally applied to a model for grocery store trips, so there is no theoretical interpretation as a repeat-buying model. However, its application to a filler trips model is as follows.

Assumptions:

1. Consider that a consumer has to buy a particular grocery item during a shopping trip, and that the probability of success, i.e., the consumer buying the item, is p, which is constant among the consumers in the market. Also assume that the trips are statistically independent.
2. The number of grocery items to be purchased over a given time period is not fixed but is a random variable which follows a Poisson distribution with mean μ.

3. Not more than one item is purchased in each trip.

Let n+k be the number of trips required to buy all the k items, then, in n+k-1 trips, the consumer purchases k-1 items and in the last trip he/she buys the kth item. The probability of this is

$$\binom{n+k-1}{k-1} p^{k-1} q^{n} p = \binom{n+k-1}{k-1} p^{k} q^{n} \qquad \text{where } q = 1-p.$$

Mixing the distribution with respect to k, which follows a Poisson distribution, we will have

$$P(n) = \frac{1}{n!}\left\{\frac{d^{n}}{ds^{n}} \exp\left[\frac{\mu q(s-1)}{1-qs}\right]\right\}_{s=0} \qquad \text{for } n = 0,1,2,\ldots. \quad \mu>0,\ 0<q<1.$$

The mean and variance of MNBD are :

$E[n] = \mu q/p$,

$Var(n) = (\mu q/p^2)\ (1+q)$.

2.6.2 Estimation of the Parameters

Both the moment method and "mean and zeros" method will be considered to estimate μ and q, the parameters of MNBD.

For the moment method, $\hat{\mu}$ and $\hat{q}$ can be found by solving

$$m = \hat{\mu}\hat{q}/\hat{p}$$

$$v = \hat{\mu}\hat{q}/\hat{p}^2(1+\hat{q})$$

where m is the observed mean,
v is the observed variance.

We will have

$$\hat{q} = (v-m)/(v+m)$$

$$\hat{\mu} = (1-\hat{q})m/\hat{q}$$

For the "mean and zeros" method, let

$$q_0 = r\hat{q}$$

$$\mu_0 = r\hat{\mu}$$

where $\hat{q}$ and $\hat{\mu}$ are moment estimates,

q_0 and μ_0 are "mean and zeros" estimates.

r is a ratio which can be found by solving

$$b_0 = \exp\{ -\mu_0 q_0 \}$$
$$= \exp\{ -\hat{\mu}\hat{q}r^2 \},$$

here b_0 is the observed proportion of zeros and

$\exp\{ -\mu_0 q_0 \}$ is the expected proportion of zeros.

Chapter III

THE BETA BINOMIAL DISTRIBUTION

3.1 INTRODUCTION

There has been some work showing that the beta-binomial distribution (BBD) does describe some kind of buying behaviour. One example is Chatfield and Goodhardt's BBD [9]. This model states that, if a given consumer has a constant probability p to buy at least one unit of the brand in a particular week, then the number of weeks in which the consumer buys at least one unit in a time period of n weeks will follow a binomial distribution with parameters n and p. Also, the probability p for different consumers is assumed to have a beta distribution with parameters a and b say. Under these two postulates, the probability of an individual consumer buying at least one unit in exactly r out of n weeks will have a beta-binomial distribution with parameters n, a and b.

Another example of the BBD has also been developed by Chatfield and Goodhardt [11]. This model states that, if a given consumer buys an item from the product field, and the probability that the item bought is of brand i (which is one of the brands in the product field) has the constant value p, then the number of times that the consumer makes r purchases of brand i out of n purchases of the product will have a binomial distribution with parameters n and p. Once again, the probability p for different consumers is assumed to have a beta distribution with parameters a and b. Under the two postulates, the probability that an individual consumer makes r purchases of brand i out of n purchases of the product will have a BBD with parameters n, a and b.

The two BBD examples above can be considered to be special cases of Pyatt's BBD [57]. In 1969, Pyatt had suggested a compound binomial distribution (i.e. a BBD) which is very similar to Ehrenberg's NBD as a consumer behaviour model. In his model, he assumes that each household conducts a series of n Bernoulli trials with a probability p of purchasing a particular brand in each trial. Also, p varies among households and has a beta distribution. However, Pyatt gave no unequivocal interpretation of n, one of the parameters of the BBD, although he states:

> 'the question of whether or not [n] is known brings to the fore some ambiguities as to how this parameter should be interpreted. A simple interpretation of [n] would be that it is the number of times a housewife goes out shopping. Accordingly, on each such trip a Bernoulli trial is conducted, with "success" being achieved by buying the product in question.'

The second example of Chatfield and Goodhardt's BBD may be developed from Pyatt's idea [57]. If we assume no household makes more than one purchase of a particular brand in a ticular brand in a week, then the first example of Chatfield and Goodhardt's BBD would become a special case of Pyatt's BBD by interpreting n as the number of weeks.

Pyatt has also compared the compound binomial distribution with Ehrenberg's compound Poisson distribution. The earlier result has shown that the NBD is a limiting form of the BBD by letting n and b approach infinity in a certain way [64]. Pyatt has also shown that the theoretical variance of the BBD, fitted by the "mean and zeros" method is positively associated with the value of the parameter n. This implies that the BBD may have a better fit than the NBD if the fit is judged by "variance discrepancy". The implication will be true given that both distributions are fitted by "mean and zeros" and b approaches infinity in a certain way. This follows from the fact that, when n approaches infinity, the theoretical variance of the BBD will be higher and, when a NBD is fitted by the "mean and zeros" method, the "variance discrepancy", viz. the tendency of the theoretical variance for certain heavily-bought items to exceed the observed variance is likely to emerge. However, the two examples of Chatfield and Goodhardt's BBD above are not similar to Ehrenberg's NBD as a consumer behaviour model. In order for the BBD to be a candidate for replacing the NBD, we shall, in the next section, give an interpretation of the parameter n of the BBD which may bring the problem of "variance discrepancy" under descriptive control.

3.2 BUILDING THE MODEL

The following are the assumptions leading to a BBD model which is different from those of Chatfield and Goodhardt :

1. Let R, the number of purchases of a particular item, have a binomial distribution given by

$$P[R=r\,|\,p] = \binom{n}{r} p^{r} (1-p)^{n-r} ,$$

where p is the probability that a given consumer will buy the item on a given day and n is the total number of days in the analysis-period.

2. The probability P varies from consumer to consumer and has a beta distribution given by

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} \qquad \text{for } 0 < p < 1; \; a > 0 , \; b > 0; \; \beta \text{ is the beta function.}$$

Hence, the probability that an individual consumer makes r purchases is :

$$P[R = r] = \int_0^1 f(p)\binom{n}{r} p^{r} (1-p)^{n-r}\, dp$$

$$= \binom{n}{r} \frac{\beta(a+r,n+b-r)}{\beta(a,b)} \qquad \text{for } r = 0,1,2,\ldots. \; a,b > 0.$$

The distribution determined by this probability function is called the Beta Binomial Distribution (BBD).

The mean and variance of the BBD are:

$$E(R) = \frac{na}{a + b},$$

$$\mathrm{Var}(R) = \frac{(nab)(a+b+n)}{(a+b)^2(a+b+1)}.$$

The following will explain why substituting the BBD for the NBD may avoid the shortcomings of the latter with respect to heavily-bought items. The main feature of the BBD is that the number of purchases R, given p, follows a binomial distribution. In the first place, as the Poisson distribution is a limiting case of the binomial distribution, using this distribution implies that we are considering light buyers so we cannot use this distribution in the case of heavy buyers. Second, as a binomial distribution allows a given buyer to buy only once on any day with probability p, this is consistent with the analyzed data which are usually referred to as daily purchases. In case the number of purchases, R, follows a Poisson distribution, repeat buying is allowed immediately. This is criticized by Chatfield and Goodhardt [10], as discussed earlier. Third, as the NBD is a limiting case of BBD when n and b both approach infinity, and as b approaching infinity implies that most of the buyers are light buyers (see Theorem 1), we can understand why variance discrepancy may prevent the NBD from fitting heavily-bought items. The empirical results of testing the difference between BBD and NBD will be discussed in chapter 4.

Theorem 1

Let the density of a beta distribution be defined as

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} \qquad \text{for } 0 < p < 1; \; a > 0 \,,\; b > 0, \; \beta \text{ is a beta function.}$$

Then, if b approaches infinity, most of the p values will be clustered near the point zero.

Proof:

If b approaches infinity, the mean and variance of the beta distribution will approach

$$\lim_{b \to \infty} \frac{a}{a+b} = 0$$

and $$\lim_{b \to \infty} \frac{ab}{(a+b+1)(a+b)^2} = 0$$ respectively.

From the shape of the distribution it is not difficult to see that most of the p values are clustered at near the point zero. This implies that, when b approaches infinity, most of the consumers in the market are light buyers.

3.3 ESTIMATION OF THE PARAMETERS

Case (i) : n is known

A variety of methods have been discussed about the parameter estimation of the BBD when n is known [9,39,41,46,64]. The method of moments may be the easiest method for calculation. The efficiency of this method usually exceeds 70 per cent. However, the efficiency will decrease as n increases [46]. Chatfield and Goodhardt [9] suggest that the "mean and zeros" method may have a higher efficiency than the moment method if the distribution is reverse J-shaped. This idea is based on the fact that the NBD is a limiting case of the BBD; and the NBD always admits of higher efficiency when the "mean and zero" method is used. However, no formal proof has been given about the comparison of "mean and zeros" and "method of moments". In order to compare them, three methods will be used to fit the distribution to the purchase data. They are MLE, method of moments and "mean and zeros". If "mean and zeros" is better than the method of moments, we will expect the "mean and zeros" estimates to be closer values to the ML estimates because of the asymptotic efficiency property of the ML estimator.

1. Method of moments

Estimates of a and b can be easily found by solving

$$m = \frac{n\hat{a}}{\hat{a}+\hat{b}}$$

$$v = \frac{n\hat{a}\hat{b}(n+\hat{a}+\hat{b})}{(\hat{a}+\hat{b})^2(1+\hat{a}+\hat{b})}$$

where m is the observed mean and v is the observed variance.

We will have the parameter estimates

$$\hat{a} = \frac{-m\{v-m(n-m)\}}{nv - m(n-m)}$$

and $\hat{b} = (n-m)\hat{a}/m.$

2. <u>The mean and zeros method</u>

Chatfield has given an iterative technique used to fit the BBD to the data by the "mean and zeros" method. Initial estimates for a and b can be found by method of moments. Let a' and b' denote the initial estimates, i.e. let

$$a' = \frac{-m\{v-m(n-m)\}}{nv - m(n-m)}$$

$$b' = (n-m)a'/m;$$

here m is the observed mean and v is the observed variance.

Let $a'' = ra'$
$b'' = rb'$,

where a'' and b'' are improved "mean and zeros" estimates, and r is a ratio which can be determined by equating b_o, the observed proportion of zeros, to the proportion of zeros expected on the basis of the parameters ra', rb'; that is, by solving

$$\frac{\beta(ra', n+rb')}{\beta(ra', rb')} = b_o \text{ for r by iterations.}$$

3.<u>Maximum likelihood estimation</u>

The MLE of the parameters of the BBD has been discussed by various authors [39,40,41,64]. Here we present the method given in [41].

Let N_j be the number of households in the sample who make j purchases of the particular item. The log-likelihood function L of the BBD can be expressed as

$$L(N_0, N_1, \ldots\ldots, N_n; a, b, n)$$
$$= \sum_{j=0}^{k} N_j \left[\sum_{r=0}^{j-1} \ln(a+r) + \sum_{r=0}^{k-j-1} \ln(b+r) - \sum_{r=0}^{k-1} \ln(a+b+r) \right]$$

The ML equations, $\frac{\partial L}{\partial a} = 0$ and $\frac{\partial L}{\partial b} = 0$ then are

$$\sum_{j=0}^{k} N_j \left[\sum_{r=0}^{j-1} 1/(a+r) - \sum_{r=0}^{k-1} 1/(a+b+r) \right] = 0$$

and $$\sum_{j=0}^{k} N_j \left[\sum_{r=0}^{k-j-1} 1/(b+r) - \sum_{r=0}^{k-1} 1/(a+b+r) \right] = 0.$$

The ML estimates $\hat{a}$ and $\hat{b}$ can now be found by solving the ML equations. This can be done on a computer by means of the Newton-Raphson iterative method [33], using moment estimates or "mean and zeros" estimates as initial values.

Case (ii) : n is not known

Since there is no clear interpretation about the parameter n in BBD, the estimation of n will also be considered in this thesis. The BBD model in case (i) interprets n as the number of days in the analysis-period. This interpretation of n is not universal. There could be several interpretations of n. One interpretation of n was given as the number of times a housewife goes out shopping, discussed earlier in this chapter. Since there is no standardized interpretation of n, we consider n as another parameter of BBD and therefore estimate it.

Skellam [64] , Tripathi and Gurland [66] have discussed the estimation method which involves three parameters. The method that we discuss here is to combine the ML estimates of a and b with Skellam's estimates of n. After applying Skellam's method to calculate n, a and b can be obtained by MLE.

Let $u'_{(j)}$ be the observed value of the jth factorial moment of the random variable X, that is, the observed value of $E[X(X-1)\ldots(X-j+1)]$ [45]. Also let $R_j = u'_{(j)} / u'_{(j-1)}$. Then an estimate of n can be obtained by solving

$$A\hat{n}^2 + B\hat{n} + c = 0$$

where $A = R_3 - 2R_2 + R_1$

$$B = R_3 R_2 - 2R_3 R_1 + R_1 R_2 - R_3 + 4R_2 - 3R_1$$

$$C = 2R_1 (R_3 - R_2 + 1).$$

Substituting the closest positive integer for n into the BBD, the ML estimates of a and b can be derived as discussed earlier. However, if the estimate for n is negative, this will imply that the distribution is not a BBD but, rather, a type of generalized hypergeometric distribution.

