

 <p>ISSN NO. 2320-5407</p>	<p>Journal Homepage: -www.journalijar.com</p> <h2>INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)</h2> <p>Article DOI:10.21474/IJAR01/8598 DOI URL: http://dx.doi.org/10.21474/IJAR01/8598</p>	 <p>INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR) ISSN 2320-5407 Journal Homepage: http://www.journalijar.com Journal DOI:10.21474/IJAR01</p>
---	--	---

RESEARCH ARTICLE

A SIMPLE MODEL FOR ANALYZING THE CUSTOMER RETENTION COMPARING RURAL AND URBAN STORE.

Manash Pratim Kashyap.

Department of Business Administration, NERIM Group of Institutions.

Manuscript Info

Manuscript History

Received: 02 January 2018

Final Accepted: 04 February 2019

Published: March 2019

Key words:-

sBG Model, Beta distribution, Beta geometric, Beta binomial.

Abstract

Many businesses follow the repeat operation of a particular customer with a same store in different discrete time point. Customer retention comprises the probability of customer that alive in particular store in different purchase occasion. Different purchase occasion with a particular store occur at regular time intervals. The main goal of this paper is to predict future purchase patterns for rural and urban customer that can be described by the structural characteristics. The discrete time beta geometric model allows to heterogeneity of each of rural and urban customer behavioral process. The model is applied in 194 household from urban and 196 household from rural. Each household visited three times once in every three months. The study reveals that customer retention of rural customer is more than the urban customer.

Copy Right, IJAR, 2019,. All rights reserved.

Introduction:-

The increasing economic importance of staple goods in grocery store in rural and urban inspired many researchers. The term customer retention is the probability of a customer being “alive” or “repeat” buying from a firm (Gupta and Zeithams, 2006). Oliver (1997) define customer retention is deeply held commitment to re-buy or repatronize preferred product or service consistently in the future, despite situational influences and marketing efforts having the potential to cause switching behavior. Every business needs to survive, so it undertakes every possible action to get them. According to Peter Drucker (1973), the main purpose of a business is to create a customer. Marketing has traditionally focused on market share and the customer acquisition rather than on retaining existing customers and on building long lasting relationships with them (Kotler, 2003). Gupta et al. (2004) found that a 1 percent increase in customer retention had almost five times more impact on firm value than a 1 percent change in discount rate or cost of capital. Kotler and Keller (2006) give a reason for popularity of customer retention is the recognition that losing a customer means in fact more than a single sale i.e. losing the entire stream of purchases that this particular customer would make over a lifetime of patronage (Kotler and Keller, 2006).

The significance of customer retention was first introduced by Reichhel and Sasser (1990). They found that profits in service industries and credit card companies increased in direct proportion to the length of customer retention.

It is obvious that customer retention is more economical to keep existing customers than to acquire new one customer. Reichheld and Kenny (1990) observed that acquiring new customers is calculated as being five times more costly than the expenses of retaining existing customers. They observed that, long term customers buy more. The companies can clearly benefit from the life time spending of customers (Kotler, 2003 and Payne, 2006).

Corresponding Author:-Manash Pratim Kashyap.

Address:-Department of Business Administration, NERIM Group of Institutions.

In literature the concept of customer retention and customer loyalty is treated as same. But Cohen et al. (2006) argue that customer loyalty and customer retention are not same. According to Cohen et al. (2006) consumer inertia implies that some customers are only being retained, rather than expressing loyalty. According to them customer loyal are described as being less price sensitive and more inclined to increase the number of purchase.

