

## A Brand Choice Model Using Multinomial Logistics Regression, Bayesian Inference and Markov Chain Monte Carlo Method

**Deshmukh Sachin, Manjrekar Pradip\* and Gopal R.**

<sup>1</sup>YMT College of Management, Kharghar, Navi Mumbai, 410210 and Ph.D. Scholar (Business Management), Department of Business Management, Padmashree Dr. D. Y. Patil University, Navi Mumbai, 410614

<sup>2</sup>Professor & Head, Research & Consultancy & Extension Centre, Department of Business Management, Padmashree Dr. D. Y. Patil University, Navi Mumbai, 410614,  
[drpradipm@gmail.com](mailto:drpradipm@gmail.com)

<sup>3</sup>Director, Department of Business Management, Padmashree Dr. D. Y. Patil University, Navi Mumbai, 410614

**Abstract-** This paper explains a Brand Choice model using Multinomial Logistics Regression(MNL), Bayesian Inference and Markov Chain Monte Carlo (MCMC) Simulation method. The model enables a marketer to forecast probability of choosing a particular brand by a consumer at a purchase occasion. Three hundred and twenty households in one tier (Mumbai, Pune), two tier (Kolhapur, Nasik, Nagpur), and three tier (Parbhani, Latur, Amravati) cities/towns of Maharashtra participated in the survey. Data were collected on three leading detergent brands – Surf, Ariel and Rin. The basis for our brand choice model is the multinomial logistics regression which was then converted into a software program. The parameters of the model are then estimated with the help of WinBUGS (Windows Version for Bayesian Inference Using Gibbs Sampling) software that uses Bayesian Analysis and Markov Chain Monte Carlo method. The utility function in the model used five attributes – a dummy term for brand2 and brand3 each, price term giving the price in rupees, a ‘feature’ term (advertisement) that was “1” when the respondents were exposed to the advertisement between two purchases, otherwise a “0” was assigned, the promotion term taking values “1” or “0” when the product was specially promoted/not promoted. The calculated parameters are further used to estimate the probability for a brand which is priced at Rs. 55. A marketer can thus forecast the probability of purchase of a brand at different price levels.

**Key Words :** Brand Choice Model, Multinomial Logistics Regression(MNL), Bayesian Inference, Markov Chain Monte Carlo (MCMC) Simulation, WinBUGS Software.

### Introduction

Consumers' choice is a dynamic process. Whether the marketer is observing routine purchase of brands or once in life time choices of housing, choices made in the present reflects what has been learned from the past. Yet while the importance of understanding the nature of dynamic influences on choice is widely recognized, our knowledge of how choice evolves over time remains highly fragmented. The view of decision making that pervades most of the literature on choice theory remains largely static, quite adept at describing the momentary relationships that exist between preferences and actions, but less able to describe the dynamics that gave rise to these relationships. Since the first analyses of systematic analyses of consumer panel data in the early 1960's, it has been widely

observed that the probability that a household will choose a particular brand or product on a choice occasion is conditioned by the choice made on the previous occasion (e.g., Kuehn, 1962). These early observations spawned the development of a large contemporary literature of statistical models for discrete panel-data analysis, all offering approaches to representing temporal dependencies in choice probabilities (e.g., Guadagni and Little 1983; Heckman 1981; Keane 1996; Massy, Montgomery, and Morrison 1970; Papatla and Krishnamurthi 1992). Although this literature is diverse, most formulations follow a common approach: household brand-choices are initially represented by a static or cross-sectional random-utility model, and dynamics are then introduced by allowing preferences to be state-dependent, such as

evolving as a Markov process (e.g., Guadagni and Little 1983). Consumers are then assumed to act as intuitive utility theorists, on each occasion choosing the option that has the highest expected utility, and then using Bayes' rule to update prior beliefs about the utility distributions of each option. Consumer

develops actual attributes value profile of the brand. One can consider the calculation of probability of choosing a brand in a particular product category on the basis of this attributes profile. Multinomial logistics regression model can be considered for the calculation of this probability.