3.4 PREDICTIONS FOR FUTURE PERIOD

As with the NBD, many predictions based on the BBD given by Chatfield and Goodhardt [9] are easily calculated. Using the same notation as before, we have

$$m_T = Tm$$

$$b_T = 1 - \frac{\beta(a, Tn+b)}{\beta(a,b)}$$

in particular, $b_2 = 1 - \dfrac{\beta(a, 2n+b)}{\beta(a,b)}$

$$b_L = b_N = b_2 - b_1$$

$$b_R = 2b_1 - b_2$$

$$m_L = m_N = \frac{n\{\beta(a+1, n+b)\}}{\beta(a,b)}$$

$$m_R = m - m_N$$

$$w_R = m_R/b_R$$

$$w_N = w_L = m_N/b_N = m_L/b_L$$

Also, the BBD allows estimation of the loyalty index [58]. This index was originally used as a measure of brand loyalty. However, it can be extended as a measure of the strength of a given consumer's preference for (i.e., purchase) and against (i.e., do not purchase) a particular item on a given day. This is an advantage over the NBD. In case of a BBD, the loyalty index is $1/(a+b+1)$.

3.5 DISCUSSION

A main weakness of BBD is the Bernoulli assumption, that is, that the probability of a consumer buying a particular item is constant and is independent of previous purchases. Generally speaking, a consumer will have a smaller probability of buying the same item if he or she has bought the item on the previous day. This is the same kind of weakness as that of the Poisson assumption for the NBD. In a Poisson process, purchases of a given consumer are regarded as independent drawings at successive points in time.

However, the weakness of the BBD is related to what is to be considered as a trial. In order to decrease the strength of independence, we will not only be testing the BBD, on the basis of one-day trials, but we shall also try a different time-length for each trial. These different lengths will be integers that can divide the total number of days in the analysis-period. So the BBD can also be tested for independence.

We have seen that, in the NBD model, the mean purchase rate among consumers is assumed to follow a gamma distribution. This assumption has been supported by strong evidence [23]. In the BBD model, the corresponding probability p of purchasing an item is assumed to follow a beta distribution. We also pointed out that the NBD is a limiting case of the BBD. It is therefore of interest to show a certain relevant relation between the gamma and beta distribution. This is what we wish to discuss in the next section.

3.6 RELATION BETWEEN GAMMA AND BETA DISTRIBUTIONS

Chatfield and Goodhardt [9] have noted that the gamma distribution can be obtained from the beta distribution in a special way. Define the beta distribution by its density function as follows:

$$f(p) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} \qquad 0 < p < 1.$$

Put $p(b-1) = z$ and let b approach infinity and p approach zero in such a way that z remains constant. By transformation, the probability of z becomes

$$g(z) = \frac{z^{a-1}e^{-z}}{\Gamma(a)}.$$

This is a special form of gamma distribution. The above relationship can be extended as follows:

Theorem 2

Consider a beta distribution with density function

$$f(p;\ k,n/t+1) = \frac{p^{k-1}(1-p)^{n/t}}{\beta(k,n/t+1)} \qquad \text{for } 0 < p < 1; \quad k > o,\ n/t+1 > 0.$$

Let μ = np. If n and n/t approach infinity such that n is sufficiently larger than μ and t remains a positive constant, then the individual probabilities of μ can be approximately estimated from a gamma distribution with parameters t and k and probability density function

$$g(\mu) = \frac{\mu^{k-1} e^{-\mu/t}}{t^{k} \Gamma(k)} .$$

Proof : Let μ = np. By transformation, we have

$$g(\mu) = \frac{1}{n}\left(\frac{\mu}{n}\right)^{k-1} \frac{\left[\left(1-\frac{\mu}{n}\right)^{n}\right]^{1/t}}{\int_0^1 x^{k-1}(1-x)^{\frac{n}{t}}\,dx}$$

Since $\lim_{x\to 0} (1+cx)^{1/x} = e^{c}$ for constant c,

$$g(\mu) = \frac{\lim_{n\to\infty} \frac{1}{n}\left(\frac{\mu}{n}\right)^{k-1} e^{-\mu/t}}{\int_0^1 x^{k-1}(1-x)^{\frac{n}{t}}\,dx}$$ by letting $n\to\infty$ and n is sufficiently larger than μ

$$= \frac{\mu^{k-1} e^{-\frac{\mu}{t}}}{\lim_{\substack{\frac{n}{t}\to\infty \\ n\to\infty}} t^{k}\left(\frac{n}{t}\right)^{k} \int_0^1 x^{k-1}(1-x)^{\frac{n}{t}}\,dx}$$ by letting $\frac{n}{t}\to\infty$

Since $$\lim_{m\to\infty} m^{k} \int_0^1 x^{k-1}(1-x)^{m}\,dx$$

$$= \lim_{m\to\infty} \int_0^1 (mx)^{k-1}(1-x)^{m}\,d(mx)$$

$$= \lim_{m\to\infty} \int_0^1 (mx)^{k-1}\left(1-\frac{mx}{m}\right)^{m}\,d(mx)$$

$$= \lim_{m\to\infty} \int_0^\infty y^{k-1} \left(1 - \frac{y}{m}\right)^m dy$$

$$= \int_0^\infty y^{k-1} e^{-y} dy = \Gamma(k)$$

we have
$$\lim_{\substack{n\to\infty \\ \frac{n}{t}\to\infty}} q(\mu) = \frac{\mu^{k-1} e^{-\mu/t}}{t^k \Gamma(k)} \qquad \#$$

It is well known that the Poisson distribution is a limiting case of the binomial distribution with parameters n and p obtained by letting n approach infinity and p approach zero so that np = μ remains finite. If we combine this fact with the theorem, it is not difficult to see that the NBD is a limiting case of BBD by replacing the BBD parameter b by n/t + 1, and letting n/t and n approach infinity, while the parameter t remains a constant.

The theorem also implies that the corresponding limiting values of the BBD parameters a and b can be estimated by the NBD parameters k and t such that a $\approx$ k and b $\approx$ n/t + 1. Of the four sets of data, we see that the NBD parameter k comes closer to the BBD parameter a as n becomes larger (refer to tables A1.3, A2.3, A3.3, A4.3, A1.8, A2.8, A3.8, A4.8 in the appendix). Also, if we let n = 1000 in the fourth set of data, it makes the ML estimates of the BBD parameters a and b equal to 2088.5868 and 1.5956 respectively. If we estimate a and b from the NBD parameters m and k, the above theorem will give a $\approx$ k = 1.601 and b $\approx$ 1000k/m + 1 = 2098.1967.

Chapter IV

THE MAIN RESULTS - TESTING THE FIT OF THE MODELS

4.1 THE FOUR SETS OF DATA

All six purchasing behaviour models described in the preceding sections will be applied to each of the four sets of data. Then, with respect to each set of data, the performance and the goodness of fit of the models will be compared. These are the data given in [17], [39], [19], and [56]. The fact that no information about the product field of the data has been given in the journals (except for the fourth set of data) has created some difficulties in that we do not know what items are involved. However, all of the data were related to consumer purchases of frequently-bought goods.

The first set of data appeared in the earliest published example of fitting the NBD to consumer purchases [17]. These data, collected from 2000 households, concern purchases of a lightly-bought good during a period of 26 weeks. Later work has shown that the LSD, CNBD and lognormal distributions also give a good fit to this set, a classical example of applying the NBD.

According to Kalwani [39], the second set of data came from an NPD (National Purchase Diary) consumer panel. These data were collected from 10507 households and concern the purchases of a product class over a one-year period (from June 1, 1977 to May 31, 1978). However, Kalwani's paper only reveals the reports of 10505 households. In this thesis, we assume that the total number of households in the sample is 10505.

The third set of data, from [19], gives a typical example of "variance discrepancy" in applying the NBD. It presents the percentages of the number of households in the sample making r purchases of a weekly bought item during a 4-week period.Note that the χ^2 test needs the value of n, the total number of observations, which is not given in the data set. We therefore have to find χ^2 value in terms of percentages as follows.

Let:

p_i denote observed percentages,

$\hat{p}_i$ the theoretical percentages and

N the total number of observations.

Then $\chi^2 = \sum_i (Np_i - N\hat{p}_i)^2/N\hat{p}_i = N\{\sum_i (p_i - \hat{p}_i)^2/\hat{p}_i\}$

Since N is not known, we can not calculate numerical values of χ^2. However, because N, the number of households is the same for the different models in this data set, we can decide which is the closest fit among them by considering the ratios of the χ^2 values.

The fourth set of data, according to Paull [56], came from a continuous purchase daily panel obtained from the Consumer Panel of Canada, operated by International Surveys Limited in Canada. These data were collected from 710 Ontario households relating to purchases from a product class (facial tissue) during one month. In order to fit the BBD which requires information about the length of the analysis-period, we assume a month consists of 30 days.

In the following sections, we shall discuss the empirical findings according to each of the models, including the BBD which will be considered in more detail, with respect to their estimation methods and fit.

4.2 COMPARISON OF ESTIMATION METHODS

The "mean and zeros" method gives estimates close to those obtained by MLE for the NBD parameters in all four sets of data (refer to tables A1.8, A2.8, A3.8 A4.8 in Appendix). This is consistent with what Ehrenberg [19] says, viz. that the "mean and zeros" method is highly efficient in fitting the NBD to the purchase data. Also, the parameter estimates of the BBD according to the method of moments, the "mean and zeros" and ML method have been compared. Whether or not the parameter estimates result in a good fit, it appears that the "mean and zeros" estimates often are closer to the ML estimates than those made by moment methods (refer to tables A5.1 - A5.4). This is also consistent with what Chatfield and Goodhardt [9] say, viz. that the method of "mean and zeros" will have a high efficiency for the BBD as the NBD is a limiting case of the BBD. The only cases in which "mean and zeros" do not give a close estimate to the MLE's are n=26 in the first set of data and n=14 in the third set of data (refer to tables A5.1, A5.3).

Since the MNBD gives a poor fit to the first set of data by the method of moments, the "mean and zeros" method is used. However, it doesn't give a better fit (refer to table A1.7).

4.3 THE FIT OF THE BETA BINOMIAL DISTRIBUTION

In this section, we will be fitting the BBD to the four sets of data. Different positive integers that can divide the number of days, which is the length of the analysis-period, have been substituted for n, the parameter of the BBD. Then the distribution is fitted to each set of data by the MLE and "mean and zeros" methods (refer to tables A1.2, A1.3, A2.2, A2.3, A3.2, A3.3. A4.2, A4.3). For each set of data, the value of n appears to be positively associated with the expected variance of the BBD fitted by "mean and zeros". This is consistent with Pyatt's theorem (see [57]). However, this is not the case when the BBD is fitted by the MLE (refer to table A1.3). For estimating n Skellam's method has been used. For the first set of data, the estimate of n gives two negative values by Skellam's method, so

BBD is not applicable. Among the different values of n , the BBD has the best fit when n = 182, which is the number of days in the analysis-period (refer to tables A1.2 - 1.3). However, the BBD with p value 0.8596 has a slightly worse fit than the NBD with p value 0.8679 (refer to table A1.9). This is due to the fact that here the "variance discrepancy" is working in the opposite direction, that is, the observed variance is larger than the expected variance of NBD.

For the second set of data, the BBD gives a poor fit through various methods used (refer to tables A2.1, A2.2, A2.3). However, it gives a better fit than the NBD. The estimate of n obtained by Skellam's method doesn't give a better fit than other values of n; in fact, it gives the worst fit (table A2.3).

For the third set of data, the BBD has the best fit, with chi-square 0.2214h, when n = 4 which is exactly equal to the number of weeks in the analysis-period (refer to tables A3.1, A3.2, A3.3). Skellam's estimate gives n = 5, which is close to the number of weeks, but doesn't have a better fit. When NBD is fitted, chi-square equals 5.8486h (refer to table A3.9). These results imply that the "variance discrepancy" is due to many buyers purchasing exactly one purchase weekly and only a few buyers purchasing more than once a week.

For the fourth set of data, the BBD has the best fit, with p value 0.8921, when n = 10 (refer to tables A4.1, A4.2, A4.3). This implies that the BBD with three-day Bernoulli trials may give a suitable fit to the data of the product class of facial tissue. It may be noted that the BBD with one-day Bernoulli trials still has a better fit than the NBD. Skellam's estimate here is n=12. This gives the smallest chi-square value (0.5988), but not the smallest p value among different values of n because one degree of freedom has been lost in estimating the parameter n.

4.4 COMPARISON OF THE PREDICTIONS GIVEN BY NBD AND BBD

The predictions by NBD an BBD give very similar results based on our four sets of data (refer to tables A6.1 - A6.4). However, it will make a significant difference if we interpret n as the number of weeks in the third set of data, particularly in the estimation of b_N ,b_L , m_N and m_L . Thus, the prediction for the weekly-bought items given by the BBD may be improved if we interpret n as the number of weeks.

4.5 COMPARISON OF THE CONSUMER BEHAVIOUR MODELS

When we compare the fit of the different distributions to the purchase data, we will use the MLE for estimation of the parameters whenever possible because the chi square test requires the estimators to be BAN [50].