Dwyer (1997) and Jain and Singh (2002) observed that customer retention model exist for two classes “lost for good” and “always a share”. The “lost for good” class define as customer defection as permanent while the “always a share” class considers customer switching to competitors as transient. Kalbfleisch and Prentice (1980) used hazard models using which fall into broad groups- accelerated failure time (AFT) or proportional hazard (PH) models. Schmittgen, Morrison and Colombo (1987) and Schmittlein and Peterson (1994) proposed a NBD/Pareto model for assessing the probability that a customer is still alive. This model is used by Reinartz and Kumar (2000). Fader, Hardie and Lee (2005) modified this model. The “always share” retention models view customer switching to competitors and typically uses migration or Markov models to estimate transition probabilities of customers being in a certain state. Bitran and Mondschein (1996), Gonul and Shi (1998) and Pfeifer and Carraway (2000) defined these states based on RFM measures while Rust, Lemon and Zeithamal (2004) used brands as states and estimated transition probabilities using a logit model. Simester, sun and Tsitsiklis (2005) used a binary tree approach to define the state space and estimated the transition probabilities using a non parametric approach. As mentioned earlier, retention are categorized into two groups- lost for good and always share. Some researchers have argued that customers should be treated as renewable resource (Dreze and Bonfrer, 2005). Thomas, Blattberg and Fox (2004) build a model for recapturing the lost customers.

Objective of the Study

The main objective of the study is to compare the customer retention of the grocery store for the two population i.e. rural and urban using stochastic model. The reason is that urban and rural customers are different in their culture, nature and purchasing capacity.

Hypotheses of the Study

Based on the objective our the major hypothesis is as follows

H₀: There is no significant difference between the ‘Customer retention’ of the grocery store of rural and urban. Probabilistic model is used to compare the customer retention of rural and urban grocery store.

Sampling and Data Consideration

The data is collected from Cachar district of south Assam, India. The district covers an area of 3786 square kilometres with 273694 households according to 2001 census. Out of which 21575 households are located in urban area and the rest are rurally located. Since the study involves both rural and urban customers, so a sample of households of size 194 from urban area and 196 from rural area are selected. The sample size is calculated with 95 percent confidence level and a confidence interval of 5 percent. However, a larger sample size 250 household is taken from each of the rural and urban population which is considered as final sample size. This was done to avoid the loss of contact with the respondents in subsequent stages of the study as it is spread across more than two years period.

The sampling is performed separately for rural and urban area. In the said district there are 1020 villages. From which 52 villages were selected i.e. one out of every twenty village. From the selected 52 villages, a total of 196 households are selected randomly, in such a way that the number of households selected from each village remains proportional to the total number of households in that village. There are only two towns in the district viz Silchar and Lakhipur which are divided into 28 and 10 municipal wards respectively. From each of the wards, households are selected randomly taking the total number of households selected from urban area to 194. In this case as well, the number of households selected from each of the municipal wards is proportional to the total number of households in that ward. The available information was loaded to Microsoft Excel and accordingly the data was analyzed.

Methodology:-

The methodology adopted in the study can be used for comparing the customer retention of the grocery store for rural and urban population. For the purpose of comparison the investigator has to contact the respondents for three purchase occasion after a gap of three months. This is done to observe whether the respondent choice the grocery store at subsequent purchase occasions. Similarly, the investigator observed the same for the next purchase occasion.

Now this makes sense to discuss the customer retention. The retention rate for the period t, denoted by r_t which is defined as follows

$r_t = \text{Customer active at the end of the period } t / \text{Customer active at the end of the period } t-1$
 i.e. proportion of customers active at the end of period t-1 who are still active at the end of the period t.

Marketers particularly give importance to the concept for characterizing the distribution of customer life time. The distribution of customer life time is studied by the survivor function which is denoted by $S(t)$. The survivor function is defined as the probability that a customer has survived to time “t” i.e. the customer is still active at the period “t”. The survival function $S(t)$ is defined as follows

$S(t) = \text{Probability that a customer survival up to time } t$

$$= r_1 \times r_2 \times \dots \times r_t$$

$$= \prod_{i=1}^t r_i$$

..... (1)

Again,

$$S(t-1) = \prod_{i=1}^{t-1} r_i$$

Thus,

$$r_t = \frac{S(t)}{S(t-1)} \quad \dots(2)$$

But practically it is observed that time horizon is limited and it is not possible to observe a cohort of customers infinitely. The investigator observed it is for some fixed time.

