- '/8

ESTIMATING ZERO ORDER MODELS
GIVEN VARIABLE NUMBER OF PURCHASE
RECORDS PER HOUSEHOLD*

Manohar U. Kalwani**

ESTIMATING ZERO ORDER MODELS
GIVEN VARIABLE NUMBER OF PURCHASE
RECORDS PER HOUSEHOLD*

Manohar U. Kalwani**

WP1026-78   October 1978

**Assistant Professor of Management Science
  Sloan School of Management
  Massachusetts Institute of Technology

ABSTRACT

Assuming a zero order process, the beta-binomial distribution model is often used to describe consumers' brand switching behavior. A variety of methods have been used to estimate the parameters of the beta-binomial distribution, each of which assumes <u>fixed number of purchase occasions across households</u>. Consumer panel data, however, are typically available for a fixed duration of time (say, a year) during which consumers invariably make different number of purchases in the product class. The purpose of this paper is to present a method for estimating zero order models from panel data containing different number of purchase records per household. Empirical data are presented to demonstrate the efficiency of this estimation method.

INTRODUCTION

Marketing researchers frequently use two different approaches to present analytical results from consumer panel data: (i) the negative binomial distribution (NBD) model (see [2]), and (ii) the beta-binomial distribution (BBD) model (see [5,8]). The first method involves the construction of a frequency distribution of purchases in the sample who buy 0,1,2,..., or in general, r units of an item (a particular pack-size of a brand) in some <u>specified time period</u>. This frequency distribution information is then used to obtain two key measures of market response to this item, namely, penetration and purchase frequency. Penetration is defined as the proportion of people who buy this item at least once in the given time period. Purchase frequency is the average number of times these buyers purchase the item in the specified time period (see [2] for further details). Assuming a zero order process[1], the BBD model is used to report the proportion of households in the sample who buy a brand a given number of times out of a <u>fixed number of purchase occasions</u>. Additionally, the BBD model is employed to describe consumers' brand switching behavior. Since each person is defined by a single parameter, p. these zero order models are completely defined by specifying the parameters of the mixing beta distribution. The conditional and unconditional switching and repeat purchase probabilities representing measures of market response to the brand can be easily obtained from a knowledge of these parameters (see [4,5,8]). A variety of methods have been used to estimate the parameters of the BBD model, each of which

assumes fixed number of purchase occasions across consumers (see [1,3,4,7]),
Consumer panel data, however, are commonly available for fixed durations of
time (say, a year) during which households invariably make a different number
of purchases in the product class. The purpose of this paper is to present
a method for estimating zero order models from panel data containing
differing number of purchase records per household. Empirical data are pre-
sented to demonstrate the efficiency of this estimation method.

THE METHOD

In the BBD model, each consumer is assumed to buy a given brand (say,
Brand 1) with probability, p, or some brand in the aggregate "all other"
class with probability, 1-p. Further, each consumer is assumed to make a
fixed number of purchases in the product class. The model allows for
heterogeneity in the population by letting the p values differ across in-
dividuals according to a beta distribution. The marginal distribution of
the number of successes (i.e., purchases of Brand 1), x, on k Bernoulli
trials (i.e., purchases of the product class) is then beta-binomial with

$$P(x|\alpha,\beta,k) = \binom{k}{x} \frac{B(\alpha+x,\beta+k-x)}{B(\alpha,\beta)} , \qquad 0 < p < 1; \alpha,\beta > 0, \qquad (1)$$

where α and β are the parameters of the beta distribution and B(.) denotes
the beta function. The mean and variance of the beta-binomial distribution
are given by

$$E[x] = k\frac{\alpha}{\alpha+\beta} ,$$

$$Var[x] = \frac{k\alpha\beta(k+\alpha+\beta)}{(\alpha+\beta)^2(\alpha+\beta+1)} . \qquad (2)$$

As indicated earlier, a variety of methods have been used to estimate the parameters of the beta-binomial distribution, each of which assumes a fixed number of trials across consumers (see [1,3,4,7]). The maximum likelihood estimation procedure presented in Kalwani and Morrison [3,4] can be extended to the case of variable number of trials per household. Denote the number of purchase records of household j by k_j, where j=1,2,...N. Then, assuming independence between k_j and household j's probability to purchase Brand 1, the probability of household j making x_j purchases of Brand 1 is given by

$$P(x_j|\alpha,\beta,k_j) = \binom{k_j}{x_j} \frac{B(\alpha+x_j,\beta+k_j-x_j)}{B(\alpha,\beta)} . \qquad (3)$$

The likelihood function for the N households can be written as a product of N probabilities using equation (3). Ignoring the constant terms, this likelihood function for the overall sample data is given by

$$\ell(x_j\text{'s}|\alpha,\beta,k_j\text{'s}) = \prod_{j=1}^{N} \frac{B(\alpha+x_j,\beta+k_j-x_j)}{B(\alpha,\beta)} .$$