**The Multinomial Logistics Regression (MNL) Model**

The Multinomial Logistics Model is used when there is a choice between two or more alternatives. Let us consider only two alternatives (products) from which a choice has to be made. The consumer will select that product that will help him to maximize his utility (Ben – Akiva & Lerman 1985). For a consumer, the utility of product i, from a set of I products and described on a set of “A” attributes as  $x_{i1}, x_{i2}, \dots, x_{iA}$  is as follows.

$$U_i = \sum b_a x_{ia} + e_i$$

Where  $e_i$  is the error term. The consumer chooses the option “i” which maximizes the utility and the probability that the consumer chooses i is as follows.

$$P(U_i > U_j, \text{ for all } j \neq i) = p \left[ \sum b_a x_{ia} + e_i > \sum b_a x_{ja} + e_j, \text{ for all } j \neq i \right]$$

If we assume that stochastic component  $e_1, e_2, \dots, e_i$  are drawn from extreme value distribution, the choice probability is given by

$$P(\text{choice} = i) = \frac{\exp(X_i \beta_k)}{1 + \sum_{k=1} \exp(X_i \beta_k)}$$

Where  $X_i$  are the explanatory variables and  $\beta_k$  are unknown parameters which are estimated by the likelihood.

**Bayesian Analysis**

A probability model relates observed data y to a set of unknown parameters  $\theta$ , with the possible inclusion of fixed, known covariates x. Our data are usually a collection of observations indexed  $i = 1, \dots, n$ :  $y = \{y_1, y_2, \dots, y_n\}$ . The observed data need not be scalars; in fact, they can be anything supported by the probability model, including vectors and matrices. The probability model has k parameters, which are represented  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ . The covariates, or independent variables, are typically a collection of column vectors:  $x = \{x_1, x_2, \dots, x_n\}$ . The probability model can be written  $f(y|\theta, x)$ , or, suppressing the conditioning on the covariates,  $f(y|\theta)$ . It is important to stress the importance of choosing appropriate probability models. While canonical models

exist for certain types of dependent variables, the choice of model is rarely innocuous, and is thus something that should be tested for adequacy.

The purpose of statistical inference is to learn about parameters that characterize the data generating process given observed data. In the conventional, frequentist approach to statistical inference, one assumes that the parameters are fixed, unknown quantities, and that the observed data y are a single realization of a repeatable process, and can thus be treated as random variables. The goal of the conventional approach is to produce estimates of these unknown parameters. These estimates are denoted  $\theta^\wedge$  (theta cap). The most common way to obtain these

estimates is by the method of maximum likelihood.

One maximizes the likelihood function  $L(\cdot)$  with respect the parameters to obtain the maximum likelihood estimates; i.e., the parameter values most likely to have produced the observed data. To perform inference about the parameters, the frequentist recognizes that the estimated parameters  $\hat{\theta}$  result from a single sample, and uses the sampling distribution to compute standard errors, perform hypothesis tests, construct confidence intervals, and the like.

When performing Bayesian inference, the foundational assumptions are quite different. The unknown parameters  $\theta$  are treated as random variables, while the observed data  $y$  are treated as fixed, known quantities. (Both the Bayesian and frequentist approach treat the covariates  $x$  as fixed, known quantities). These assumptions are much more intuitive; the unobservable parameters are treated probabilistically, while the observed data are treated deterministically. Indeed, the quantity of interest is the distribution of the parameter  $\theta$  after having observed the data  $y$ . This posterior distribution can be written  $f(\theta|y)$ , and can be computed using Bayes' Theorem:

$$f(\theta|y) = \frac{f(y|\theta)f(\theta)}{f(y)}$$

One of the three quantities on the right-hand side of the above equation is familiar:  $f(y|\theta)$  is the likelihood function dictated by the probability model. The second expression in the numerator  $f(\theta)$  is called the prior distribution. This distribution contains all ex ante information about the parameter values available to the researcher before observing the data. Oftentimes researchers use non-informative (or minimally informative parameters) such that the amount of prior information included in the analysis is small. The denominator of the above equation contains the prior predictive distribution:

$$f(y) = \int_{\theta} f(y|\theta)f(\theta)d\theta$$

While this quantity is useful in some settings, such as model comparison, most of

the time researchers work up to a constant of proportionality  $f(\theta|y) \propto f(y|\theta)f(\theta)$ . The posterior distribution, in essence, translates the likelihood function into a proper probability distribution over the unknown parameters, which can be summarized just as any probability distribution; by computing expected values, standard deviations, quantiles, and the like. What makes this possible is the formal inclusion of prior information in the analysis.

**Model Fitting Via Simulation**

It is impossible to analytically summarize the posterior distribution. For the last twenty years, Bayesian statisticians have used Monte Carlo method (Metropolis and Ulam, 1949) to perform the summarization numerically. The Monte Carlo method is based on a simple idea: one can learn anything about a target Distribution by repeatedly drawing from it and empirically summarizing those draws.