Introducing the Discrete Time Model

According to Fader and Hardie (2005), the duration of customer life times is based on the following two assumptions:

Every customer has a constant retention probability “ $1-\theta$ ”. The duration of the customer’s relationship with the grocery store denoted by the random variable T. So, T follows (shifted) geometric distribution with probability mass function

$$p(T = t/\theta) = \theta(1-\theta)^{t-1}; \quad t=1,2,3,\dots$$

..... (3)

And survivor function is given by

$$S(T = t/\theta) = (1-\theta)^{t-1} \quad t=1, 2, 3,\dots$$

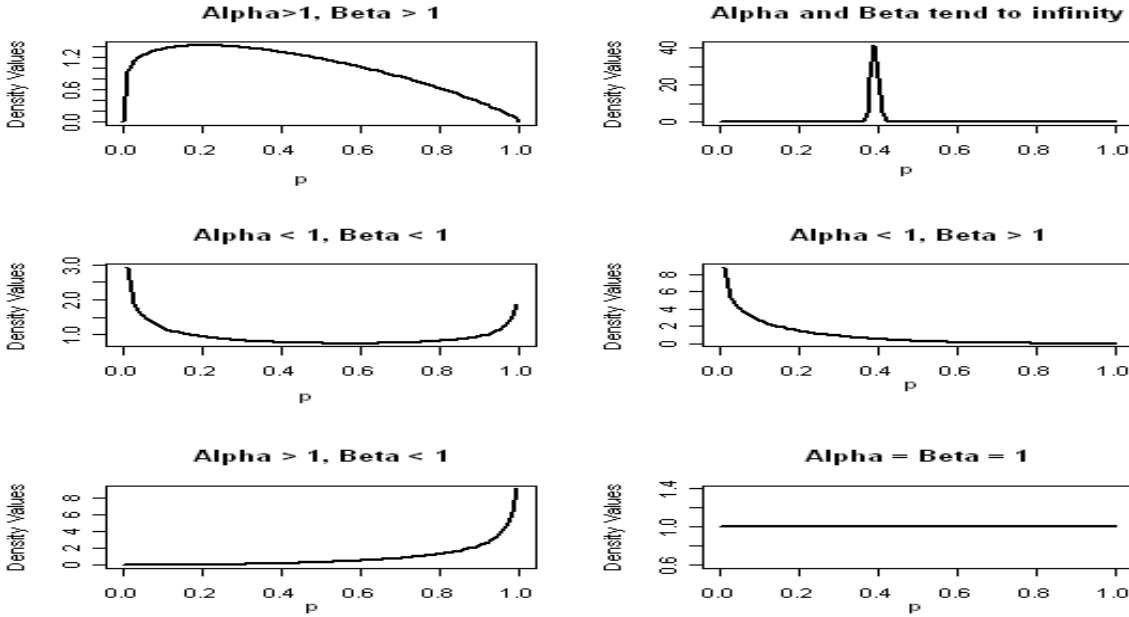
..... (4)

However θ cannot be considered as a constant but shall vary from customer to customer. Thus θ is assumed to probability density function

$$f(\theta/\alpha, \beta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha, \beta)}; \quad 0 < \theta < 1; \alpha > 0, \beta > 0$$

..... (5)

The reason to take heterogeneity coefficient θ to follow the beta distribution is for the flexibility of the distribution and also as it is bounded between zero and one. If both the parameters of the beta distribution (α and β) are small i.e. ($\alpha, \beta < 1$) the $f(\theta/\alpha, \beta)$ is “U- Shaped” or highly polarized across the customers. If both the parameters are relatively large ($\alpha, \beta > 1$); then $f(\theta/\alpha, \beta)$ are fairly homogeneous. The different shapes of the beta curve for different values of α and β can be seen in the figure 1



Since θ is unobserved, i.e. for a given individual the value of θ shall remain constant over the time period of the study,

$$\begin{aligned}
 P(T = t / \alpha, \beta) &= \int_0^1 P(T = t / \theta) f(\theta / \alpha, \beta) d\theta \\
 &= \int_0^1 \theta(1 - \theta)^{t-1} \frac{\theta^{\alpha-1} (1-\theta)^{\beta-1}}{B(\alpha, \beta)} d\theta \\
 &= \frac{1}{B(\alpha, \beta)} \int_0^1 \theta^\alpha (1 - \theta)^{\beta+t-2} d\theta \\
 &= \frac{B(\alpha + 1, \beta + t - 1)}{B(\alpha, \beta)} \quad \dots(6)
 \end{aligned}$$