All of the distribution models, except MNBD, have given a good fit to the first set of data with p value greater than 0.8 (refer to table A1.9). Among the six models, CNBD has the best fit. This is consistent with Chatfield and Goodhardt's empirical finding that the CNBD will have a better fit than the NBD for lightly-bought items [10]. Also, the lognormal distribution has a better fit than the NBD. This result is different from that of Lawrence who compared the lognormal distribution with the NBD over the same data set [47]. This is due to a different grouping of the data.

For the second set of data, all of the distribution models give a poor fit (refer to A2.9). However, the records of the heavy buyers are suspect. If we compare the first set of data with the second, we note that the mean frequency of purchases from the first is 0.636 in 26 weeks and the mean frequency of purchases from the second is 2.0785 in one year. Under the stationary condition, the predicted mean frequency of purchases from the first set of data will be 2(0.636) = 1.272 in 52 weeks (approximately one year). So the second set of data suggests a more heavily-bought item than the first. It is strange that no household out of the sample of 10505 makes more than 20 purchases in one year as there are still 5 households, out of the sample of 2000 from the first set of data, making more than 20 purchases in 26 weeks [see 19].

There are three possible way of accounting for this. The first could be mistakes in recording the information from the heavy buyers. The second could be that the panel of the second set of data ignored households purchasing more than 20 purchases. The third could be that the buying behaviour patterns between the two sets of data are very different. However, some of the distribution models do indeed have a good fit to the purchase of the light buyers. Even if the parameters are estimated from the full distribution, the BBD, with p value 0.4705, gives the best fit of the six distribution models (refer to table A2.10).

For the third set of data, referring to a typical weekly-bought item, it is not possible to justify the fit of all models, since the number of households is not given. However, the lognormal has the lowest chi-square value, 3.4373h (refer to table A3.9). But then the lognormal has lost one more degree of freedom than the other distribution models since it has been truncated. If we interpret the BBD parameter n as being the number of weeks, that is 4, then the BBD will have the lowest chi-square value, viz. 0.2214h (refer to table A3.3). But in this way the BBD will have one degree of freedom less than the lognormal because of different grouping. However, it is obvious that the fit has been much improved if the BBD parameter n is interpreted as the number of weeks for the weekly-bought items.

The MNBD, BBD and NBD give a good fit to the fourth set with p value greater than 0.75 (refer to table A4.9). It is surprising that the MNBD gives the best fit with p value 0.866, since it has a very poor fit to the first set of data. In this case, the LSD gives the worst fit because of the NBD parameter k , which is 1.6, large enough to distinguish the NBD from the LSD.

4.6 CONCLUSION

As a result, we see that the CNBD performs better than the NBD for the lightly-bought goods such as in the case of the first set of data and the BBD performs better than the NBD for the heavily-bought goods such as in the case of the other sets of data. However, the difference in the fit of the NBD and the BBD is small and so is the difference in their predictions for a future period even when the value of n is moderate (say n = 30 in the fourth set of data). This agrees with Skellam's finding that the BBD and NBD give almost identical results even for moderate values of n [64]. The LSD, a special case of NBD, can be used when the NBD parameter k is small. Over the four sets of data, the lognormal distribution, a left truncated distribution, has a better fit than the LSD, a specially truncated NBD distribution. Though the MNBD gives the best fit in the fourth set of data, it remains a question whether it can properly be used as a consumer behaviour model.

For the weekly-bought goods, the modification of interpreting the BBD parameter n as the number of weeks seems suitable. But the BBD may become worse if more consumers in the population buy more than once a week. In this case, Chatfield and Goodhardt's BBD [9], which concerns the number of consumers who make at least one purchase of a particular brand in exactly r out of n weeks, is more suitable. However, a marketing manager is more interested in the daily purchases than the weekly purchases.

The BBD with daily Bernoulli trials fails to solve the problem of "variance discrepancy" because the Bernoulli assumption is broken by the excessively heavy buyers. Thus it is our intention to introduce a new model, focused on the problem of the Bernoulli assumption, in chapter 5.

After all, it is not easy to give strong justifications because of the limitation of the data. More empirical tests are needed to justify the fit and the choice of the consumer behaviour models.

Chapter V

THE COMPOUND HYPERGEOMETRIC DISTRIBUTIONS

There can be several explanations for the failure of the NBD and BBD to model the purchase behaviour of consumers buying heavily-bought items. We will concentrate on one assumption in these two models which, we think, is the major contributor to this failure. Consider the assumption that each purchase of a given consumer is regarded as an independent drawing from a Poisson distribution (in the case of NBD) or a binomial distribution (in the case of BBD). The suitability of this assumption in real world, however, will depend on the length of the interpurchase times. If a smoker purchases a pack of cigarettes on a given day, the probability of buying a pack the next day may not be independent and would be lower than that probability was on the preceding day; similarly, if the smoker did not buy a pack on a given day, the probability of buying a pack the next day would be higher. Thus the probabilities of purchase on two successive day may not be independent. In order to construct a distribution in which these ideas are taken into account, we consider two urn models which would resolve the problem of independence, and, we think, may give a better fit in the case of heavily-bought items. In each of these models, every drawing from the urn representing a purchase incidence depends on previous drawings. The theoretical development of these models is similar to that of the Polya distribution [5] which we will discuss in the next section.

5.1 THE POLYA DISTRIBUTION

Suppose there are M black balls and N - M white ones in an urn. Also suppose that everytime we draw a ball, with replacement, we add to the urn c balls of the same colour as that of the ball drawn. Here c is a specified number. If X represents the number of black balls in n successive drawings, X has a Polya-Eggenberger distribution [37]. Bosch [5] calls it a Polya distribution. It is also a Type IIA generalized hypergeometric distribution or BBD [43].

The Polya distribution can be expressed as

$$Po\ (x;\ r,s,n) = \frac{\binom{n}{x} \beta(r+x, s+n-x)}{\beta(r,s)} \qquad \text{for } x = 0,1,2,\ldots;\ r = M/c,\ s = (N-M)/c,\ r,s,n > 0,\ n \text{ integer.}$$

In the following sections, we will introduce two urn models that may explain the purchasing behaviour of a given consumer for whom the assumption regarding the independence of drawings is not valid. We have already given an example of the frequently-bought items in the introduction of chapter 5.

5.2 URN MODEL A

The procedure is basically the same as that generating the Polya distribution. Suppose there are M black and N-M white ones in an urn. Everytime we draw a ball, with replacement, we now remove from the urn c balls of the same colour as the ball drawn, before the next drawing is performed. With this modification, the distribution of X , the number of black balls in n successive drawings, can be derived as follows :

Case (i)

Suppose there are still some black and white balls in the urn at the nth drawing. Then we have

$$P_1(x;\ N,M,c,n)$$

$$= \frac{\binom{n}{x} M(M-c)(M-2c)\ldots\{M-(x-1)c\}(N-M)(N-M-c)\ldots\{N-M-(n-x-1)c\}}{N(N-c)\ldots\ldots\ldots\{N-(n-1)c\}}$$

$$\text{for } x = 0,1,2,\ldots\ ;\ M > (n-1)c,\ N\text{-}M > (n-1)c,\ n,\ N,\ M,\ c \text{ positive integers.} \tag{5.1}$$

Let $M/c = r$ and $(N-M)/c = s$.

$$P_1(x;\ N,M,c,n)$$

$$= \frac{\binom{n}{x} r(r-1)\ldots\ldots(r-x+1)s(s-1)\ldots\ldots(s-n+x+1)}{(r+s)(r+s-1)\ldots\ldots\ldots\{r+s-(n-1)\}}$$

$$= \frac{n!}{(n-x)!x!} \times \frac{r!}{(r-x)!} \times \frac{s!}{(s-n+x)!} \Big/ \frac{(r+s)!}{(r+s-n)!}$$, where $a! := \Gamma(a+1)$

$$= \frac{r!}{x!\,(r-x)!} \times \frac{s!}{(s-(n-x))!\,(n-x)!} \Big/ \frac{(r+s)!}{(r+s-n)!\,n!}$$

$$= \frac{\binom{r}{x}\binom{s}{n-x}}{\binom{r+s}{n}}$$

for $x = 0,1,2,\cdots,n$; $r>0$, $n>0$, n integer, $r > n-1$, $s > n-1$.

The above distribution is called the Type IA(i) Generalized Hypergeometric Distribution (GHgIA(i)) defined by Kemp and Kemp [43,67]. However, Kemp and Kemp did not derive it through an urn model.

Let $T = r+s$ and $p = r/(r+s)$, then the GHgIA(i) distribution can be expressed as

$$P_1(x;\ T,p,n) = \frac{\binom{Tp}{x}\binom{T-Tp}{n-x}}{\binom{T}{n}} \qquad (5.3)$$

for $x = 0,1,2,\ldots n$; $T > 2n-2$, $n > 0$, n integer, $(n-1)/T < p < 1 - (n-1)/T$.

Case (ii)

Suppose that, at the nth drawing, there are no black balls but some white balls in the urn. This implies that c divides M, $M \leq (n-1)c$ and $N-M > (n-1)c$. We have

$P_1(x;\ N,M,c,n)$

$$= \frac{\binom{n}{x} M(M-c)(M-2c)\ldots\{M-(x-1)c\}(N-M)(N-M-c)\ldots\{N-M-(n-x+1)c\}}{N(N-c)\ldots\ldots\ldots\{N-(n-1)c\}}$$

for $x = 0,1,2,\ldots.,M/c$;
$0 \leqslant M \leqslant (n-1)c$, $(n-1)c < N - M \leqslant N$,
N, c, n positive integers, M integer. (5.4)

Let $M/c = r$ and $(N-M)/c = s$. As in the first case, we have

$$P_1(x;\ r,s,n) = \frac{\binom{r}{x}\binom{s}{n-x}}{\binom{r+s}{n}} \tag{5.5}$$

for $x = 0,1,2,\ldots.,r$; $r \geqslant 0$, $n > 0$,
r,n integers, $r \leqslant n-1$, $s > n-1$

The above distribution belongs to the Type IA(ii) Generalized Hypergeometric Distribution (GHgIA(ii)) defined by Kemp and Kemp [43] (whose GHgIA(ii) distribution, however, allows n to be non-integer).

Let $T = r+s$ and $p = r/(r+s)$. Then the GHgIA(ii) defined as above can be expressed as

$$P_1(x;\ T,p,n) = \frac{\binom{Tp}{x}\binom{T - Tp}{n-x}}{\binom{T}{n}} \tag{5.6}$$

for $x = 0,1,2,\ldots.,Tp$; $T > n-1$, n positive integer,
$0 \leqslant p \leqslant (n-1)/T$, $0 \leqslant p < 1-(n-1)/T$.

The moments of the GHgIA(i) and GHgIA(ii) distributions are as follows [43] :

$$E[X] = rn/(r+s),$$

$$E[X-E(X)]^2 = E[X]\{(r+s-n)/(r+s-1)\}\{s/(r+s)\},$$

$$E[X-E(X)]^3 = E[X-E(X)]^2\{(r+s-2n)/(r+s-2)\}\{(s-r)/r+s)\}.$$

5.3 URN MODEL B

Suppose there are M black and N-M white balls in an urn. Also, suppose everytime we draw a ball, with replacement, we also add to the urn c balls of the opposite colour as the ball drawn, before the next drawing is performed. The distribution of X, the number of black balls in n successive drawings following this procedure, can be derived as follows:

$$P_2(x;\ N,M,c,n)$$

$$= \frac{\binom{n}{x} M(M+c)(M+2c)\ldots\{M+(n-x-1)c\}(N-M)(N-M+c)\ldots\{N-M+(x-1)c\}}{N(N+c)(N+2c)\ldots\ldots\{N+(n-1)c\}}$$

for $x = 0,1,2,\ldots.n$; N,M,c,n positive integers.