To use the Monte Carlo method to summarize posterior distributions, it is necessary to have algorithms that are well-suited to producing draws from commonly found target distributions. Two algorithms—the Gibbs sampling and Metropolis-Hastings Algorithms— have proven to be very useful for applied Bayesian work. Both of these algorithms are Markov chain Monte Carlo methods, which means that the draws  $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(g)}$  are dependent, each draw  $\theta^{(g+1)}$  depends only on the previous draw  $\theta^{(g)}$ . The sequence of draws thus forms a Markov chain. Algorithms are constructed such that the Markov chain converges to the target density (its steady state) regardless of the starting values.

The Gibbs sampling algorithm (Geman and Geman, 1984; Gelfand and Smith, 1990) uses a sequence of draws from conditional distributions to characterize the joint target distribution. Suppose that our parameter vector  $\theta$  has three components, making our target distribution  $f(\theta_1, \theta_2, \theta_3|y)$ . To use the Gibbs sampler, one begins by choosing starting values  $\theta^{(0)}_2$  and  $\theta^{(0)}_3$  (starting values are usually chosen near the posterior mode or the maximum likelihood Estimates). One then repeats, for  $g = 1, \dots, G$  iterations (making sure to store the sequences of draws at each iteration):

Draw  $\theta_1^{(g)}$  from  $f(\theta_1|\theta_2^{(g-1)}, \theta_3^{(g-1)}, \mathbf{y})$

Draw  $\theta_2^{(g)}$  from  $f(\theta_2|\theta_1^{(g)}, \theta_3^{(g-1)}, \mathbf{y})$

Draw  $\theta_3^{(g)}$  from  $f(\theta_3|\theta_1^{(g)}, \theta_2^{(g)}, \mathbf{y})$

Since we are always conditioning on past draws, the resultant sequence results in a Markov chain. When computing Monte Carlo estimates of quantities of interest, like the posterior mean, one discards the first set of "burn-in" iterations to ensure the chain has reached steady state. For the posterior distribution of many common models, these conditional distributions take known forms; e.g., multivariate normal, truncated normal, Gamma, etc. So, while the joint posterior distribution is difficult to simulate from directly, it is easy to simulate from these conditionals.

The second algorithm that enjoys common use in applied Bayesian statistics, is the Metropolis-Hastings algorithm, first introduced by Metropolis et al. (1953) and generalized by Hastings (1979). Chib and Greenberg (1995) provide an accessible introduction to this algorithm. The algorithm has many applications beyond Bayesian

statistics; it is commonly used for all sorts of numerical integration and optimization. (It is also the case that the Gibbs sampling algorithm is a special case of the Metropolis-Hastings algorithm.) To simulate from our target distribution  $f(\theta|\mathbf{y})$ , we again start with sensible starting values:  $\theta^{(0)}$ . For each iteration of the simulation  $g = 1, \dots, G$ , we draw a proposal  $\theta^*$  from a known proposal distribution  $p_g(\theta^*|\theta^{(g-1)})$ . One chooses a proposal distribution from which it is easy to sample, such as a uniform distribution over a particular region, or a multivariate normal or multivariate-t, centered at the current location of the chain, the posterior mode, or perhaps elsewhere. It is important to choose a proposal distribution such that the chain "mixes well"; i.e., adequately explores the posterior distribution. The convergence diagnostics, discussed below, can be used to determine how well the chain is mixing. For each iteration, we set:

$$\theta^{(g)} = \begin{cases} \theta^* & \text{with probability } \alpha^* \\ \theta^{(g-1)} & \text{with probability } (1 - \alpha^*) \end{cases}$$

With  $\alpha^*$  defined:

$$\alpha^* = \min \left\{ \frac{f(\theta^*|\mathbf{y}) p_g(\theta^{(g-1)}|\theta^*)}{f(\theta^{(g-1)}|\mathbf{y}) p_g(\theta^*|\theta^{(g-1)})}, 1 \right\}$$