This forward recursion formula used to compute the probabilities is derived by the following manner
 If we put $T = 1$ in (3) then

$$\begin{aligned}
 P(T = 1 / \alpha, \beta) &= \frac{B(\alpha + 1, \beta)}{B(\alpha, \beta)} \\
 &= \frac{\alpha}{\alpha + \beta}
 \end{aligned}$$

Now for $T=2, 3 \dots$

$$\begin{aligned}
 P(T = t) &= \frac{P(T = t)}{P(T = t - 1)} \times P(T = t - 1) \\
 \Rightarrow \frac{P(T = t)}{P(T = t - 1)} &= \frac{B(\alpha + 1, \beta + t - 1)}{B(\alpha, \beta)} \times \frac{B(\alpha, \beta)}{B(\alpha + 1, \beta + t - 2)} \\
 &= \frac{B(\alpha + 1, \beta + t - 1)}{B(\alpha + 1, \beta + t - 2)} = \frac{\beta + t - 2}{\alpha + \beta + t - 1}
 \end{aligned}$$

Finally, the customer retention rate is obtained using equation number (2). Substitute shifted Beta Geometric survivor function into (2) gives,

$$\begin{aligned}
 r_t &= \frac{S(t)}{S(t-1)} \\
 &= \frac{B(\alpha, \beta + t)}{B(\alpha, \beta + t - 1)} \\
 &= \frac{\beta + t - 1}{\alpha + \beta + t - 1}
 \end{aligned}$$

Estimation of the Parameters

The aim of this section is to compute the sBG model parameters using maximum likelihood estimate. Let, there are n_{ut} and n_{rt} customers who are still active at time t in grocery store for urban and rural population respectively. This means number of customer who buys products from the same store for the consecutive purchase occasion under study. Now, n_{u1} and n_{r1} are the urban and rural customer who drop out in the first period. Similarly, n_{u2} and n_{r2} are urban and rural customer who drop out in the second period. After the end of the second period, there are $n_u - (n_{u1} + n_{u2})$ and $n_r - (n_{r1} + n_{r2})$ customer still being active at the end of the second period for urban and rural population respectively. Assume that one customer is drop out is independent of the behavior of the other customer. It gives the specific values of the model parameter α and β . The joint probability of n_{u1} and n_{u2} urban customer and $n_u - (n_{u1} + n_{u2})$ urban customers still being active at the end of the second period is given by

$$P(Data_u / \alpha, \beta) = P(T = 1 / \alpha, \beta)^{n_{u1}} \times P(T = 2 / \alpha, \beta)^{n_{u2}} \times S(2 / \alpha, \beta)^{n_u - (n_{u1} + n_{u2})}$$

Similarly for rural,

$$P(Data_r / \alpha, \beta) = P(T = 1 / \alpha, \beta)^{n_{r1}} \times P(T = 2 / \alpha, \beta)^{n_{r2}} \times S(2 / \alpha, \beta)^{n_r - (n_{r1} + n_{r2})}$$

To estimate the model parameters, maximum likelihood estimation is used. So, the likelihood function is given by,

$$L(\alpha, \beta / Data_u) = P(T = 1 / \alpha, \beta)^{n_{u1}} \times P(T = 2 / \alpha, \beta)^{n_{u2}} \times S(2 / \alpha, \beta)^{n_u - (n_{u1} + n_{u2})}$$

Taking Log both side,

$$LL(\alpha, \beta / Data_u) = n_{u1} \ln[p(T = 1 / \alpha, \beta)] + n_{u2} \ln[p(T = 2 / \alpha, \beta)] + [n_u - (n_{u1} + n_{u2})] \ln[S(2 / \alpha, \beta)]$$

Similarly for the rural, it as follows,

$$LL(\alpha, \beta / Data_r) = n_{r1} \ln[p(T = 1 / \alpha, \beta)] + n_{r2} \ln[p(T = 2 / \alpha, \beta)] + [n_r - (n_{r1} + n_{r2})] \ln[S(2 / \alpha, \beta)]$$

After estimating the parameters the customer retention is calculated for the different discrete time period.