Taking logarithms and expressing the beta functions in terms of the gamma functions, the above expression simplifies to the following log-likelihood function

$$L(x_j\text{'s}|\alpha,\beta,k_j\text{'s}) = \sum_{j=1}^{N} \left[\sum_{r=0}^{x_j-1} \log(\alpha+r) + \sum_{r=0}^{k_j-x_j-1} \log(\beta+r) - \sum_{r=0}^{k_j-1} \log(\alpha+\beta+r) \right] . \qquad (4)$$

In many zero order applications, the knowledge of a brand's market share $\mu=\alpha/(\alpha+\beta)$ and loyalty index, $\phi=1/(1+\alpha+\beta)$ is adequate (see [4]). The conditional as well as unconditional repeat and switching probabilities can be obtained from a knowledge of these two parameters. It is shown in [4] that the transformed parameters, μ and ϕ, are computationally easier to estimate than the original parameters since their maximum like-lihood estimates have smaller standard errors than the maximum likelihood estimates of α and β. Substituting for α and β in terms of μ and ϕ, the log likelihood function in equation (4) can be rewritten as:

$$L(x_j\text{'s}|\mu,\phi,k_j\text{'s}) = \sum_{j=1}^{N}\left[\sum_{r=0}^{x_j-1}\log(\mu(1-\phi)+r\phi)+\sum_{r=0}^{k_j-x_j-1}\log((1-\mu)(1-\phi)+r\phi)\right.$$

$$\left. - \sum_{r=0}^{k_j-1}\log(1-\phi+r\phi)\right] . \tag{5}$$

The maximum likelihood estimates of μ and ϕ (or α and β) can be obtained by maximizing the log-likelihood function in equation (5) (or equation (4)) through numerical optimization.

Another way of representing the log-likelihood function in equation (5) (or equation (4)) involves classifying the households in the panel according to number of purchases of the product class, k, where k=1,2,...K. The log-likelihood function in equation (5) now becomes

$$L(n_x^k|\mu,\phi,k\text{'s}) = \sum_{k=1}^{K}\left[\sum_{x=0}^{k}n_x^k\left\{\sum_{r=0}^{x-1}\log(\mu(1-\phi)+r\phi) + \sum_{r=0}^{k-x-1}\log((1-\mu)(1-\phi)+r\phi)\right.\right.$$

$$\left.\left. - \sum_{r=0}^{k-1}\log(1-\phi+r\phi)\right\}\right] , \tag{6}$$

where n_x^k is the number of consumers who make x purchases of Brand 1 given
that they make a total of k purchases in the product class. The numerical
optimization procedure based on maximizing the log-likelihood function in
equation (6) is computationally much faster than that based on equation
(5) since the inside expression containing the three summations is es-
timated many fewer times. Empirical data are presented below to illustrate
the application of the above estimation method.

ILLUSTRATIVE APPLICATION

The data presented below are from an MRCA (Market Research Corporation
of America) consumer panel. They represent margarine purchases of 2496 house-
holds during the fourth quarter of 1976. The frequency distribution denoting
the number of households with margarine purchases ranging from 1 to 20 is
displayed in Table 1. Examination of Table 1 reveals that methods based on

----

INSERT TABLE 1 HERE

----

fixed number of purchases per household would be inefficient in that they
would ignore considerable purchase information. Consider for instance, the
commonly used sequence length of 5 purchase records. Estimation methods based
on fixed number of purchase occasions would utilize only five purchase records
of each of 1161 households (or 46.5% of 2496 households) who made more than
five purchases of margarine. This would result in the exclusion of two
types of purchase information: (i) all purchase records of 1335 (=2496-1161)
households with four or fewer purchases of margarine, and (ii) a number of
purchase records of those households with more than five purchases of margarine.
The maximum likelihood estimation method presented in this paper on the other
hand, would utilize all purchase records of all the 2496 households in the
panel. Does this full utilization of purchase information translate into

increased efficiency for the estimation method based on variable number of purchase records per houshold? If yes, how much is the gain in efficiency?

Recall that under rather general regularity conditions (see [6, pp. 295-302]), maximum likelihood estimates are best asymptotically normal (BAN). That is, they are consistent, asyptotically normal, and asymptotically efficient. It is shown in [6] that BAN estimators obey the "inverse square root of n relationship"; that if, for large sample sizes, standard error of a BAN estimator is reduced by $1/\sqrt{n}$ with an n-fold increase in sample size. Assuming that the maximum likelihood functions presented in this paper satisfy the regularity conditions discussed in [6], the standard errors of maximum likelihood estimates using the estimation method based on variable number of purchase records would be $\sqrt{\dfrac{n_F}{n_V}}$ times the standard errors of maximum likelihood estimates based on fixed number of purchase records per household where n_F and n_V denote the corresponding sample sizes.

