**Customer Lifetime Value: Evolution of Approaches to Predicting Customer Behavior**

**Introduction.** Predicting customers is incredibly hard. The state-of-art approaches of probabilistic modeling make CLV estimation easy. Buy-till-you-die models focus on forecasting customer’s attrition, transaction and spending processes. The study compares customer behavior predicting models in noncontractual business settings (Pareto/NBD, BG/NBD and MBG/NBD). The estimates and approximation quality are evaluated by using a transaction dataset gathered by a Russian restaurant chain. Static and dynamic control variables make up a generalized approach used in the study. This paper aims at filling the existing gap in Russian marketing theory, as it provides examples of Pareto/GGG parameter estimation and evaluation.

**Overview.** For quite a while, marketing theory’s been trying to answer whether to attract or retain customers. Historically, aggressive marketing strategies have prevailed, justifying the need for constant customer attraction. Conversely, supporters of relationship marketing preached the idea of customer retention. A target strategy depends on the cost-to-benefit ratio of both to reach a perfect balance. The estimated ratio varies across industries, companies and products. An extensive retention strategy is associated with high costs. A company should recognize the benefits of long-term cooperation with its customers, taking into account a different level of an individual contribution made to its steady income. A CLV is a discounted company’s future cashflow contributed to the whole future relationship with a customer. The strategy of long-term cooperation is successful when certain criteria are met: high market saturation; high initial costs of attracting customers; the benefits of maintaining relationships that allow to gain long-term competitive advantages. CLV models differ depending on business settings (contractual and noncontractual). Whilst in contractual business settings, it’s easy to identify customer’s churn risk by logistic regression or clustering analysis. Contractual settings introduce greater uncertainty, specific distribution assumptions need to be done.

**Methods.** The state-of-art CLV models applicable only in noncontractual business settings (Table 1) were addressed in the study. Buy-till-you-die models help companies tell the story of a customer from the very moment she was attracted until she’s gone for good.

**Data.** The dataset used in the study contains 326768 historic transaction records of 215518 Russian company’s customers who were active in a 6-month period (July through December 2022). The company owns a chain of Japanese restaurants located in 114 cities of Russia. The sample contains the buyer’s ID, transaction date, product set and monetary value.

Table 1. Model Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature | Pareto/NBD | BG/NBD | MBG/NBD | Pareto/GGG |
| Schmittlein et al. (1987) | Fader et al.(2005) | Batislam et al.(2007) | Reutterer et al. (2016) |
| Modeling transaction process | Negative binomial distribution | Negative binomial distribution | Negative binomial distribution | A mixture of 3 gamma distributions |
| Modeling individuals' attrition | Pareto distribution | Beta-geometric distribution | Modified beta-geometric distribution | Pareto distribution |
| Data requirements | Recency, frequency and age of the customer |
| Peculiarities | Dropout can occur at any point in time, independent of the occurrence of actual purchases | After any transaction, a customer becomesinactive | Considers an additional chance of dropout immediately after the first purchase | Considers observed intertransaction timing patterns |

*Source:* compiled by the author based on works of Schmittlein et al. (1987), Fader et al. (2005), Batislam et al. (2007), Reutterer et al. (2016).

**Results.** To achieve the parameter estimation of Pareto/NBD, BG/NBD, MBG/NBD and Pareto/GGG models, Python and R libraries have been used. The results are printed below (Table 2). Parameters of the gamma-gamma model were evaluated as well to forecast an average monetary value per transaction per customer.

Table 2. Parameter Estimates

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameters** | **Pareto/NBD** | **BG/NBD** | **MBG/NBD** | **Pareto/GGG** |
| r | 0.4408 | 0.2685 | 0.4456 | 0.6205 |
| α | 63.0917 | 39.6832 | 47.2182 | 8.1278 |
| s | 0.1826 |  |  | 0.5784 |
| β | 0.1826 |  |  | 4.3313 |
| a |  | 0.3686 | 0.4946 |  |
| b |  | 1.6272 | 1.2589 |  |
| t |  |  |  | 19.2993 |
| γ |  |  |  | 12.8286 |
| SMAPE, % | 22.68 | 22.64 | 22.48 | 10.18 |

Estimates are significant at the 1 percent level.

*Source:* compiled by the author.

The Pareto/GGG model showed the most excellent SMAPE and coefficient of determination (0.97). Figure 1 compares actual and predicted by Pareto/GGG frequency of repeat transactions.



Figure 1. Frequency of Repeat Transactions

The combination of Pareto/GGG and Gamma-Gamma models gives a powerful tool for predicting CLV throughout the entire customer dataset (Table 3).

Table 3. CLV Estimates

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameters** | **Mean** | **Min** | **Median** | **Max** |
| CLV, rubles | 32658.83 | 16.81 | 11789.34 | 6257790.28 |
| Frequency, days | 2.07 | 1.00 | 1.00 | 121.00 |
| Recency, days | 62.42 | 1.00 | 58.00 | 152.00 |
| Customer’s age, days | 110.65 | 1.00 | 121.00 | 153.00 |
| Monetary value, rubles | 3948.50 | 100.00 | 3052.00 | 69786.00 |
| Predicted number of purchases in 30 days | 0.36 | 0.00 | 0.25 | 20.27 |
| Probability of being alive in 30 days | 0.72 | 0.00 | 0.75 | 0.97 |

*Source:* compiled by the author.

**Conclusion and discussions.** The rapid development of the buy-till-you-die models is inconsistent with only a handful number of applications by researchers who have actually implemented them. A general criticism is that models are invariant to external factors (customer demographics or customer acquisition information). The study attempts to evaluate parameters by including covariates into probabilistic models. The parameters of the Pareto/GGG model were evaluated, several time-invariant and time-variant covariates were specified such as the dummy variable for ‘Gender’ (male = 0, female = 1) and for ‘Channel’ (online = 0, offline = 1). The results showed that women are more likely to purchase at a higher rate; customers acquired online make more purchases and have a lower churn rate. The assumption of independence between transaction and attrition might be addressed as well. Pareto/GGG was extended by adding the correlation parameter, after its estimation it turned out that it is insignificant.

The models may be extended by specifying additional factors of interest such as the customer’s age, address, the marketing mix and other variables considered significant under the research. The generalized models push the boundaries of applicability and usefulness of these probabilistic approaches.