Let $M/c = r$ and $(N-M)/c = s$, we have

$$P_2(x;\ N,M,c,n)$$

$$= P_2(x;\ r,s,n)$$

$$= \frac{\binom{n}{x} r(r+1)\ldots.(r+n-x-1)s(s+1)\ldots.(s+x-1)}{(r+s)(r+s+1)\ldots(r+s+n-1)}$$

$$= \frac{\binom{n}{x} \Gamma(s+x)\ \Gamma(r+n-x)\ \Gamma(r+s)}{\Gamma(r)\ \Gamma(s)\ \Gamma(r+s+n)}$$

$$= \frac{\binom{n}{x}\ \beta(s+x,\ r+n-x)}{\beta(r,s)} \qquad (5.8)$$

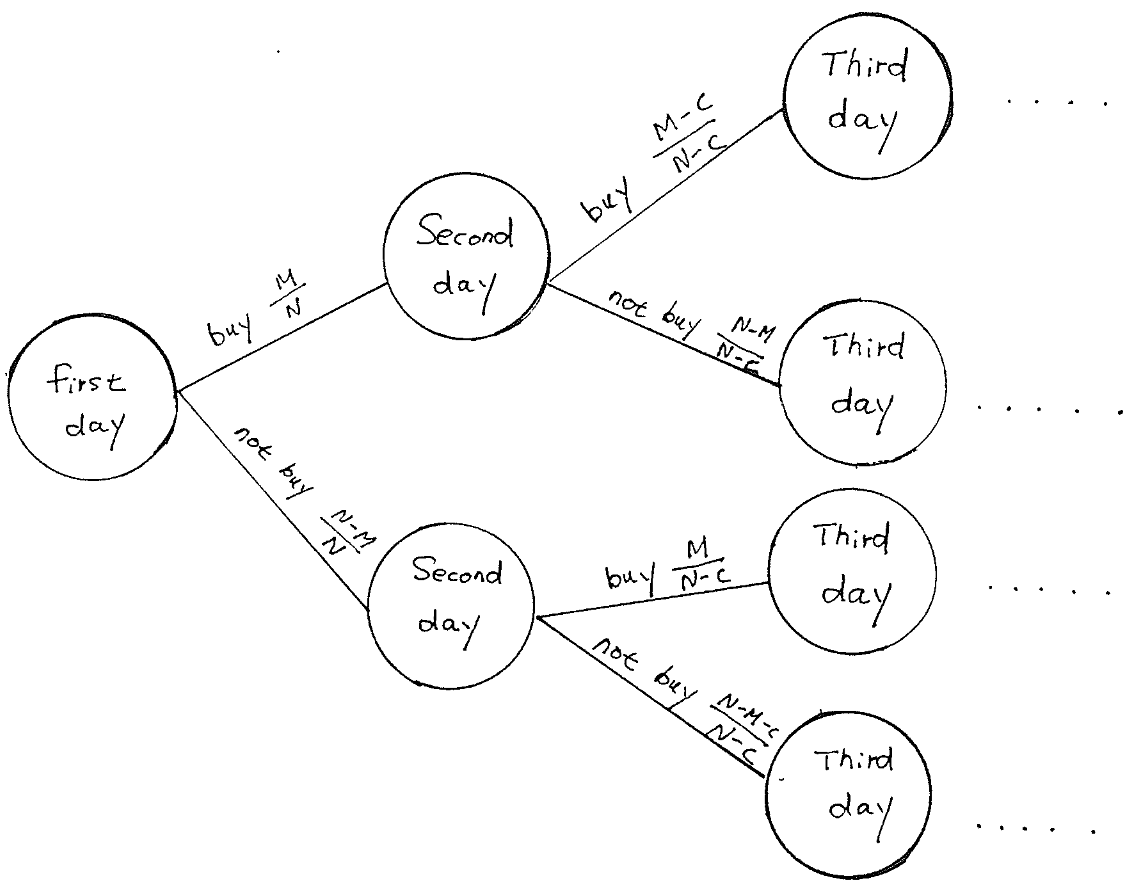
for $x = 0,1,2,\ldots,n$; $r,s,n > 0$, n integer.

Here we have the Polya distribution again.

5.4 THE RELATION BETWEEN THE URN MODELS AND THE BUYING BEHAVIOUR

Comparing urn model A with the buying behaviour of a given consumer, we let the drawing of a black ball from the population be analogous to a consumer buying a particular item on a given day whereas drawing a white ball is analogous to a consumer not buying the item on that day. With this interpretation, n is the number of days analyzed in the time-period. Therefore, the probability that a consumer buys an item on the first day is M/N. If the consumer buys the item the first day, the urn model will imply that the probability that he will buy the item the next day will be smaller, that is $(M-c)/(N-c)$ $(<M/N)$. If the consumer doesn't buy the item the first day, the urn model will imply that the probability that the consumer will buy the item the second day will be larger, that is $M/(N-c)$. Clearly, $M/(N-c) > M/N$. Similarly, if a consumer buys the item on the first two days, the probability that the consumer will buy the item on the third day will be still smaller, that is $(M-2c)/(N-2c)$. Note that $(M-2c)/(N-2c) < (M-c)/(N-c) < M/n$.

The above interpretation of an urn model as the buying behaviour of a given consumer can be represented in a tree diagram as follows :



We can use the urn model B as describing consumer behaviour in a similar way. As a result, the two urn models explain that the larger the number of consecutive days a consumer does not buy the item, the higher the probability that he will buy the item the next day. Also, the larger the number of consecutive days a consumer does buy the item, the smaller the probability that he will buy the item the next day. Thus these urn models can be applied to purchases of frequently-bought items or undurable goods.

5.5 COMPARISON OF THE URN MODELS

As can be seen, the two models mentioned above can allow the probability of purchasing a particular item by a given consumer on a day to depend on the time lapsed since the last purchase. However, there are still a few questions to be answered in order that the urn models represent the purchase behaviour satisfactorily.

A possible disadvantage of the GHgIA(i) distribution is the restriction of the parameter p, whose range is from $(n-1)/T$ to $1 - (n-1)/T$. That is, the model doesn't allow for never-buyers, who may form a large part of the population. However, if the parameter T is sufficiently large, the range of p can approximate that from 0 to 1.

The GHgIA(ii) may be a possible remedy for this since the range of the GHgIA(ii) parameter p starts from zero. That is, it includes the never-buyers in the population. Although the range of the parameter p does not extend to 1, this will not be a problem since it is rare that a consumer buys a particular grocery item every day. A minor disadvantage of the GHgIA(ii) distribution may be the number of black balls M in the urn, which is required to be a multiple of c, the number of balls taken out in each drawing. However, approximate estimates can be obtained even if M is not a multiple of c.

Also, the length of the analysis-period, n, will be sensitive to the degree of independence on each drawing. It can be explained by an example. Suppose that a typical consumer, who makes one purchase of a particular item every t days on the average, will make exactly s purchases during the analysis-period of n days, so that $n=st$. Now think of urn model A, where we put sc black balls and $tsc - sc$ white balls in an empty urn. Then the initial probability (i.e., the probability at the first drawing) of choosing a black ball from the urn is equal to $sc/stc = 1/t$, which means that the consumer purchases a particular item once every t days on the average. Given that a black ball has been chosen on the first drawing (that is, the consumer makes a purchase on the first day), the probability of choosing a black ball again on the second drawing (that is, the consumer makes another purchase on the second day) will be $(sc-c)/(stc-c) = (s-1)/(ts-1)$. If s is large, then $(s-1)/(ts-1)$ will be close to $1/t$, which is the initial probability of choosing a black ball. This implies that each day has approximately the same probability of purchase being made if the analysis-period is long enough. The following two tables display the probabilities of purchase of a weekly-bought item and a biweekly bought item respectively, for a typical consumer on the first two days, for different values of n.

TABLE 1

The probabilities of a purchase of the weekly-bought item for a typical consumer on the first two days based on the urn model A sampling procedure (M=sc, N=stc)

	s			
t=7 n=ts	1 7	2 14	3 21	4 28
P[B]	0.1429	0.1429	0.1429	0.1429
P[B|B]	0	0.0769	0.10	0.1111
P[B|not B]	0.1667	0.1538	0.15	0.1481

TABLE 2

The probabilities of a purchase of the biweekly bought item for a typical consumer on the first two days based on the urn model A sampling procedure (M=sc, N=stc).

	s			
t=14 n=ts	1 14	2 28	3 42	4 56
P[B]	0.0714	0.0714	0.0714	0.0714
P[B|B]	0	0.037	0.0488	0.0545
P[B|not B]	0.0769	0.0741	0.0732	0.0727

The symbols of P[B], P[B|B] and P[B|not B] in tables 1 and 2 denote the following probabilities:

P[B] = P[purchase on the first day] = cs/cst = 1/t

P[B|B]

= P[purchase on the second day|purchase on the first day]

= (cs-c)/(cst-c)

= (s-1)/(st-1)

$P[B|\text{not } B]$

$= P[\text{purchase on the second day}|\text{no purchase on the first day}]$

$= cs/(cst-c)$

$= s/(st-1)$

In order to compare urn models A and B, we have to make the following assumptions regarding the common buying behaviour pattern:

1. The number of white balls in the urn, N-M, is much larger than the number of black balls M. This is always true in the case of lightly-bought items. Even in the case of the heavily-bought items such as the weekly-bought goods, for a typical consumer N-M will be about six times as large as M.
2. With $P[B]$, $P[B|B]$ and $P[B|\text{not } B]$ defined as above, the two urn models will imply $P[B] > P[B|B]$ and $P[B|\text{not } B] > P[B]$. Furthermore, we assume that $P[B] - P[B|B]$ is much larger than $P[B|\text{not } B] - P[B]$.

Consider the urn model B. Under assumption 1, if a black ball is chosen on the first drawing, we need a great amount of white balls added to the urn in order to satisfy assumption 2, that the probability of choosing a black ball is much smaller in the second drawing. This will require c to be a large number. However, assumption 2 cannot hold. This can be explained as follows : if a white ball is drawn, we will add c black balls to the urn, however, $P[B|\text{not } B] - P[B]$ cannot be small compared with $P[B] - P[B|B]$ if c is large.

It can be easily illustrated by an example. For a weekly-bought item, we assume there are 60 white balls and 10 black balls in an urn so that the initial probability of choosing a black ball is 1/7 or 0.1429. Suppose c = 10, then $P[B|B] = 10/80 = 0.125$ and $P[B|\text{not } B] = 20/80 = 0.25$. But $P[B] - P[B|B] = 0.0179 < P(B|\text{not } B] - P[B] = 0.1071$. Therefore assumption 2 is not satisfied.

In order to fit the consumer behaviour, the urn model B can be improved by modifying the sampling procedure as follows:

"Let there be M black balls and N-M white ones in the urn, from which we draw balls one at a time. If a black ball is drawn, we add to the urn c white balls, together with the black one which is replaced. If a white ball is drawn, then we add to the urn d black balls, together with the white one which is replaced."

Unfortunately, the probability distribution of X, the number of black balls in n successive drawings, based on this sampling procedure, will not have a convenient form.

As the urn model A doesn't have the same problem as urn model B, from a theoretical point of view, urn model A is suggested to be a better model. However, because of the disadvantage of urn model A discussed earlier, an empirical investigation of urn models A and B fitted to different kinds of frequently-bought item is needed. With respect to urn model A, it also seems that, as a consumer behaviour model, the GHgIA(ii) distribution is more reasonable than the GHgIA(i) distribution, because of the restrictions on the range of the parameter p.

5.6 AN APPROXIMATION OF THE HYPERGEOMETRIC DISTRIBUTION

If r and s are positive integers, then the GHgIA(i) and GHgIA(ii) distributions derived from urn model A can be regarded as the hypergeometric distribution. It is interesting to note that the hypergeometric probabilities can be approximately estimated by the binomial distribution and Poisson distribution since the former distribution is part of the NBD model and the latter distribution is part of the NBD model [6,15,37,55]. The following are two cases in which the probabilities involved in urn model A can be approximately estimated by the binomial distribution.

Case (i) : r+s approaches infinity

In (5.2) and (5.5), if r+s is sufficiently large, the hypergeometric distribution approximates the binomial distribution [37], that is

$$\frac{\binom{r}{x}\binom{s}{n-x}}{\binom{r+s}{n}} \longrightarrow \binom{n}{x}\left(\frac{r}{r+s}\right)^{x}\left(1-\frac{r}{r+s}\right)^{n-x}$$

when $n < 0.1\,(r+s)$

This implies that, if r+s is sufficiently large, successive drawings from urn model A will be almost independent. Also, if r+s approaches infinity and r remains a constant, the probability that a given consumer purchases a particular item on the first day (= r/(r+s)) will be close to zero, so that, in the case of the lightly-bought items for which the BBD and NBD give a good fit, the probabilities involved in urn model A can be approximately estimated by the binomial distribution.

Case (ii) : c = 0

If c=0, (5.1) and (5.4) become the probability functions of binomial distributions with parameters n and p = M/N since the drawings become independent.

5.7 MIXING THE URN MODELS

We have seen the urn models A and B as two candidates for representing the buying behaviour of a given consumer. However, if we want them to represent a complete model of consumer behaviour, we need to mix them as a convolution of two distributions.

1. Mixing urn model A

It is apparently difficult to mix the urn model A through direct derivation. For the literature related to the generalized hypergeometric distributions and their mixings, refer to [28,43,44,60,61,62,63,67]. However, an approximation can be obtained. If the GHgIA(i) distribution can be represented by a binomial distribution, we can assume that p, the probability that a given consumer buys a particular item on the initial day, varies among consumers and follows a beta distribution. Therefore, we will have the BBD again. On the other hand, Hald [28] has found an interesting theorem about the compound hypergeometric distribution if the GHgIA(i) and GHgIA(ii) distributions are usual hypergeometric distributions. It is stated as follows:

"Let X denote the number of elements having a certain attribute in a population of N elements and let x and y = X - x denote the corresponding numbers of elements in a random sample (drawn without replacement) of size n and in the remainder of the population respectively. If the distribution of X is a hypergeometric, a binomial, a rectangular, a Polya, or a mixed binomial distribution, or any weighted average of these distributions with weight independent of N and X, then for any N the distribution of x is the same as the distribution of X with n substituted for N, and the distribution of y for given x is also of the same type but with parameters depending on x and n."

An interesting part of the theorem is that if we assume that the hypergeometric parameter r of (5.2) or (5.5) follows a Polya distribution (i.e., BBD), then the compound hypergeometric distribution will be the Polya again, that is, the hypergeometric-Polya distribution is identical to the binomial-beta distribution. Not only this, but even more interesting is a theorem of Bosch [5], which states that, if r follows a Polya distribution, r/(r+s) will approach a beta distribution as r+s approaches infinity.

2. Mixing urn model B

A beta-binomial distribution (5.8) can be derived based on urn model B. As in the case with urn model A, it is also difficult to derive urn model B mixed with a suitable distribution. However, it is well known that the NBD is a limiting case of the beta-binomial distribution by letting n and r approach infinity in a certain way. We define the NBD as follows:

$$P[X=x] = NBn(c,p) = \binom{c+x-1}{x} p^{c} (1-p)^{x}$$

for $x = 0,1,2,\ldots$; $c > 0$, $0 < p < 1$.