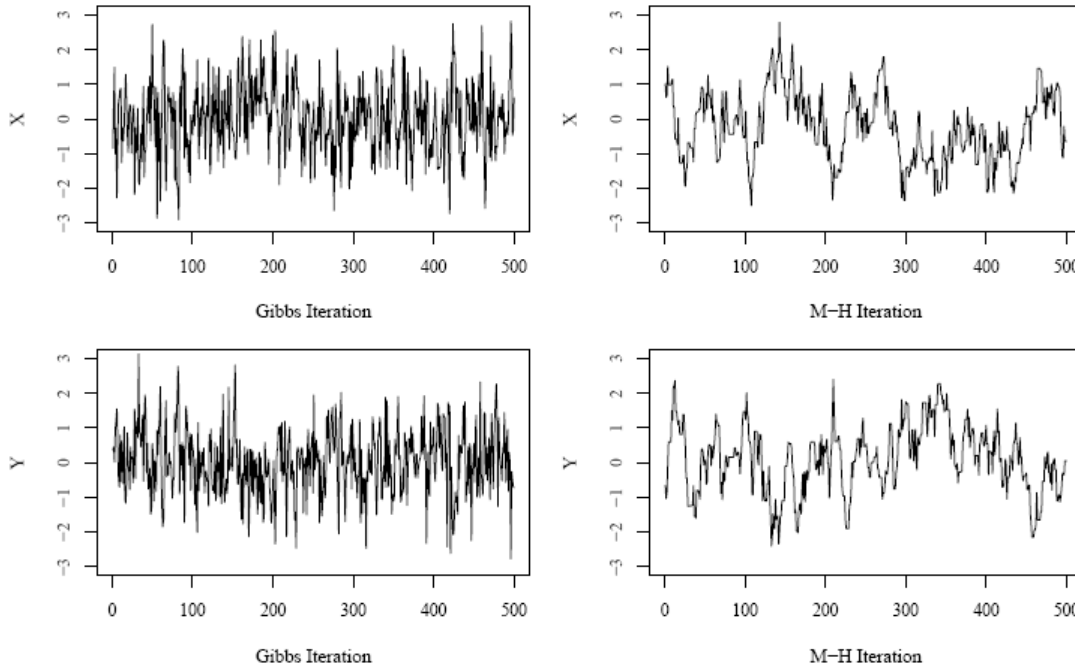
Unlike the Gibbs sampling algorithm, when each move is automatically accepted, one accepts the proposal distribution probabilistically, sometimes moving to a value with a higher density value, sometimes moving to one with a lower density value. Just as with the Gibbs sampling algorithm, the steady state of the

Markov chain characterized by this algorithm is the target distribution, in this case, the posterior distribution. Figure 1 illustrates sampling from the same bivariate normal distribution using Metropolis-Hastings with a uniform random walk proposal with width of two units. In Figure 2 we see that this Metropolis-Hastings sampler traverses the

space more slowly; in fact, 30% of the time the sampler does not move at all. The size of the proposal distribution is also somewhat small, which keeps the sampler in certain parts of the distribution longer than a better conditioned sampler.

Unlike the Gibbs sampling algorithm, when each move is automatically accepted, one accepts the proposal distribution probabilistically, sometimes moving to a location using Metropolis-Hastings with a uniform random walk proposal with width of two units.

value with a higher density value, sometimes moving to one with a lower density value. Just as with the Gibbs sampling algorithm, the steady state of the Markov chain characterized by this algorithm is the target distribution, in this case, the posterior distribution. Figure 1 illustrates sampling from the same bivariate normal distribution



**Fig.1**

Without being able to plot the target distribution, which in most applications is of high dimension, it would be difficult, if not impossible, to assess whether the chain is sampling from the target distribution. An important part of any Bayesian analysis is assessing the convergence of the simulation results. Indeed, no Monte Carlo estimates of posterior density summaries can be trusted unless the chain has reached its steady state. But just as it is impossible to know, in most circumstances, that a numerical optimizer has reached the global maximum likelihood estimate, so, too, is it impossible to know for sure whether a Markov chain Monte Carlo algorithm has converged. However, there are a number of methods that can be used to take output from a MCMC sampler and test whether the

sequence of draws is consistent with convergence (see the review pieces by Cowles and Carlin, 1995; Brooks and Roberts, 1998). Each of these convergence diagnostics is based on one of a number of criteria; some look at the marginal posterior distributions to see if the trace plots are stationary (using a number of different tests). Others compare multiple runs from different starting values and using different random number seeds to determine whether the chains converge in such a way as to make their different starting values irrelevant. It is important to note that finding nonconvergence of any parameter means that the entire chain has not converged, and it needs to be run longer (or re-implemented using a different algorithm). Every analyst should use these diagnostic tools before

computing any posterior density summaries or reporting results, and results from these diagnostics should be presented in research papers.

**The WinBUGS Software**

WinBUGS is a statistical software for Bayesian Analysis using Markov Chain Monte Carlo (MCMC) methods. It is based on BUGS (Bayesian Inference Using Gibbs Sampling). It runs on Microsoft Windows and Linux using Wine. The software has eighteen built in probability distributions. This make our working easier as one is not required to write code for these probability distributions and the MCMC procedure.