Findings and Analysis:-

Based on the methodology the different parameters are calculated using Microsoft Excel 2007. The survival data presented in table 1 are for segments of customer. One rural customer and another urban customer for grocery store type of business.

Table 1:-Observed Percentage of Rural and Urban Customer

Period	Rural Customer		Urban Customer	
	Alive (%)	Die (%)	Alive (%)	Die (%)
0 (First Purchase Occasion)	100	-	100.00	-
1 (Second Purchase Occasion)	75	25	55.37	44.63
2 (Third Purchase Occasion)	62.04	12.96	38.84	16.53

From Table 1 it is revealed that, in case of rural, 75 percent customers are alive in a grocery store in the second purchase occasion. It is assume that the number of customer in first purchase occasion is considered as cohort of customer in particular store. In third purchase occasion, 62.04 percent customers are alive. It is also observed from table 1, that 25 percent and 12.96 percent customers are die (the customer is switched to another shop) in the second purchase occasion and third purchase occasion respectively. In case of urban, 55.37 percent and 38.84 percent customer is alive in urban grocery store in second purchase occasion and third purchase occasion respectively. Similarly, 44.63 percent and 16.53 percent customer is dying in the second purchase occasion and third purchase occasion respectively.

Now our approach is to fit the sBG model for the three purchase occasion of rural and urban customer. The following table gives the different value of the model parameters which is calculated using Microsoft Excel 2007.

Table 2:-Parameter Estimation

Places	$\hat{\alpha}$	$\hat{\beta}$	θ
Rural	0.559988	1.679963	0.25
Urban	0.901502	1.118529	0.446281

Table 2 gives the value of the model parameters and the churn probabilities of rural and urban customer. The churn probability of urban consumer is more than the rural consumer. This indicates that more rural customer is active with the same store than the urban customer. The urban customer is more chance to switch the other grocery store than the rural customer. It is clearer from the figure 1.