Furthermore, we shall assume that p follows a beta distribution, namely

$$Be(a,b) = \frac{p^{a-1}(1-p)^{b-1}}{\beta(a,b)} \quad \text{for } 0 < p < 1;\ a,b > o.$$

Mixing the NBD with respect to p, we will have the type B3 generalized hypergeometric (GHgB3) distribution defined by Shimizu [60]. This work was done by Sibuya [61]. In symbols, we write this as follows:

$$GHgB3(x;a,b,c) = NBn(c,p) \underset{p}{\wedge} Be(a,b)$$
$$= \frac{\Gamma(c+a)\,\Gamma(b+a)\,(c)_x\,(b)_x}{\Gamma(a+b+c)\,\Gamma(a)\,x!\,(a+b+c)_x}$$

for $x = 0,1,2,\ldots$; $a,b,c > 0$,

$$(z)_x = \begin{cases} 1 & \text{if } x = 0 \\ z(z+1)\ldots\ldots(z+x-1) & \text{if } x = 1,2,3,\ldots. \end{cases}$$

The GHgB3 distributions are also inverse Polya-Eggenberger distributions, generalized Waring or negative binomial beta distributions [61]. If a and b approach infinity subject to $a/(a+b) = p$, then the negative binomial distribution is a limiting case of the GHgB3 [61]. Sibuya [61] has introduced two new probability distributions named "digamma" and "trigamma" which are obtained as limits of the zero-truncated GHgB3 distribution. If we let $\psi(z) = \frac{d \ln \Gamma(z)}{d\ z}$, then the digamma distribution is defined by

$$q_1(x;a,c) = \frac{1}{\psi(a+c) - \psi(a)} \frac{(c)_x}{x(a+c)_x}$$

for $x = 1,2,....; a > 0 ,c > -1 \ (c \neq 0), \ a+c > 0.$

The trigamma distribution is defined by

$$q_2(x;c) = \frac{1}{\psi'(a)} \frac{(x-1)!}{x(a)_x}$$

for $x = 1,2,....; a > 0.$

These two distributions can be strong competitors of the LSD as consumer behaviour models. As mentioned by Sibuya [61], "Our new distributions are useful as substitutes of the logarithmic series when the observed frequency data have such a long tail that cannot be fitted by the latter distributions."

Based on the two urn models, we can thus generate the GHgB3 distribution, the beta-binomial distribution, the GHgIA distribution , the digamma distribution and the trigamma distribution as consumer behavior models. However, from a theoretical point of view, the validity of most of these models derived from approximations will depend on the category of the frequently-bought items. For example, the beta-binomial distribution is only valid in the case of lightly-bought items. We have already seen the empirical result of this distribution, but, for the other models, more empirical investigation is needed for their justification.

5.8 FURTHER AREAS OF RESEARCH

We have described the two phases in the development of the consumer models. They are from NBD to BBD, and from BBD to compound hypergeometric distributions. However, the compound generalized hypergeometric distributions are only an initial step of the research regarding consumer behaviour models. Further research will involve finding a suitable compound generalized hypergeometric distribution and improving the urn models to better represent consumer behaviour.

REFERENCES

1. Aitchison, J. & Brown, J.A.C. (1957) **The Lognormal Distribution,** Cambridge,England: University Press

2. Anscombe, F.J. (1950) "Sampling Theory of the Negative Binomial and Logarithmic Series Distribution", **Biometrika, 37** ,pp 358-82

3. Bass, Frank M. (1974) "The Theory of Stochastic Preference and Brand Switching", **J.M.R., 11** ,pp 1-20

4. Bhatt, S.K. & Vedanand (1982) "On a Modified Negative Binomial Distribution Applied to Grocery Store Trips", **ASAC Proceedings,** Management Science/Recherche Operationelle Division

5. Bosch, A.J. (1963) "The Polya Distribution", **Statistica Neerlandica, 17** ,pp. 201-213

6. Brunk, H.D., Holstein, J.E. & William, F. (1968) "A Comparison of Binomial Approximations to the Hypergeometric Distribution", **American Statistician, 22** , pp. 24-26

7. Chatfield, C., Ehrenberg, A.S.C. & Goodhardt, G.J. (1966) "Progress on a Simplified Model of Stationary Purchasing Behaviour", **J.R. Statist. Soc., A , 129** , pp 316-367

8. Chatfield, C. (1969) "On Estimating the Parameters of the Logarithmic Series and Negative Binomial Distributions", **Biometrika, 56** ,pp 411-414

9. Chatfield, C. & Goodhardt, G.J. (1970) "The Beta-Binomial Model for Consumer Purchasing Behaviour", **Appl. Statist., 19** ,pp 240-250

10. ------------- (1973) "A Consumer Purchasing Model with Erlang Inter-Purchase Times", **J. Amer. Statist. Ass., 68** ,pp 828-835

11. ------------- (1975) "Results Concerning Brand Choice", **J.M.R., 12** ,pp 110-113

12. Cohen, Jr. A. C. (1959) "Simplified Estimators for the Normal Distribution when Samples are Singly Censored or Truncated", **Technometrics, 1** ,pp 217 - 237

13. Cox, D.R. (1962) **Renewal Theory,** London: Methuen & Co. Ltd.

14. Cox, D.R. & Lewis, P.A.W. (1966) **The Statistical Analysis of Series of Events,** London: Methuen & Co. Ltd.

15. Duncan, A.J. (1974) **Quality Control and Industrial Statistics,** Illinois : Richard D. Irwin ,Inc.

16. Dunn, R., Reader, S. & Wrigley, N. (1983) "An Investigation of the Assumptions of the NBD Model as Applied to Purchasing at Individual Stores", **Appl. Statist., 32** ,pp 249-259

17. Ehrenberg, A.S.C (1959) "The Pattern of Consumer Purchases", **Appl. Statist, 8** ,pp 26-41

18. ------------- (1970) "A Note on Never-Buyers", **J.M.R., 7** ,pp 536-8

19. ------------- (1972) **Repeat-Buying: Theory and Applications,** Amsterdam: North-Holland Publishing Co.

20. Fisher, R.A. (1941) "The Negative Binomial Distribution", **Ann. Eugen., Lond., 11** ,pp 182-187

21. Frisbie, Jr. A.G. (1980) "Ehrenberg's Negative Binomial Model Applied to Grocery Store Trips", **J.M.R., 17** ,pp 385-90

22. Goodhardt, G.J., Ehrenberg, A.S.C. & Chatfield, C. (1984) "The Dirichlet: A Comprehensive Model of Buying Behaviour", **J.R. Statist. Soc., 147 ,Part 5,** pp 621-655

23. Goodhardt, G.J. & Chatfield, C. (1973) "Gamma Distribution in Consumer Purchasing", **Nature, 244,** pp 316

24. Grahn, Gary L. (1969) "The Negative Binomial Distribution Model of Repeat-Purchase Loyalty: An Empirical Investigation", **J.M.R., 6** ,pp 72-8

25. Grundy, P.M. (1952) "The Fitting of Grouped Truncated and Grouped Censored Normal Distributions", **Biometrika, 39** ,pp 252-9

26. Haight, F.A. (1967) **Handbook of the Poisson Distribution,** New York: John Wiley & Sons,Inc.

27. Hald, A. (1949) "Maximum Likelihood Estimation of the Parameters of a Normal Distribution which is Truncated at a Known Point", **Skandinavisk Aktuarietidskrift,** pp 119-134

28. -------- (1960) "The Compound Hypergeometric Distribution", **Technometrics, 2 ,** pp 275

29. Haldane, J.B.S. (1941) "The Fitting of Binomial Distributions", **Ann. Eugen., Lond., 11 ,** pp 179-81

30. Herniter, J.D. (1971) "A Probabilistic Model of Purchasing Timing and Brand Selection", **Mang. Sci., 18** ,pp 102-13

31. ---------- (1971) "An Entropy Model of Brand Purchase Behaviour", **J.M.R., 10** ,pp 361-75

32. ----------- (1974) "A Comparison of the Entropy Model Model and the Henry Model", **J.M.R., 11** ,pp 21-29

33. Hildebrand, F.B. (1976) **Advanced Calculus for Applications,** Englewood Cliffs, N.J.: Prentice-Hall Inc.

34. Hogg, R.V. & Craig, A.T. (1978) **Introduction to Mathematical Statistics,** 4th Edition,New York: Macmillan Publishing Co.,Inc.

35. Ishii, G. & Hayakawa, R. (1960) "On the Compound Binomial Distribution", **Ann. Inst. Statist. Math., 12** ,pp 69-80

36. Jewell, W.S. (1959) "The Properties of Recurrent-event Processes", **O.R., 8** ,pp 446-72

37. Johnson, N.L. & Kotz, S. (1969) **Discrete Distributions,** Boston: Houghton Mifflin Company

38. -------------- (1978) **Urn Models and Their Applications,** New York: John Wiley & Sons,Inc.

39. Kalwani, M.U. (1980) "Maximum Likelihood Estimation of Zero-order Models Given Variable Numbers of Purchase Per Household", **J.M.R., 17** ,pp 547-51

40. Kalwani, M.U. & Morrison, D.G. (1977) 'Estimating the Proportion of "always Buy" and "never buy" Consumers: A Likelihood Ratio Test with Sample Size Implications', **J.M.R., 14** ,pp 601-6

41. ------------------ (1980) "Sample Size Requirements for Zero-Order Models", **J.M.R., 17** ,pp 221-7

42. Kapur, J.N. & Bhatt, S.K. (1983) "Derivation of Some Purchase Incidence Stochastic Models from Entropy-Maximization Principle", **Proceedings of the Eleventh Conference of the Administrative Science Association of Canada, 4 ,Part 2,** Vancouver, Canada

43. Kemp, C.D. & Kemp, A.W. (1956) "Generalized Hypergeometric Distributions", **J. Roy. Statist. Soc., B , 18** ,pp 202-211

44. -------------- (1974) "A Family of Discrete Distributions Defined Via their Factorial Moments", **Communications In Statistics, 3(12)** ,pp 1187-1196

45. Kendall, M. & Stuart, A. (1977) **The Advanced Theory of Statistics, Volume 1,** 4th Edition, New York: Macmillan Publishing Co.,Inc.

46. Kleinman, J.C. (1973) "Proportions with Extraneous Variance : Single and Independent Samples", **J. Amer. Statist. Ass, 68** , pp 46-54

47. Lawrence, Raymond J. (1980) "The Lognormal Distribution of Buying Frequency Rates", **J.M.R., 17,** pp 212-220

48. ------------------ (1981) "Comment on 'Modeling Consumer Purchase Events: A Reply to Lawrence'", **J.M.R., 18** ,pp 470-72

49. Massy, W.F.,Montgomery, D.B. & Morrison, D.G. (1970) **Stochastic Models of Buying Behavior,** Massachusetts: The M.I.T. Press

50. Mood, A.M., Graybill, F.A. & Boes, D.C. (1974) **Introduction to the Theory of Statistics,** 3rd Edition,New York: McGraw-Hill Book Company

51. Morrison, D.G. (1969) "Conditional Trend Analysis : A Model that Allows for Nonusers", **J.M.R., 6** , pp 342-6

52. -------------- (1981) "Modeling Consumer Purchase Events: A Reply to Lawrence", **J.M.R., 18** ,pp 465-469

53. Morrison, D.G. & Schmittlein (1981) "Predicting Future Random Events Based on Past Performance", **Mang. Sci., 27** ,pp 1006-1025

54. Mosimann, J.E. (1962) "On the Compound multinomial distribution,the Multivariate ß-Distribution and Correlations among Proportions", **Biometrika, 49** ,pp 65-82

55. Patil, G.P. & Joshi, S.W. (1968) **A Dictionary and Bibliography of Discrete Distributions**, Edinburgh: Oliver and Boyd LTD

56. Paull, A.E. (1978) "A Generalized Compound Poisson Model for Consumer Purchase Panel Data Analysis", **J. Amer. Statist. Ass., 73** ,pp 706-13

57. Pyatt, G. (1969) "A Model of Brand Loyalties", **Proc 1968 CEIR/SCS Symposium on Model-building in Business and Industry**, London:Griffin

58. Sabavala, D.J. & Morrison, D.G. (1977) "A Model of TV Show Loyalty", **Journal of Advertising Research , 17** , pp 35-43

59. Sharma, S. (1981) "The Theory of Stochastic Preferences:Further Comment and Clarifications", **J.M.R., 18** ,pp 364-9

60. Shimizu, R. (1968) "Generalized Hypergeometric Distributions", **Proc. Inst. Statist. Math., 16** , pp 147-165, (in Japanese)

61. Sibuya, M. (1979) "Generalized Hypergeometric Digamma and Trigamma Distributions", **Ann. Inst. Statist. Math., 31 , Part A** ,pp 373-390

62. Sibuya, M. & Shimizu, R. (1981) "The Generalized Hypergeometric Family of Distributions", **Ann. Inst. Statist. Math., 33 , part A** ,pp 177-190

63. Slater, L.J. (1966) **Generalized Hypergeometric Functions**, London: Cambridge University Press

64. Skellam, J.G. (1947) "A Probability Distribution Derived from the Binomial Distribution by Regarding the Probability of Success as Variable between the Sets of Trials", **J.R. Statist. Soc., B, 10** ,pp 257-261

65. Tallis, G.M. & Young S.S.Y. (1962) "Maximum Likelihood Estimation of Parameters of the Normal, Log-normal, Truncated normal and Bivariate Normal Distributions from Grouped Data", **Australian Journal of Statistics, 4** ,pp 49-54

66. Tripathi, R.C. & Gurland, J. (1976) "Estimation in the 2- and 3-parameter Binomial ^ Beta distribution", **Technical Report #460,** Department of Statistics, University of Wisconsin, Madison.