However, we will use the most common priors for MNL model parameters which are of the form

**The Likelihood Function**  $\beta_j \sim N(\mu_j, \sigma_j^2)$

In logistics regression, we know that

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p}}$$

So that the maximum likelihood function for the n<sup>th</sup> consumer will be

$$\text{likelihood}_i = \left( \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{y_i} \left( 1 - \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{(1-y_i)}$$

Since the individual consumers are assumed independent of each other, likelihood function over the data set of n consumers is then

$$\text{likelihood} = \prod_{i=1}^n \left[ \left( \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{y_i} \left( 1 - \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{(1-y_i)} \right]$$

**Posterior Distribution via Baye’s Theorem**

The posterior distribution is derived by multiplying the prior distribution over all the parameters by the full likelihood function

$$\begin{aligned} \text{posterior} &= \prod_{i=1}^n \left[ \left( \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{y_i} \left( 1 - \frac{e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip}}} \right)^{(1-y_i)} \right] \\ &\times \prod_{j=0}^p \frac{1}{\sqrt{2\pi}\sigma_j} \exp \left\{ -\frac{1}{2} \left( \frac{\beta_j - \mu_j}{\sigma_j} \right)^2 \right\} \end{aligned}$$

**Data Collection**

Three hundred and twenty five households from the state of Maharashtra participated in the survey. Data collection was done for three purchase occasions of the sampled households which makes a total of nine hundred and seventy five observations collectively. Three leading brands of detergents (Surf, Ariel, Rin) of one kg. pack were considered for our study.

Sr.no	Name of Place	No of families participated
-------	---------------	-----------------------------

WinBUGS is a open source and can be downloaded easily.

**Prior Distribution for MNL Model**

The set of unknown parameter consist of  $\beta_0, \beta_1, \beta_2, \dots \beta_p$ . In general any prior distribution can be used, depending upon the available prior information. The choice can include informative prior distributions if something is known about the likely values of the unknown parameters, or diffused or non-informative prior if either little is known about the coefficient values or if one wishes to see what the data themselves provide as inferences.



1	Beed	25
2	Parhbani	25
3	Kolhapur	25
4	Sholapur	25
5	Aurangabad	25
6	Amravati	25
7	Pune	25
8	Mumbai	75
9	Nagpur	25
10	Thane	25
11	Latur	25

Total	325
-------	-----

### Model Estimation

Our logistics regression model studied the choice made among these three brands. The utility function used five attributes – a dummy term for brand2, a dummy term corresponding to brand3, a price term giving price in rupees, a feature term that was “1” when the product was advertised and the respondent saw the advertisement, and “0”

when the respondent did not see the advertisement in the period between two purchases. A promotion term that was “1” when the product enjoyed special promotion and “0” when it was not.

We estimated the model by expressing it as a WinBUGS code and getting it executed by the software. The WinBUGS code is given below for the reader’s reference.

### WinBUGS software code for the brand choice model

#### Model.

```
{
logit(p[i]) <- alpha + b.brand2*brand2[i] + b.brand3*brand3[i] + b.Price*Price[i]+
b.Feat*Feat[i] +b.Disp*Disp[i]
#likelihood function for each data point

low[i] ~ dbern(p[i])

alpha ~ dnorm(0.0,1.0E-2) #prior for intercept
b.brand2 ~ dnorm(0.0,1.0E-2) #prior for slopes
b.brand3 ~ dnorm(0.0,1.0E-2)
b.Price ~ dnorm(0.0,1.0E-2)
b.Feat ~ dnorm(0.0,1.0E-2)
b.Disp ~ dnorm(0.0,1.0E-2)

pred.price <- exp(alpha + b.brand2 + b.brand3 + b.Price + b.Feat + b.Disp) / (1 + exp(alpha +
b.brand2 + b.brand3 + b.Price + b.Feat + b.Disp))
}
```