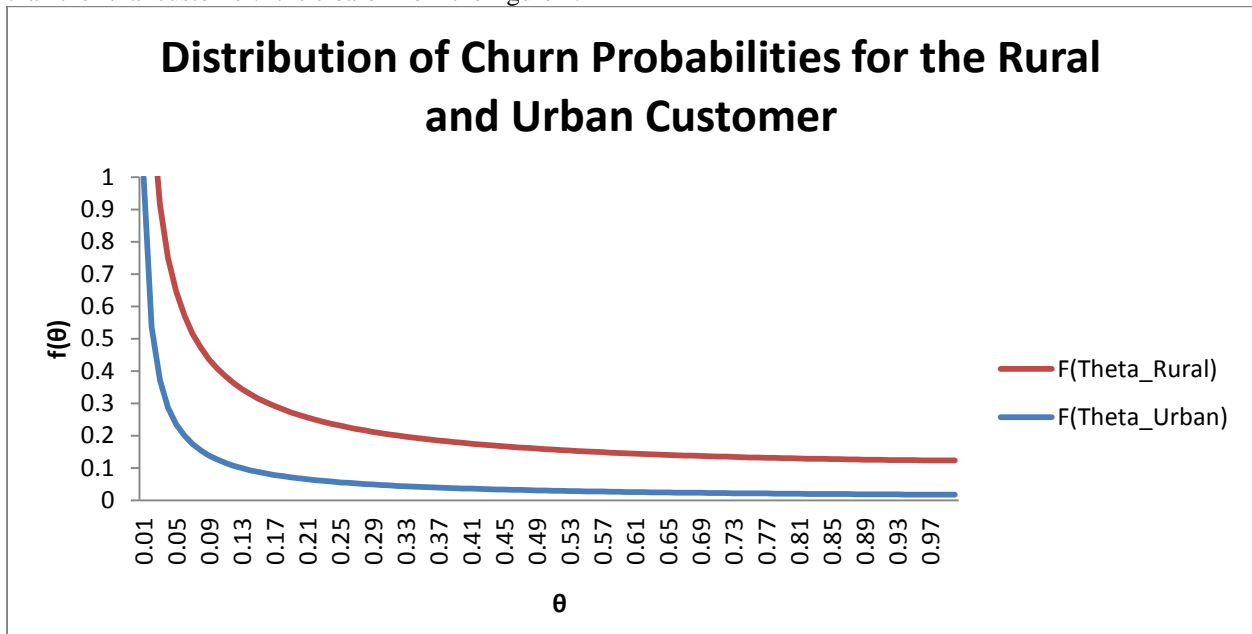


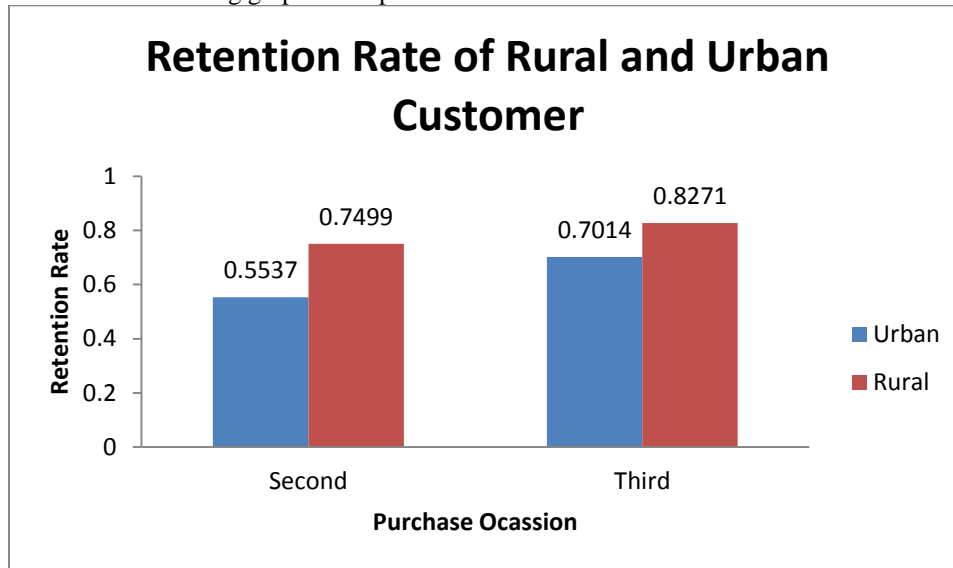
Figure 2:-Distribution of Churn Probabilities for the Rural and urban Customer

From the figure 2 it is observed that the distribution of churn rate rural and urban customers is “reverse J-Shape”. These gives that within rural and urban customer have fairly low churn probabilities and follow the beta distribution. Closer examination shows that urban consumer has more churn rate than the rural consumer. After estimating the sBG model parameter the retention rate of each of rural and urban consumer is computed the following table no 3 gives the customer retention rate of rural and urban for both the discrete time point second and third purchase occasion. Since, first purchase occasion is considered as cohort of customer for both rural and urban. The following table 3 gives the customer retention rate of rural and urban customer.

Table 3:-Customer retention of Rural and urban

Place	Purchase Occasion	
	Second	Third
Rural	0.75	0.827162
Urban	0.553719	0.701492

Table 3 shows that retention rate of rural customer is more than the urban customer in the second purchase occasion. Similarly, in the third purchase occasion rural customer has more retention rate than the urban customer. This indicates that rural customer has more chance to stay alive in the particular store than the urban customer. It is more comprehensible from the following graphical depiction

**Figure 3:-**Retention rate of Rural and Urban Customer**Conclusion:-**

The main purpose of the study is to compare the customer retention of rural and urban population. This study shows that customer retention of rural customers is greater than the urban customer. This means that rural customer is more attach with the same store than the urban customer at the different time horizon. This study also focuses the heterogeneity of customer at different time points. The urban consumer has more tendencies to leave the grocery store last purchase than the rural customer. The reason is that, (Vyas, 1997) it is observed that rural consumers vary not only in their behavior and practices but also in conviction and belief.