67. ----------------- (1979) "Some Aspects of the Kemp Families of Distributions", **Commun. Statist. - Theor. Meth., A8(9)** ,pp 855-869

68. Walters, C.G. **Consumer Behavior : Theory and Practice,** London: Irwin-Dorsey international

69. Williamson, E. & Bretherton, M.H. (1963) **Tables of the Negative Binomial Probability Distribution,** London : John Wiley & Sons Limited

Appendix A

THE EMPIRICAL FINDINGS

TABLE A1.1

BBD fitted to the first set of data

Number of Purchases	Frequencies O	E	E
0	1612	1612	1613.4
1	164	155.8	154.8
2	71	74	73.5
3	47	44.4	44.2
4	28	29.4	29.3
5	17	20.6	20.5
6	12	14.9	14.9
7	12	11	11
8	5	8.3	8.3
9	7	6.4	6.4
10	6	4.9	4.9
11 11-12	3 6	3.8 6.8	3.8 6.8
12	3	3	3.0
13	5	2.4	2.4
14	0 5	1.9 7	1.9 7
15 13-16	0	1.5	1.5
16	0	1.2	1.2
17+	8	4.5	5

Method of Estimation		Mean and zeros	MLE
Parameters:			
n		182	182
a(estimated)		0.1133	0.1123
b(estimated)		32.3086	31.9654
variance	4.5025	4.0661	4.1092
d.f.		11	11
chi-square		7.0596	6.2018
p-value		0.7942	0.8596

O = observed

E = expected

TABLE A1.2

BBD fitted to the first set of data
with different values of n by 'mean and zeros'

Number of Purchases		Frequencies O		E		E		E	
0		1612		1612		1612		1612	
1		164		148.8		154.8		155.8	
2		71		73		73.8		74	
3		47		45.3		44.5		44.4	
4		28		30.9		29.6		29.4	
5		17		22.2		20.8		20.6	
6		12		16.5		15.1		14.9	
7		12		12.4		11.2		11	
8		5		9.5		8.5		8.3	
9		7		7.3		6.5		6.4	
10		6		5.9		5		4.9	
11	11-12	3	6	4.3	7.6	3.9	6.9	3.8	6.8
12		3		3.3		3		3	
13		5		2.5		2.4		2.4	
14	13-16	0	5	1.9	6.8	1.9	7	1.9	7
15		0		1.4		1.5		1.5	
16		0		1		1.2		1.2	
17+		8		1.8		4.3		4.5	

Method of Estimation		Mean & zeros	Mean & zeros	Mean & zeros
Parameters:				
n		26	91	182
a(estimated)		0.1034	0.1117	0.1133
b(estimated)		4.1234	15.8709	32.3086
variance	4.5025	3.5882	3.9784	4.0661
d.f.		11	11	11
chi-square		28.7161	7.8199	7.0596
p-value		0.0025	0.7293	0.7942

TABLE A1.3

BBD fitted to the first set of data
with different values of n by MLE

Number of Purchases	Frequencies O	E	E	E
0	1612	1616.3	1611.7	1613.4
1	164	140.4	154.9	154.8
2	71	69.8	73.9	73.5
3	47	44	44.6	44.2
4	28	30.6	29.6	29.3
5	17	22.5	20.8	20.5
6	12	17	15.1	14.9
7	12	13.1	11.2	11
8	5	10.3	8.5	8.3
9	7	8.1	6.5	6.4
10	6	6.4	5	4.9
11 } 11-12	3 } 6	5.1 } 9.1	3.9 } 6.9	3.8 } 6.8
12	3	4	3	3
13	5	3.1	2.4	2.4
14 } 13-16	0 } 5	2.5 } 8.9	1.9 } 6.9	1.9 } 7
15	0	1.9	1.5	1.5
16	0	1.4	1.1	1.2
17+	8	3.5	4.4	5
Method of Estimation		MLE	MLE	MLE
Parameters:				
n		26	91	182
a(estimated)		0.0954	0.1103	0.1123
b(estimated)		3.5635	15.6004	31.9654
variance	4.5025	4.2031	4.0511	4.1092
d.f.		11	11	11
chi-square		18.7842	7.5176	6.2018
p-value		0.065	0.7558	0.8596

Table A1.4

CNBD fitted to the first set of data

Number of Purchases		Frequencies O		E	
0		1612		1612	
1		164		166.8	
2		71		70.1	
3		47		41.3	
4		28		27.3	
5		17		19.1	
6		12		14	
7		12		10.5	
8		5		8	
9		7		6.2	
10		6		4.8	
11	11-12	3	6	3.8	6.8
12		3		3	
13		5		2.4	
14	13-16	0	5	2	7.3
15		0		1.6	
16		0		1.3	
17+		8		5.8	

Method of Estimation		Mean & Zeros
parameters:		
m (estimated)		0.636
k (estimated)		0.0993
variance	4.5025	4.4262
d.f.		11
chi-square		4.7756
p-value		0.9415

TABLE A1.5

Lognormal Distribution fitted to the first set of data

Number of Purchases		Frequencies O		Frequencies E	
0		(1612)		(290)	
1		164		159.3	
2		71		79.5	
3		47		45.2	
4		28		28	
5		17		18.5	
6		12		12.8	
7		12		9.2	
8		5		6.8	
9		7		5.2	
10		6		4	
11	11-12	3	6	3.1	5.6
12		3		2.5	
13	13-16	5	5	2	6.1
14		0		1.6	
15		0		1.4	
16		0		1.1	
17+		8		7.8	

Method of Estimation		MLE
Parameters:		
μ(estimated)		0.2088
σ^2(estimated)		1.3276
variance	14.5462	15.8753
d.f.		10
chi-square		4.4746
p-value		0.9234

TABLE A1.6

LSD fitted to the first set of data

Number of Purchases		Frequencies O		E	
0		1612		-	
1		164		165.5	
2		71		72	
3		47		41.7	
4		28		27.2	
5		17		19	
6		12		13.7	
7		12		10.2	
8		5		7.8	
9		7		6	
10		6		4.7	
11	11-12	3	6	3.7	6.7
12		3		3	
13		5		2.4	
14	13-16	0	5	1.9	7.2
15		0		1.6	
16		0		1.3	
17+		8		6.3	

Method of Estimation		MLE
Parameters:		
q(estimated)		0.8699
variance	14.5462	14.4513
d.f.		11
chi-square		4.1992
p-value		0.9379

TABLE A1.7

MNBD fitted to the first set of data

Number of Purchases	Frequencies O	E	E
0	1612	1708.7	1612
1	164	66.6	41.4
2	71	51.4	37
3	47	39.7	33
4	28	30.6	29.5
5	17	23.6	26.4
6	12	18.2	23.6
7	12	14.0	21.1
8	5	10.8	18.8
9	7	8.3	16.8
10	6	6.4	15
11 11-12	3 6	5.0 8.8	13.4 25.4
12	3	3.8	12
13	5	2.4	9.6
14	0	1.7	8
15 13-16	0 5	1.2 6.2	7 30.8
16	0	.9	6.2
17+	8	6.7	69.2

Method of Estimation		Method of Moments	Mean and Zeros
Parameters:			
μ(estimated)		0.2092	0.2448
q(estimated)		0.7525	0.8810
variance	4.5025	4.5025	28.6578
d.f.		11	11
chi-square		165.9159	525.0944
p-value		0	0

TABLE A1.8

NBD fitted to the first set of data

Number of Purchases		O		Frequencies E		E	
0		1612		1612		1613.2	
1		164		156.8		156.1	
2		71		74.1		73.7	
3		47		44.2		44.1	
4		28		29.2		29.1	
5		17		20.3		20.3	
6		12		14.7		14.7	
7		12		10.9		10.9	
8		5		8.2		8.2	
9		7		6.2		6.3	
10		6		4.8		4.8	
11	11-12	3	6	3.8	6.7	3.8	6.8
12		3		2.9		3	
13		5		2.3		2.3	
14	13-16	0	5	1.8	6.8	1.9	6.9
15		0		1.5		1.5	
16		0		1.2		1.2	
17+		8		5.1		4.9	

Method of Estimation		Mean and zeros	MLE
Parameters:			
m(estimated)		0.636	0.636
k(estimated)		0.1149	0.1142
variance	4.5025	4.1564	4.1780
d.f.		11	11
chi-square		5.681	6.0804
p-value		0.8938	0.8679

TABLE A1.9

A comparison of NBD, LSD, Lognormal Distribution, CNBD, MNBD and BBD fitted to the first set of data

Number of Purchases	O	NBD E	LSD E	Frequencies Lognormal E	CNBD E	MNBD E	BBD E
0	1612	1613.2	-	-	1612	1708.7	1613.4
1	164	156.1	165.5	159.3	166.8	66.6	154.8
2	71	73.7	72	79.5	70.1	51.4	73.5
3	47	44.1	41.7	45.2	41.3	39.7	44.2
4	28	29.1	27.2	28	27.3	30.6	29.3
5	17	20.3	19	18.5	19.1	23.6	20.5
6	12	14.7	13.7	12.8	14	18.2	14.9
7	12	10.9	10.2	9.2	10.5	14	11
8	5	8.2	7.8	6.8	8	10.8	8.3
9	7	6.3	6	5.2	6.2	8.3	6.4
10	6	4.8	4.7	4	4.8	6.4	4.9
11-12	6	6.8	6.7	5.6	6.8	8.8	6.8
13-16	5	6.9	7.2	6.1	7.3	6.2	7
17+	8		6.3	7.8	5.8	6.7	5
Method of Estimation		MLE	MLE	MLE	Mean & Zeros	Method of Moments	MLE
Parameters:		m=.636 k=.01142	q=.8699	μ=.2088 σ^2=1.3276	m=.636 k=.0993	μ=.2448 q=.881	n=182 a=.1123 b=31.9654
variance	4.5025	4.178	-	-	4.4262	4.5025	4.1092
d.f.		11	11	10	11	11	11
chi-square		6.0804	4.1992	4.4746	4.7756	165.9159	6.2018
p-value		0.8679	0.9379	0.9234	0.9415	0	0.8596

TABLE A2.1

BBD fitted to the second set of data

Number of Purchases	Frequencies O	E	E
0	4848	4848	4830.7
1	1788	1776.5	1781.3
2	1025	1059.2	1063.4
3	685	712.4	715.5
4	480	506.2	508.3
5	374	371.1	372.6
6	293	277.6	278.6
7	229	210.6	211.3
8	164	161.4	161.8
9	125	124.7	124.9
10	125	96.9	97
11	85	75.7	75.7
12	69	59.3	59.3
13	44	46.7	46.6
14	48	36.8	36.7
15	26	29.1	29
16	27	23	23
17	21	18.3	18.2
18	23	14.5	14.4
19	16	11.5	11.5
20 } 20+	10 } 10	9.2 } 45.5	9.1 } 45.2
21+	0	36.3	36.1

Method of Estimation		Mean and zeros	MLE
Parameters:			
n		365	365
a(estimated)		0.4433	0.4466
b(estimated)		77.5684	78.0122
variance	10.6342	11.5632	11.5293
d.f.		18	18
chi-square		56.4242	56.794
p-value		0	0

TABLE A2.2

BBD fitted to the second set of data
with different values of n by 'Mean and Zeros'

Number of Purchases	Frequencies O	E	E
0	4848	4848	4848
1	1788	1776.5	1747.2
2	1025	1059.2	1049.1
3	685	712.4	712.5
4	480	506.2	511.3
5	374	371.1	378.3
6	293	277.6	285.2
7	229	210.6	217.7
8	164	161.4	167.6
9	125	124.7	129.8
10	125	96.9	101.0
11	85	75.7	78.7
12	69	59.3	61.5
13	44	46.7	48.1
14	48	36.8	37.6
15	26	29.1	29.4
16	27	23	23
17	21	18.3	17.9
18	23	14.5	14
19	16	11.5	10.9
20 } 20+	10 } 10	9.2 } 45.5	8.4 } 36.2
21+	0	36.3	27.8

Method of Estimation		Mean and zeros	MLE
Parameters:			
n		365	73
a(estimated)		0.4433	0.4275
b(estimated)		77.5684	14.5859
variance	10.6342	11.5632	11.0994
d.f.		18	18
chi-square		56.4242	44.6955
p-value		0	0.0005

TABLE A2.3

BBD fitted to the second set of data
with different values of n by MLE

Number of Purchases	Frequencies O	E	E	E
0	4848	4830.7	4847.2	4885.8
1	1788	1781.3	1749.4	1677.0
2	1025	1063.4	1050.3	1017.9
3	685	715.5	713.2	704.5
4	480	508.3	511.6	516.1
5	374	372.6	378.3	389.6
6	293	278.6	285.1	299.2
7	229	211.3	217.5	232.1
8	164	161.8	167.4	181.0
9	125	124.9	129.5	141.4
10	125	97.0	100.7	110.5
11	85	75.7	78.5	66.1
12	69	59.3	61.2	66.8
13	44	46.6	47.9	51.5
14	48	36.7	37.4	39.4
15	26	29	29.1	29.8
16	27	23	22.8	22.4
17	21	18.2	17.8	16.5
18	23	14.4	13.8	12.0
19	16	11.5	10.8	8.6
20	10	9.1	8.4	6.0
21+	0	36.1	27.1	30.8
20+	10	45.2	35.5	36.8

Method of Estimation		MLE	MLE	MLE(except n)
Parameters:				
n		365	73	30*
a(estimated)		0.4466	0.4283	0.3919
b(estimated)		78.0122	14.6413	5.2529
variance	10.6342	11.5293	11.0476	10.3971
d.f.		18	18	17
chi-square		56.7940	44.9805	64.0205
p-value		0	0.0004	0