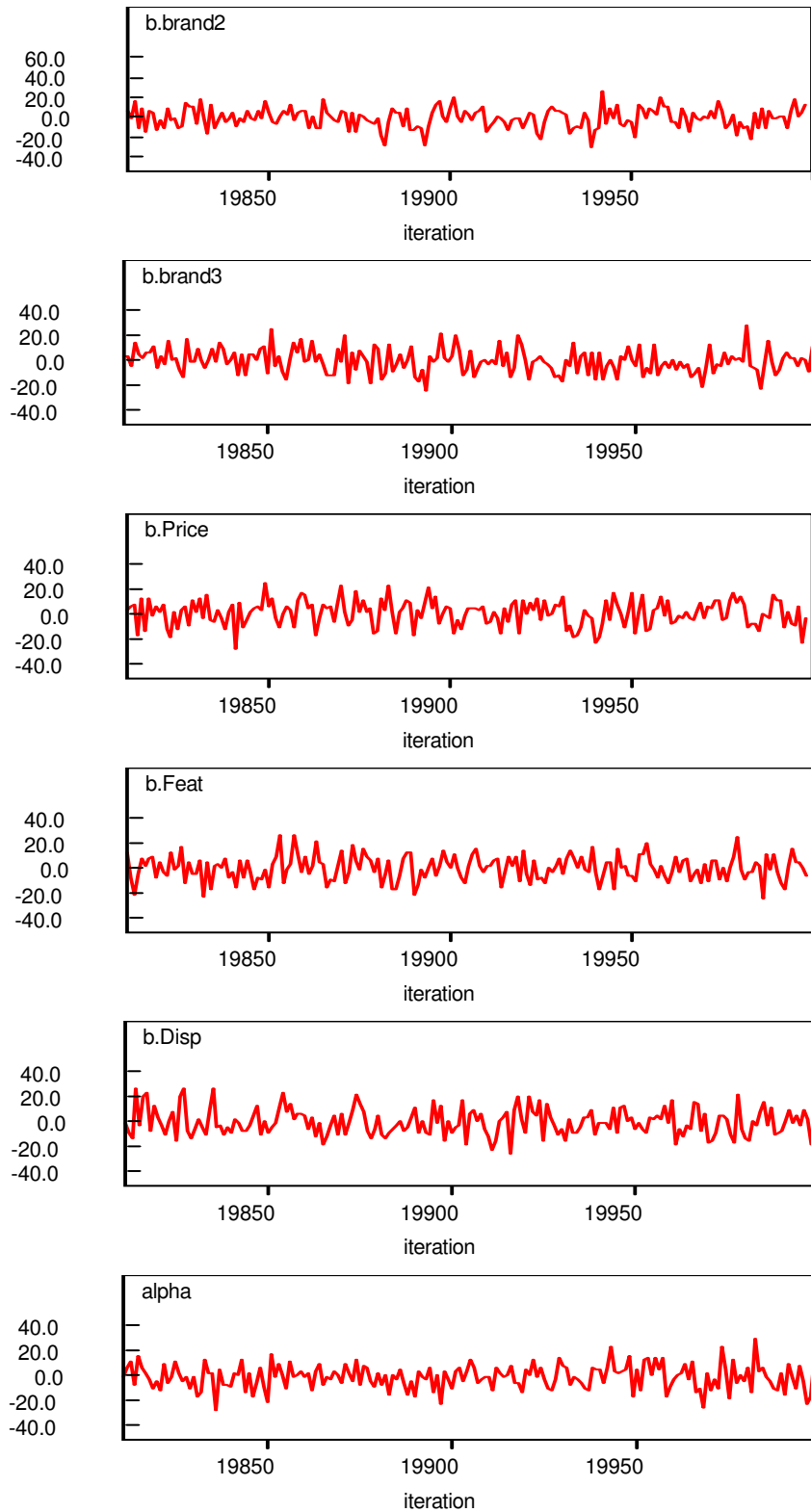
#### Exhibit 1

We ran the software code in WinBUGS using MCMC and discarded the first ten thousand draws and used the next ten thousand draws to assess the convergence and draw inference.

### Findings

1. After discarding the first 10,000 draws, we assessed the data convergence which was found to be satisfactory as shown below

A Brand Choice Model Using Multinomial Logistics Regression



The above graphs indicate that the chain has reached the steady state and posterior distributions of the parameters converge satisfactorily.

2. The parameters estimated by WinBUGS software are given in the table below.  
(Parameter can be read from the “mean” column in the table)



**Table 1- showing calculation of parameters of the MNL model**

Node	Mean	Sd	MC error	2.5%	Median	97.5%	Start	Sample
b.brand2	-0.03686	10.01	0.1014	-19.64	9.46E-4	19.47	10001	10000
b.brand3	0.07863	10.03	0.09409	-19.53	0.06272	19.83	10001	10000
b.Price	-0.09079	9.946	0.1044	-19.4	-0.09129	19.3	10001	10000
b.Feat	-0.1108	9.924	0.1068	-19.75	-0.1622	19.11	10001	10000
b.Disp	-0.05104	10.07	0.09432	-19.7	-0.09232	19.76	10001	10000
Pred.price	0.4973	0.4842	0.004926	6.913E-22	0.47	1.0	10001	10000

**Table 1**

The parameters of the multinomial logistics regression model are

b.brand2 = -0.03683

b.brand3 = 0.07863

b.Price = -0.09079 (Price of the brand)

b.Feat = -0.1108 (advertisement of the brand)

b.Disp = -0.05104 (promotion of the brand)

pred.price = 0.4973 (predicted probability for a brand priced at Rs. 55)

The authors then estimated the probability of purchase for a detergent brand priced at Rs.55 and the probability was estimated to be 0.4973 (see table 1, pred.price). Thus for different price levels, we can predict the probability of purchase of the brands.