References:-

1. Bitran, Gabriel and Susana Mondschein (1996). Mailing Decisions in the Catalog Sales Industry. *Management Science*, 42(9).
2. Cohen, David; Gan, Christopher; Au Yong, Hua Hwa and Choong, Esther (2006). Customer Satisfaction: A Study of Bank Customer Retention in New Zealand. Commerce Division, Discussion Paper No. 109, (March).
3. Dréze, Xavier and André Bonfrer.(2005). Moving From Customer Lifetime Value to Customer Equity. Working Paper, University of Pennsylvania.
4. Drucker, F. Peter. (1973). *Management: Tasks Responsibilities Practices*. (Harper and Row).
5. Dwyer, F Robert (1997).Customer lifetime valuation to support marketing decision making. *Journal of Interactive Marketing*, 11(4), Autumn: 6-13.
6. Fader, Peter, Bruce Hardie, and Ka Lee (2005). RFM and CLV: Using Iso-value Curves for Customer Base Analysis. *Journal of Marketing Research*, 42(4): 415-430.
7. Gonul, Fusun, Byung-Do Kim and Mengze, Shi. (2000). Mailing Smarter to Catalog Customers. *Journal of Interactive Marketing*, 14 (2): 2-16.
8. Gupta, S., Lehmann, D. R., Stuart, J. A.(2004). "Valuing customer", *Journal of Marketing Research*, 41 (1): 7-18.

9. Gupta, Suni and Zeithaml, Valarie. (2006). Customer Metrics and Their Impact on Financial Performance. working paper, Columbia University, New York.
10. Jain, Dipak and Singh, Siddhartha. (2002). Customer Lifetime Value Research in Marketing: A Review and Future Directions. *Journal of Interactive Marketing*, 16(2).
11. Kalbfleisch, John and Prentice, Ross.(1980). *Statistical Analysis of Failure Time Data*. Wiley, New York.
12. Kotler, Philip and Kevin L. Keller (2006). *Marketing Management*. Upper Saddle River, NJ: Prentice Hall.
13. Kotler, P. (2003). *Marketing Management*. PHI, Tenth Edition, pp. 397.
14. Oliver, Richard. (1997). *Satisfaction: A Behavioral Perspective on the Consumer*. New York: McGraw-Hill.
15. Payne, Adrian (2006). *Handbook of CRM: Achieving Excellence in Customer Management*. Great Britain: Butterworth-Heinemann.
16. Pfeifer Phillip and Robert Carraway (2000). Modeling customer relationships as Markov chains. *Journal of Interactive Marketing*, 14(2): 43-55.
17. Reichheld, Frederick F. and W. Earl Sasser Jr. (1990). Zero Defections: Quality Comes to Services. *Harvard Business Review*, (September-October), 105-111.
18. Reinartz, Werner and V. Kumar (2000). On the Profitability of Long-life
19. Rust, Roland, Katherine Lemon and Zeithaml, Valarie (2004). Return on Marketing: Using Customer Equity to Focus Marketing Strategy. *Journal of Marketing*, 68(1): 109-126.
20. Schmittlein, D.C, Morrison, D.G and Colombo, R (1987). Counting Your Customers: Who are They and What will They do Next?. *Management Science*, 33 (1): 1-24.
21. Schmittlein, David and Peterson, Robert. (1994). Customer Base Analysis: An Industrial Purchase Process Application. *Marketing Science*, 13 (1): 41-68.
22. Simester, Duncan, Peng Sun and John Tsitsiklis (2005). *Dynamic Catalog Mailing Policies*. Working Paper, MIT.
23. Thomas, Jacquelyn, Blattberg, Robert and Fox, Edward. (2004), *Recapturing Lost Customers*. *Journal of Marketing Research*, 16 (February): 31-45.
24. Vyas, S. C. (1997). *Rural Marketing: Thrust and Challenges (First ed.)*. New Delhi: National Publishing House.