* n=30 is estimated by Skellam's method

TABLE A2.4

CNBD fitted to the second set of data

Number of Purchases		Frequencies O		Frequencies E	
0		4848		4848	
1		1788		1924.7	
2		1025		1026.3	
3		685		669.6	
4		480		471.4	
5		374		345.5	
6		293		259.7	
7		229		198.7	
8		164		154	
9		125		120.5	
10		125		95	
11		85		75.4	
12		69		60.1	
13		44		48.2	
14		48		38.7	
15		26		31.2	
16		27		25.3	
17		21		20.5	
18		23		16.6	
19		16		13.6	
20	20+	10	10	11.1	62
21+		0		50.9	

Method of Estimation		Mean and Zeros
Parameters:		
m(estimated)		2.0785
k(estimated)		0.3729
variance	10.6342	12.7109
d.f.		18
chi-square		84.3922
p-value		0

Table A2.5

Lognormal distribution fitted to the second set of data

Number of Purchases	Frequencies O	Frequencies E
0	4848	-
1	1788	1657.6
2	1025	1166.9
3	685	778.7
4	480	527.9
5	374	367.2
6	293	262.1
7	229	191.3
8	164	142.5
9	125	108.1
10	125	83.3
11	85	65
12	69	51.4
13	44	41.1
14	48	33.2
15	26	27
16	27	22.2
17	21	18.4
18	23	15.3
19	16	12.8
20 } 20+	10 } 10	10.8 } 85
21	0	74.2
Method of Estimation		MLE
Parameters:		
μ(estimated)		0.871
σ^2(estimated)		0.904
variance	12.8721	20.7158
d.f.		17
chi-square		172.3709
p-value		0

TABLE A2.6

LSD fitted to the second set of data

Number of Purchases	Frequencies O	E
0	4848	-
1	1788	2224.5
2	1025	998.9
3	685	598.1
4	480	402.9
5	374	289.5
6	293	216.6
7	229	166.8
8	164	131.1
9	125	104.6
10	125	84.6
11	85	69.1
12	69	56.9
13	44	47.1
14	48	39.3
15	26	32.9
16	27	27.7
17	21	23.5
18	23	19.9
19	16	16.9
20 } 20+	10 } 10	14.4 } 106.1
21+	0	91.7

Method of Estimation		MLE
Parameters:		
q(estimated)		0.8981
variance	12.8722	22.9790
d.f.		18
chi-square		317.7138
p-value		0

TABLE A2.7
MNBD fitted to the second set of data

Number of Purchases	Frequencies O	E
0	4848	5323.7
1	1788	1183.2
2	1025	927.8
3	685	722.6
4	480	559.5
5	374	431
6	293	330.5
7	229	252.3
8	164	191.9
9	125	145.5
10	125	109.9
11	85	82.8
12	69	62.2
13	44	21.5
14	48	7.3
15	26	4.1
16	27	3
17	21	2.1
18	23	1.5
19	16	0.9
20 } 20+	10 } 10	0.7 } 141.7
21+	0	141.0

Method of Estimation		Method of Moments
Parameters:		
μ(estimated)		1.001
q(estimated)		0.673
variance	10.6342	10.6342
d.f.		18
chi-square		1812.3133
p-value		0

TABLE A2.8

NBD fitted to the second set of data

Number of Purchases		Frequencies O		E		E	
0		4848		4848		4826.9	
1		1788		1780.8		1788.7	
2		1025		1060.1		1066.3	
3		685		711.6		715.9	
4		480		504.7		507.5	
5		374		369.5		371.2	
6		293		276.1		277.1	
7		229		209.4		209.9	
8		164		160.4		160.6	
9		125		123.9		123.9	
10		125		96.4		96.2	
11		85		75.4		75.1	
12		69		59.2		58.9	
13		44		46.6		46.4	
14		48		36.9		36.6	
15		26		29.2		29	
16		27		23.2		23	
17		21		18.5		18.3	
18		23		14.8		14.6	
19		16		11.8		11.6	
20		10		9.4		9.3	
21	20+	0	10	7.6	48.5	7.4	47.3
22		0		6.1		6	
23+		0		25.4		24.6	

Method of Estimation		Mean and zeros	MLE
Parameters:			
m(estimated)		2.0785	2.0785
k(estimated)		0.4462	0.4510
variance	10.6342	11.7606	11.6576
d.f.		18	18
chi-square		59.0377	59.9236
p-value		0	0

TABLE A2.9

A Comparison of NBD, LSD, Lognormal distribution CNBD, MNBD and BBD fitted to the second set of data

Number of Purchases	O	NBD E	LSD E	LOGNORMAL E	CNBD E	MNBD E	BBD E
				Frequencies			
0	4848	4826.9	-	-	4848	5323.7	4830.7
1	1788	1788.7	224.5	1657.6	1924.7	1183.2	1781.3
2	1025	1066.3	998.9	1166.9	1026.3	927.8	1063.4
3	685	715.9	598.1	778.7	669.6	722.6	715.5
4	480	507.5	402.9	527.9	471.4	559.5	508.3
5	374	371.2	289.5	367.2	345.5	431	372.6
6	293	277.1	216.6	262.1	259.7	330.5	278.6
7	229	209.9	166.8	191.3	198.7	252.3	211.3
8	164	160.6	131.1	142.5	154	191.9	161.8
9	125	123.9	104.6	108.1	120.5	145.5	124.9
10	125	96.2	84.6	83.3	95	109.9	97
11	85	75.1	69.1	65	75.4	82.8	75.7
12	69	58.9	56.9	51.4	60.1	62.2	59.3
13	44	46.4	47.1	41.1	48.2	21.5	46.6
14	48	36.6	39.3	33.2	38.7	7.3	36.7
15	26	29	32.9	27	31.2	4.1	29
16	27	23	27.7	22.2	25.3	3	23
17	21	18.3	23.5	18.4	20.5	2.1	18.2
18	23	14.6	19.9	15.3	16.6	1.5	14.4
19	16	11.6	16.9	12.8	13.6	0.9	11.5
20+	10	47.3	106.1	85	62	141.7	45.2

		NBD	LSD	LOGNORMAL	CNBD	MNBD	BBD
Method of Estimation		MLE	MLE	MLE	Mean & Zeros	Method of moments	MLE
Parameters:		m=2.0785 k=0.451	q=.8981	μ=.871 σ^2=.904	m=2.0785 k=0.3729	μ=1.001 q=0.673	n=365 a=.4466 b=78.0122
variance	10.6342	11.6576	-	-	12.7109	10.6342	11.5293
d.f.		18	18	17	18	18	18
chi-square		59.9236	317.7138	172.3709	84.3922	1812.3133	56.794
p-value		0	0	0	0	0	0

TABLE A2.10

A Comparsion of NBD, LSD, Lognormal distribution CNBD, NMBD and BBD fitted to the light-buyers of the Second Set of data

Number of Purchases	O	NBD E	LSD E	Lognormal E	CNBD E	MNBD E	BBD E
				Frequencies			
0	4848	4826.9	-	-	4848	5323.7	4830.7
1	1788	1788.7	2224.5	1657.6	1924.7	1183.2	1781.3
2	1025	1066.3	998.9	1166.9	1026.3	927.8	1063.4
3	685	715.9	598.1	778.7	669.6	722.6	715.5
4	480	507.5	402.9	527.9	471.4	559.5	508.3
5	374	371.2	289.5	367.2	345.5	431	372.6
6	293	277.1	216.6	262.1	259.7	330.5	278.6
7	229	209.9	166.8	191.3	198.7	252.3	211.3
8	164	160.6	131.1	142.5	154	191.9	161.8
9	125	123.9	104.6	108.1	120.5	145.5	124.9

	NBD	LSD	Lognormal	CNBD	MNBD	BBD
Method of Estimation	MLE	MLE	MLE	Mean & Zeros	Method of moments	MLE
parameters:	m=2.0785 k=0.451	q=0.8981	μ=0.871 σ^2=0.904	m=2.0785 k=0.3729	μ=1.001 q=0.673	n=365 a=0.4466 b=78.0122
d.f.	7	7	6	7	7	7
chi-square	7.2692	200.7553	60.2195	22.2804	395.9789	6.6118
P value	0.4014	0	0	0.0022	0	0.4705

TABLE A3.1

BBD fitted to the third set of data

Number of Purchases		Percentages O		E		E	
0		79.3		79.3		78.9	
1		8.4		10.4		10.8	
2		5.1		4.5		4.6	
3		3.0		2.4		2.4	
4		3.6		1.4		1.3	
5		.2		.8		.8	
6	6+	.4	.4	.5	1.2	.5	1.2
7+		0		.7		.7	

Method of Estimation		Mean and zeros	MLE
Parameters:			
n		28	28
a(estimated)		.1768	.1884
b(estimated)		10.7268	11.5442
variance	1.1219	1.4598	1.3805
d.f.		4	4
chi-square		5.0551h	5.7922h

100h is the number of households analysed.

TABLE A3.2

BBD fitted to the third set of data
with different values of n by 'mean and zeros'

Number of Purchases	Percentages O	E	E	E	E
0	79.3	79.3	79.3	79.3	79.3
1	8.4	10.4	10.2	9.5	8
2	5.1	4.5	4.5	4.7	4.8
3	3.0	2.4	2.5	2.8	3.8
4	3.6	1.4	1.4	1.8	4.1
5	.2	0.8	0.9	1.1	0 } 4.1
6	.4	0.5 } 1.2	0.5 } 1.2	0.6 } .8	0
7+	0	0.7	0.7	0.2	0
Method of Estimation		Mean & Zeros	Mean & Zeros	Mean & Zeros	Mean & Zeros
Parameters:					
n		28	14	7	4
a(estimated)		.1768	.1635	.1365	.094
b(estimated)		10.7268	4.8784	1.9679	.7342
variance	1.1219	1.4598	1.3844	1.2452	1.0629
d.f.		4	4	4	2
chi-square		5.0551h	5.0326h	2.912h	.2096h

100h is the number of households analysed.

TABLE A3.3

BBD fitted to the third set of data
with different values of n by MLE

Number of Purchases	Percentages O	E	E	E	E	E
0	79.3	78.9	79	79.1	79.2	79.3
1	8.4	10.8	10.6	10	9.3	8.1
2	5.1	4.6	4.7	4.8	4.9	4.8
3	3.0	2.4	2.5	2.8	3.1	3.8
4	3.6	1.3	1.4	1.7	2.1	3.9
5	.2	.8	.8	1.0	1.4	0
6	.4	.5 } 1.2	.5 } 1	.5 } .7	0 } 1.4	0 } 3.9
7+	0	.7	.5	.2	0	0

Method of Estimation		MLE	MLE	MLE	MLE (except n)	MLE
Parameters:						
n		28	14	7	5*	4
a(estimated)		.1884	.1759	.1477	.124	.0962
b(estimated)		11.5442	5.3488	2.1935	1.2861	.7613
variance	1.1219	1.3805	1.2914	1.1567	1.0666	1.0419
d.f.		4	4	4	3	2
chi-square		5.7922h	4.8589h	3.1816h	1.6272h	.2214h

* n=5 is estimated by Skellam's method

100h is the number of households.

TABLE A3.4

CNBD fitted to the third set of data

Number of Purchases	Percentages O		Percentages E	
0	79.3		79.3	
1	8.4		11.1	
2	5.1		4.2	
3	3.0		2.2	
4	3.6		1.2	
5	.2		0.7	
6 } 6+	.4	} .4	0.5	} 1.3
7+	0		0.8	

Method of Estimation		Mean and zeros
Parameters:		
m(estimated)		0.454
k(estimated)		0.1517
variance	1.1219	1.626
d.f.		4
chi-square		6.9207h

100h is the number of households analysed

TABLE A3.5

Lognormal Distribution fitted to the third set of data

Number of Purchases	Percentages O	E
0	79.3	-
1	8.4	7.7
2	5.1	6.6
3	3.0	3.4
4	3.6	1.6
5	.2	.7
$\left.\begin{matrix}6\\7+\end{matrix}\right\}$ 6+	$\left.\begin{matrix}.4\\0\end{matrix}\right\}$ .4	$\left.\begin{matrix}.3\\.4\end{matrix}\right\}$ .7

Method of Estimation		MLE
Parameters:		
μ(estimated)		0.8114
σ^2(estimated)		0.2723
variance	1.6052	2.0823
d.f.		3
chi-square		3.4373h

100h is the number of households analysed.