**Conclusion**

The parameters of the model were calculated using multinomial logistics regression, Bayesian Inference and MCMC. The concept is further utilized to forecast the probability of purchase at different price levels. The graphs (Trace Plots) generated by the software indicate convergence of data and leads to calculation of parameters as mentioned in (1) above. A marketer can forecast the probability of purchase of a particular brand at different price levels using multinomial logistics regression, Bayesian Inference and MCMC.

**Limitations**

The model is restricted in the sense that it allows only the main effects of price and other predictor terms. The article ignores the other aspects of buying behavior such as purchase incidence and quantity decisions.

**References**

- [1] Bettman James R., Mary Francis Luce, and John Payne (1998) *Journal of Consumer Research*, 187 – 217.
- [2] Bettman James R.(1998) *Journal of Marketing Research* Vol 4, pp 465 – 471.
- [3] Bettman James R. and Michel A. Zins (1977) *Journal of Consumer Research*, Vol. 4, pp. 75-85.
- [4] Geman S., and Geman D. (1984) *IEEE Transactions of Pattern Analysis and Machine Intelligence* 6:721–741.
- [5] Metropolis Nicholas, and Ulam S. (1949) *Journal of the American Statistical Association* 44:335–341.
- [6] Hastings W. K. (1979) *Biometrika* 57:97–109.
- [7] Chib Siddhartha, and Edward Greenberg (1995) *The American Statistician* 49:327–336.
- [8] Cowles M.K. and Carlin B.P. (1995) *Journal of the American Statistical Association* 91:883–904.
- [9] Brooks S. P. and Roberts G.O. (1998) *Statistics and Computing* 8:319–335.
- [10] Jackman Simon.( 2004) *Annual Reviews of Political Science* 7:483–505.
- [11] Payne John W. (1982) *Psychological Bulletin*, Vol. 92, pp. 382-402.
- [12] Brooks S. P., and Roberts G.O. (1998) *Statistics and Computing* 8:319–335.
- [13] Ionnis Ntzoufras (2008) *Bayesian Modeling Using WinBUGS*, John Wiley & Sons Publication U.S.A. pp 1 – 150.

ANNEXURE

**The WinBUGS software program along with data**

**Model**

```
{  
  
logit(p[i]) <- alpha + b.brand2*brand2[i] + b.brand3*brand3[i] + b.Price*Price[i]+ b.Feat*Feat[i]  
+b.Disp*Disp[i] #likelihood function for each data point  
  
#low[i] ~ dbern(p[i])  
  
alpha ~ dnorm(0.0,1.0E-2) #prior for intercept  
b.brand2 ~ dnorm(0.0,1.0E-2) #prior for slopes  
b.brand3 ~ dnorm(0.0,1.0E-2)  
b.Price ~ dnorm(0.0,1.0E-2)  
b.Feat ~ dnorm(0.0,1.0E-2)  
b.Disp ~ dnorm(0.0,1.0E-2)  
  
pred.price <- exp(alpha + b.brand2 + b.brand3 + b.Price + b.Feat + b.Disp) / (1 + exp(alpha +  
b.brand2 + b.brand3 + b.Price + b.Feat + b.Disp))  
}
```

**Data**

```

list(i = 975)
brand2[ ] brand3[ ] Price[ ] Feat[ ] Disp[ ]
0      1      48      1      0
0      1      48      1      0
0      0      150     1      0
0      0      150     1      0
0      0      150     0      0
0      1      48      1      1
0      1      48      0      1
0      1      48      1      1
1      0      148     0      1
0      1      48      0      0
0      1      48      1      0
0      0      150     1      1
1      0      148     1      1
0      1      48      1      0
0      0      150     1      0
1      0      148     0      0
1      0      145     0      0
1      0      145     1      0
0      1      48      0      0
1      0      145     0      0
0      0      150     1      0
1      0      148     0      0
0      1      48      0      0
1      0      148     1      0
0      1      48      0      1
0      0      150     0      0
1      0      148     1      0
0      1      48      1      0
0      1      48      0      0
1      0      148     1      0
1      0      140     1      1
0      0      150     0      1
0      0      140     1      1
0      1      48      1      0
0      0      140     0      1
0      1      48      1      1
0      1      48      1      0
1      0      148     1      1
0      1      48      0      0
0      1      48      1      0
0      0      150     1      0
1      0      148     0      1
0      0      150     1      1
0      1      48      0      0
0      1      48      1      0
0      1      48      0      0
1      0      148     1      1
0      1      48      0      0
0      1      48      1      0
0      0      150     1      0
1      0      148     0      1
0      1      48      0      1