What are the values of n_F and n_V for the illustrative application presented in the section? The answer to this question is not clear cut. One set of values can be obtained by aggregating households with their purchase records to obtain total number of purchase records being utilized for parameter estimation. Thus, the estimation method based on five purchase occasions per household utilizes 5805 (= 1161 X 5) purchase records while the estimation method which allows for variable number of margarine purchase records. This approach to computation of sample sizes would imply a 31% $\left(\text{i.e., } 1 - \sqrt{\dfrac{5805}{12236}}\right)$ reduction in the standard error of estimate. The inadequacy of this approach becomes evident when one considers the extreme case of a sample of one household with a very large number of purchase records. Such an extreme case provides no information on heterogeneity in brand

purchase probabilities.

The second set of values can be obtained by weighting each household equally irrespective of the number of its margarine purchases. Recall that the parameters that are being estimated - α, β or μ, ϕ - refer to the mixing distribution of purchase probabilities (assumed to be beta distribution in this paper). Hence, margarine purchases for a given household represent replications and the aggregation of households forms the sample size. This approach to the computation of sample sizes would imply a 32% $\left(\text{i.e., } 1 - \sqrt{\frac{1161}{2496}} \right)$ reduction in the standard error of estimate.

It should be noted that the lack of a clear cut approach to the computation of sample size does not question the very claim of improved efficiency with the use of the new estimation method. It merely suggests that the exact gain in efficiency will depend on the interpretation of sample size. Note, that simulated purchase data can be used to obtain accurate estimates of reductions in the standard errors of estimates.

The estimation procedure presented above was employed to obtain maximum likelihood estimates of α, β and μ, ϕ for Parkay margarine. The results obtained therefrom are given below:

$$\hat{\alpha} = 0.28677, \quad \hat{\beta} = 2.09998,$$
$$\hat{\mu} = 0.12015 \quad \hat{\phi} = 0.29456.$$

Since maximum likelihood estimates are invariant, we can check the accuracy of the numerical optimization procedure by comparing $\hat{\alpha}/(\hat{\alpha}+\hat{\beta})$ with $\hat{\mu}$ and $1/(1+\hat{\alpha}+\hat{\beta})$ with $\hat{\phi}$. The results of this comparison are given below:

$$\frac{\hat{\alpha}}{\hat{\alpha}+\hat{\beta}} = 0.12015, \quad \frac{1}{1+\hat{\alpha}+\hat{\beta}} = 0.29527.$$

and they reveal accuracy to the third decimal place.

CONCLUSION

The methods used to estimate zero order models assume fixed number of purchase occasions across households. Panel data, however, are normally available for a fixed duration of time and contain different number of purchase records per household. In this paper, we presented a maximum likelihood method for estimating the parameters of zero order models given variable number of purchase occasions per household. Empirical data from a MRCA panel were presented to illustrate the application of this estimation method.

FOOTNOTE

1. In zero order models each consumer is assumed to pur-
   chase a given brand (say, Brand 1) with probability,
   p, and Brand 0 representing the aggregate "all other"
   class with the complementary probability, 1-p. These
   models allow for heterogeneity in the population by
   letting the p values differ across individuals.

Table 1

NUMBER OF HOUSEHOLDS WITH DIFFERENT NUMBER
OF PURCHASES OF MARGARINE

NUMBER OF PURCHASES OF MARGARINE	NUMBER OF HOUSEHOLDS	NUMBER OF PURCHASES OF MARGARINE	NUMBER OF HOUSEHOLDS
1	329	11	57
2	371	12	50
3	344	13	30
4	291	14	16
5	271	15	11
6	215	16	8
7	177	17	4
8	138	18	2
9	105	19	2
10	74	20	1

Total Number of Households = 2496.

Total Number of Purchase Records = 12,236.

REFERENCES

1. Chatfield, C., and G.J. Goodhardt. "The Beta-Binomial Model for Consumer Purchasing Behavior," Applied Statistics, 19 (June 1970), 240-50.

2. Ehrenberg, A.S.C., Repeat Buying. Amsterdam: North Holland Publishing Company, 1972.

3. Kalwani, M.U., and Donald G. Morrison. "Estimating the Proportion of 'Always Buy' and 'Never Buy' Consumers: A Likelihood Ratio Test with Sample Size Implications," Journal of Marketing Research, 14 (November 1977), 601-6.

4. Kalwani, M.U., and Donald G. Morrison. "Sample Size Requirements for Zero Order Models," Working Paper, Sloan School of Management, M.I.T. (October 1978).

5. Massy, W.F., D.B. Montgomery, and D.G. Morrison. Stochastic Models of Buyer Behavior. Cambridge, MA: M.I.T. Press, 1970.

6. Rao, C.R. Linear Statistical Inference and Its Applications, New York: John Wiley & Sons, Inc., 1965.

7. Shenton, L.R. "Maximum Likelihood and the Efficiency of Fitting by Moments," Biometrika, 37 (June 1950), 111-6.

8. Wierenga, B. An Investigation of Brand Choice Processes. Rotterdam: University of Rotterdam Press, 1974.