TABLE A3.6

LSD fitted to the third set of data

Number of Purchases	Percentages O	Percentages E
0	79.3	-
1	8.4	11.1
2	5.1	4.2
3	3.0	2.1
4	3.6	1.2
5	.2	.7
6 6+	.4 .4	.5 1.4
7+	0	.9

Method of Estimation		MLE
Parameters:		
q(estimated)		.7554
variance	1.6052	4.1564
d.f.		4
chi-square		7.1068h

100h is the number of households analysed

TABLE A3.7

MNBD fitted to the third set of data

Number of Purchases	Percentages O	E
0	79.3	77
1	8.4	11.6
2	5.1	5.8
3	3.0	2.9
4	3.6	1.4
5	.2	.7
6 } 6+	.4 } .4	.3 } .6
7+	0	.3

Method of Estimation		Method of Moment
Parameters:		
μ(estimated)		0.6172
q(estimated)		0.4282
variance	1.1219	1.1219
d.f.		4
chi-square		4.9203h

100h is the number of households analysed

TABLE A3.8
NBD fitted to the third set of data

Number of Purchases	Percentages O	E	E
0	79.3	79.3	78.9
1	8.4	10.6	11
2	3.1	4.5	4.6
3	3.0	2.3	2.3
4	3.6	1.3	1.3
5	.2	.8	.8
6 } 6+	.4 } .4	.5 } 1.3	.5 } 1.1
7+	0	.8	.6

Method of Estimation		Mean & Zeros	MLE
parameters:			
m(estimated)		0.454	0.454
k(estimated)		0.19	0.2004
variance	1.1219	1.5388	1.4825
d.f		4	4
chi square		5.982h	5.8486h

100h is the number of households analysed

TABLE A3.9

A Comparison of NBD, LSD, Lognormal distribution, CNBD, MNBD AND BBD fitted to the third set of data

Number of Purchases	O	NBD E	LSD E	Lognormal E	CNBD E	MNBD E	BBD E
	Percentages						
0	79.3	78.9	-	-	79.3	77	78.9
1	8.4	11	11.1	7.7	11.1	11.6	10.8
2	5.1	4.6	4.2	6.6	4.2	5.8	4.6
3	3.0	2.3	2.1	3.4	2.2	2.9	2.4
4	3.6	1.3	1.2	1.6	1.2	1.4	1.3
5	.2	.8	.7	.7	.7	.7	.8
6	.4	.5	.5	.3	.5	.3	.5
7+	0	.6	.9	.4	.8	.3	.7
6+	.4	1.1	1.4	.7	1.3	.6	1.2

Method of Estimation		MLE	MLE	MLE	Mean & Zeros	Method of Moment	MLE
Parameters:		m=.454 k=.2004	q=.7554	μ=.8114 σ^2=.2723	m=.454 k=.1517	μ=.6172 q=.4282	n=28 a=.1884 b=11.5442
Variance	1.1219	1.4825	-	-	1.626	1.1219	1.3805
d.f.		4	4	3	4	4	4
chi-square		5.8486h	7.1068h	3.4373h	6.9207h	4.9203h	5.7922h

100h is the number of households analysed.

TABLE A4.1
BBD fitted to the fourth set of data

Number of Purchases	Frequencies O	E	E
0	382	382	380.9
1	193	193.9	195.1
2	81	82.7	83
3	37	32.6	32.5
4	11	12.2	12
5	5	4.4	4.3
6	0	1.5	1.4
7 5+	1 6	.7 6.6	.8 6.5
8+	0	0	0

Method of Estimation		Mean & Zeros	MLE
parameters:			
n		30	30
a(estimated)		1.3949	1.4304
b(estimated)		53.4227	54.7901
variance	1.1074	1.1305	1.1209
d.f		3	3
chi square		0.8056	0.8188
p-value		0.8481	0.845

TABLE A4.2

BBD fitted to the fourth set of data
with different values of n by 'mean and zeros'

Number of Purchases	Frequencies O	E	E	E	E
0	382	382	382	382	382
1	193	193.9	191.4	187.9	185.5
2	81	82.7	84.6	87.3	89.1
3	37	32.6	34	36.3	38
4	11	12.2	12.4	12.7	12.8
5	5	4.4	4.1	3.4	2.7
6 } 5+	0 } 6	1.5 } 6.6	1.2 } 5.7	0.5 } 3.9	0 } 2.7
7	1	.7	.3	0	0
8+	0	0	0	0	0

Method of Estimation		Mean & Zeros	Mean & Zeros	Mean & Zeros	Mean & Zeros
parameters:					
n		30	10	6	5
a(estimated)		1.3949	1.1279	0.9365	0.8431
b(estimated)		53.4227	13.6484	6.4239	4.6789
variance	1.1074	1.1305	1.1073	1.0647	1.0436
d.f		3	3	3	3
chi square		0.8056	0.6051	1.9649	5.3524
p value		0.8481	0.8953	0.5797	0.1477

TABLE A4.3

BBD fitted to the fourth set of data
with different values of n by MLE

Number of Purchases	O	Frequencies E	E	E	E	E
0	382	380.9	382	382.4	383.9	384.8
1	193	195.1	192.2	191	185.8	182.4
2	81	83	84	84.4	86.7	88.2
3	37	32.5	33.6	34	36.5	38.4
4	11	12	12.3	12.5	13.0	13.4
5	5	4.3	4.2	4.1	3.6	2.9
6	0 6	1.4 6.5	1.3 5.8	1.2 5.7	0.6 4.2	0 2.9
7	1	0.8	0.3	0.4	0	0
8+	0	0	0	0	0	0
Method of Estimation		MLE	MLE	MLE	MLE	MLE
Parameters:						
n		30	12*	10	6	5
a		1.4304	1.2046	1.1357	0.9028	0.8003
b		54.7901	17.7328	13.7405	6.1941	4.4420
Variance	1.1074	1.1209	1.1091	1.1049	1.0775	1.0612
d.f.		3	2	3	3	3
chi square		0.8188	0.5988	0.6188	1.7491	5.0188
P value		0.845	0.7413	0.8921	0.626	0.1704

*n = 12 is estimated by Skellam's method.

TABLE A4.4
CNBD fitted to the fourth set of data

Number of Purchases		Frequencies O		Frequencies E	
0		382		382	
1		193		201.2	
2		81		75.6	
3		37		30.1	
4		11		12.3	
5		5		5.1	
6	5+	0	6	2.1	8.8
7		1		0.9	
8+		0		0.7	

Method of Estimation		Mean & Zeros
parameters:		
m(estimated)		0.7634
k(estimated)		0.7951
variance	1.1074	1.204
d.f		3
chi square		3.3299
p value		0.3435

TABLE A4.5

Lognormal distribution fitted to the fourth set of data

Number of Purchases	Frequencies O	E
0	382	-
1	193	190.9
2	81	88.2
3	37	31.1
4	11	10.9
5	5	4.0
6	0 6	1.6 6.9
7 5+	1	0.7
8+	0	0.6
Method of Estimation		MLE
parameters:		
μ (estimated)		0.4682
σ^2 (estimated)		0.289
variance	0.928	1.1412
d.f.		2
chi square		1.8485
p value		0.3968

TABLE A4.6

LSD fitted to the fourth set of data

Number of Purchases		Frequencies O		Frequencies E	
0		382		-	
1		193		213.3	
2		81		64.7	
3		37		26.1	
4		11		11.9	
5		5		5.8	
6	5+	0	6	2.9	12
7		1		1.5	
8+		0		1.8	

Method of Estimation		MLE
parameters:		
q (estimated)		0.6065
variance	0.928	1.4687
d.f.		3
chi square		13.6586
p value		0.0034

TABLE A4.7

MNBD fitted to the fourth set of data

Number of Purchases		Frequencies O		E	
0		382		380.8	
1		193		193.6	
2		81		84.8	
3		37		33	
4		11		11.9	
5		5		4	
6	5+	0	6	1.3	5.9
7		1		0.4	
8+		0		0.2	

method of Estimation		method of moments
parameters:		
μ(estimated)		3.388
q(estimated)		0.1839
variance	1.1074	1.1074
d.f.		3
chi square		0.7305
P value		0.866

TABLE A4.8

NBD fitted to the fourth set of data

Number of Purchases	Frequencies O	E	E
0	382	382	380.3
1	193	194.9	196.6
2	81	82	82.5
3	37	32.1	32
4	11	12.1	11.9
5	5	4.4	4.3
6	0	1.6	1.5
7	1	0.6	0.5
8+	0	0.3	0.4
5+	6	6.9	6.7

Method of Estimation		mean & zeros	MLE
parameters:			
m (estimated)		0.7634	0.7634
k (estimated)		1.5379	1.601
variance	1.1074	1.1423	1.1274
d.f.		3	3
chi square		1.0232	0.9961
P value		0.7956	0.8022

TABLE A4.9

A Comparison of NBD, LSD, Lognormal distribution, CNBD, MNBD AND BBD fitted to the fourth set of data

Number of Purchases	O	NBD E	LSD E	Lognormal E	CNBD E	MNBD E	BBD E
				Frequencies			
0	382	380.3	-	-	382	380.8	380.9
1	193	196.6	213.3	190.9	201.2	193.6	195.1
2	81	82.5	64.7	88.2	75.6	84.8	83
3	37	32	26.1	31.1	30.1	33	32.5
4	11	11.9	11.9	10.9	12.3	11.9	12
5	5	4.3	5.8	4.0	5.1	4	4.3
6	0	1.5	2.9	1.6	2.1	1.3	1.4
7	1	.5	1.5	0.7	0.9	0.4	0.8
8+	0	.4	1.8	0.6	0.7	0.2	0
5+	6	6.7	12	6.9	8.8	5.9	6.5

Method of Estimation	MLE	MLE	MLE	Mean & Zeros	Method of Moment	MLE
Parameters:	m=.7634 k=1.601	q=.6065	μ=.4682 σ^2=.289	m=.7634 k=.7951	μ=388 q=.1839	n=30 a=1.4304 b=54.7901
Variance 1.1074	1.1274	-	-	1.204	1.1074	1.1209
d.f.	3	3	2	3	3	3
chi square	1.0232	13.6586	1.8485	3.3299	0.7305	0.819
p value	0.7956	0.0034	0.3968	0.3435	0.866	0.8449

TABLE A5.1

Estimated parameters of BBD fitted to the first set of data with different values of n by different methods of estimation

parameters	method of moments	mean & zeros	MLE
n	182	182	182
a	0.100	0.1133	0.1123
b	28.516	32.3086	31.9654
n	91	91	91
a	0.0956	0.1117	0.1103
b	13.5881	15.8709	15.6004
n	26	26	26
a	0.0733	0.1034	0.0954
b	2.9223	4.1234	3.5635

TABLE A5.2

Estimated parameters of BBD fitted to the second set of data with different values of n by different methods of estimation.

parameters	method of moments	mean and zeros	MLE
n	365	365	365
a	0.4933	0.4433	0.4466
b	86.3118	77.5684	78.0122
n	73	73	73
a	0.4521	0.42755	0.4283
b	15.425	14.5859	14.6413

TABLE A5.3

Estimated parameters of BBD fitted to the third set of data with different values of n by different methods of estimation.

parameters	method of moments	mean & zeros	MLE
n	28	28	28
a	0.2734	0.1768	0.1884
b	16.5857	10.7268	11.5442
n	14	14	14
a	0.2389	0.1635	0.1759
b	7.127	4.8784	5.3488
n	7	7	7
a	0.1721	0.1365	0.1477
b	2.4809	1.9679	2.1935
n	4	4	4
a	0.077	0.094	0.0962
b	0.6013	0.7342	0.7613

TABLE A5.4

Estimated parameters of BBD fitted to the fourth set of data with different values of n by different methods of estimation.

parameters	method of moments	mean and zeros	MLE
n	30	30	30
a	1.4851	1.3949	1.4304
b	56.878	53.4227	54.7901
n	10	10	10
a	1.1461	1.1279	1.1357
b	13.6468	13.6484	13.7405
n	6	6	6
a	0.8336	0.9365	0.9028
b	5.718	6.4239	6.1941
n	5	5	5
a	0.8336	0.8431	0.8003
b	3.9127	4.6789	4.442

TABLE A6.1

A comparison of BBD and NBD about the the predictions in the second period based on the first set of data

	BBD	NBD
b_1	0.1933	0.1934
b_2	0.2473	0.2481
$b_N = b_L$	0.054	0.0546
b_R	0.1394	0.1388
$m_N = m_L$	0.077	0.0781
m_R	0.559	0.5579
$w_N = w_L$	4.0109	4.0191
w_R	1.4274	1.4298
method of estimation:	MLE	MLE
estimated parameters:	n*=182 a=0.1123 b=31.9654	m=0.636 k=0.1142

* n is assumed known.

TABLE A6.2

A comparison of BBD and NBD about the predictions in the second period based on the second set of data.

	BBD	NBD
b_1	0.5402	0.5405
b_2	0.6485	0.6494
$b_N = b_L$	0.1083	0.1089
b_R	0.4319	0.4316
$m_N = m_L$	0.169	0.1703
m_R	1.9095	1.9082
$w_N = w_L$	4.4211	4.421
w_R	1.5609	1.5637
method of estimation:	MLE	MLE
estimated parameters:	n*=365 a=0.4466 b=78.0122	m=2.0785 k=0.451

* n is assumed known.

TABLE A6.3

A comparison of BBD and NBD about the predictions in the second period based on the third set of data

	BBD	BBD	NBD
b_1	0.2068	0.2108	0.2111
b_2	0.2552	0.2871	0.2902
$b_N = b_L$	0.0484	0.0763	0.0791
b_R	0.1584	0.1345	0.1321
$m_N = m_L$	0.0628	0.1048	0.1097
m_R	0.3912	0.3492	0.3443
$w_N = w_L$	2.4701	2.5967	2.6071
w_R	1.2973	1.3734	1.3873
method of estimation:	MLE	MLE	MLE
estimated parameters:	n*=4 a=0.0962 b=0.7613	n*=28 a=0.1884 b=11.5442	m=0.454 k=0.2004

* n is assumed known.

TABLE A6.4

A comparison of BBD and NBD about the predictions in the second period based on the fourth set of data.

	BBD	NBD
b_1	0.4635	0.4643
b_2	0.6518	0.6577
$b_N = b_L$	0.1883	0.1934
b_R	0.2751	0.2709
$m_N = m_L$	0.267	0.2769
m_R	0.4964	0.4865
$w_N = w_L$	1.8041	1.7958
w_R	1.4179	1.4316
method of estimation:	MLE	MLE
estimated parameters:	n*=30 a=1.4304 b=47.9007	m=0.7634 k=1.601

* n is assumed known.