```

A Brand Choice Model Using Multinomial Logistics Regression

0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	1	1
0	1	48	0	0
0	0	150	0	0
1	0	148	0	0
0	0	150	1	0
0	0	150	1	1
0	1	48	1	0
0	0	150	0	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1

0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	1	48	1	0
0	1	48	1	0
0	0	150	1	0
0	0	150	1	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	1
0	1	48	1	1
1	0	148	0	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0

A Brand Choice Model Using Multinomial Logistics Regression

0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0

0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0



A Brand Choice Model Using Multinomial Logistics Regression

1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	1	1
0	1	48	0	0
0	0	150	0	0
1	0	148	0	0
0	0	150	1	0
0	0	150	1	1
0	1	48	1	0
0	0	150	0	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1

0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	1	0
0	0	150	1	0
0	0	150	1	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	1
0	1	48	1	1
1	0	148	0	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0

A Brand Choice Model Using Multinomial Logistics Regression

0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0

0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	1	1
0	1	48	0	0
0	0	150	0	0
1	0	148	0	0
0	0	150	1	0
0	0	150	1	1
0	1	48	1	0
0	0	150	0	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1

A Brand Choice Model Using Multinomial Logistics Regression

0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	1	48	1	0
0	1	48	1	0
0	0	150	1	0
0	0	150	1	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	1
0	1	48	1	1
1	0	148	0	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0

1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	1	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1

A Brand Choice Model Using Multinomial Logistics Regression

0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0



1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	1	1
0	1	48	0	0
0	0	150	0	0
1	0	148	0	0
0	0	150	1	0
0	0	150	1	1
0	1	48	1	0
0	0	150	0	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0

A Brand Choice Model Using Multinomial Logistics Regression

0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0

0	1	48	1	0
0	1	48	1	0
0	0	150	1	0
0	0	150	1	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	1
0	1	48	1	1
1	0	148	0	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	0	150	1	1
0	1	48	0	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0

A Brand Choice Model Using Multinomial Logistics Regression

1	0	148	1	0
0	1	48	0	0
0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1
1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	1	0
0	0	150	1	0
0	0	150	1	0
0	0	150	0	0
0	1	48	1	1
0	1	48	0	1
0	1	48	1	1
1	0	148	0	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	1

1	0	148	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	0	0
1	0	145	0	0
1	0	145	1	0
0	1	48	0	0
1	0	145	0	0
0	0	150	1	0
1	0	148	0	0
0	1	48	0	0
1	0	148	1	0
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	1	48	1	0
0	1	48	0	0
1	0	148	1	0
1	0	140	1	1
0	0	150	0	1
0	0	140	1	1
0	1	48	1	0
0	0	140	0	1
0	1	48	1	1
0	1	48	1	0
1	0	148	1	1
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	0	1
0	1	48	0	1
0	0	150	0	0
1	0	148	1	0
0	0	150	1	0
0	0	150	0	1
0	1	48	0	0
0	0	150	1	0
0	1	48	1	1
1	0	140	1	1
0	0	140	1	1
1	0	148	0	1
0	1	45	1	1
1	0	135	1	1
0	1	48	1	0
0	0	150	1	0
1	0	148	1	0
0	1	48	0	0

A Brand Choice Model Using Multinomial Logistics Regression

0	1	48	0	0
1	0	148	0	0
0	0	150	1	0
0	0	145	1	1
0	0	145	1	1
0	1	48	0	0
0	0	150	1	0
0	1	48	0	0
0	1	48	1	0
1	0	148	1	1
0	1	48	0	1
1	0	148	0	0
0	1	48	1	0
0	1	48	0	0
0	0	150	1	0
0	0	150	0	0
0	0	150	0	1
0	1	48	1	1
0	1	48	0	0
0	1	48	1	0
1	0	148	1	0
0	1	48	0	0
0	1	48	1	0
0	0	150	1	0
1	0	148	1	1
0	1	48	0	0
0	0	150	0	0

**END**

Inits

list(alpha=0, b.brand2 = 0, b.brand3 = 0, b.Price = 0, b.Feat = 0, b.Disp = 0)