Do Intensive Buyers Perform Differently? An Extension of BG/BB Model

Fan Liu, and Yelin Fu

Abstract—Researchers have put their effort in tracking customer future purchases based on their past transaction behavior, such as recency, the time that has passed since the customer last purchase, and frequency, the number of purchases in the observation period. We extend the beta-geometric/beta-Bernoulli (BG/BB) model by incorporating another important metric, i.e., intensity which measures the whether a customer concentrates his purchases in a short period or purchases with a steadily. Though complex, our model can still be implemented easily in Microsoft Excel. Via an empirical application, we show that our model outperforms the BG/BB model in predicting the number of donations made by a targeted cohort of donators.

Keywords—Customer-base, Data Mining, repeat buying, probability models.

I. INTRODUCTION

With the development of data mining technique, collecting and storing rich customer-level datasets become easier, stimulating customer managers to further investigate their customer base. The core task is to answering the following questions which is directly connected to the customer lifetime value:

1) Will the customer with a particular purchase history leave the company in the future?
2) How many times of purchases would he make is he is still active?

A number of probability models have been developed to deal with such task. One of the exploratory works is made by [1], where the best-known Pareto/NBD model is proposed. It makes the use of the recency, measuring how recent the customer made his last purchase, and frequency, measuring how often he purchases, and forecasts and tracks his future transactions. Such RF framework is widely applied in area of customer base analysis, for it is convenience and explicit to describe customer past purchase history with the framework.

One extension under RF framework is [2], where the assumption for the death time is changed. They assume that death only happens immediately after a purchase instead of occurring at any time assumed in [1], leading to a discrete beta-geometric death process. Their work is further extended in [3] who also let the transaction time be discrete and build a beta-binomial purchase distribution. [4] introduces another model under RF framework which assumes that death occurs periodically over the calendar time while keep the assumption of transaction time the same with [1]. They show that such periodical death opportunity (PDO) model can switch from NBD model to Pareto/NBD model, making it more general and flexible.

There are also other papers extending the basic RF framework, one of which is the famous RFM which additionally takes into account the money the customer spends on each transaction. [5] extends Pareto/NBD model to incorporate the dolor volume of past purchases. [6] develops a joint model of purchase time and quantity to assist customer selection using CLV, which is followed by [7] who applies a hierarchical Bayes approach to jointly model the purchase time, purchase amount and risk of defection. Other extensions include [8] who considers the correlation between purchase time and death time, [9] who assumes that customer will not permanently leave but return with a fixed rate, and [10] where online complaints are incorporated.

Though the RF framework has been highlighted by many researchers, it may not be sufficient to predict customer behavior today even in the discrete-basic situation which is discussed in [3]. For example, as depicted in Fig 1, there are two customers who both have 7 transaction opportunities and make total 4 purchases where both of the last purchases occur at the 7th transaction time. Thus, with the same recency and frequency data in the calibration period, their estimated purchases in the forecast period are also the same when using BG/BB model. However, from the intuition, two customers may have a diverse behavior. Customer 1 seems like to purchase more regularly, while customer 2 repurchase consecutively at first and then decreases the buying frequency, which may be a signal of leaving this market.

Fan Liu is with the School of Management, University of Science and Technology of China (corresponding author’s phone: (00186) 18823452037; e-mail: liuf1989@mail.ustc.edu.cn).

Yelin Fu is with the Department of Management Science, City University of Hong Kong.

http://dx.doi.org/10.15242/IIE.E0214044

Figure 1: Example
In this paper, we extend BG/BB model by taking into account the intensity of buying. We define intensity as the time that the consecutive transactions are made. We expect that customer behaves not only depend on the recency and frequency, but also on how the purchases distribute in the calibration period. For example, the probability that an active customer would repurchase may be related to whether he has made a purchase in the last buying opportunity.

Our model retains most of the settings in the BG/BB model but models the purchase time, the consecutive purchase time and the dropout date together. The new model is compared with BG/BB model via an application on a data set containing donations made by the supporters of a nonprofit organization. Results show that the RFI framework is preferred in identifying customer purchase patterns.

The rest of our paper is organized as follows. In the next section, we first formalize the assumptions and then specify our model incorporating the recency, frequency and intensity. Section 3 presents an empirical analysis, which is followed by the conclusions in Section 4. We also discuss the direction of future research at last.

II. MODEL DEVELOPMENT

A. Assumptions

Consider a customer who purchases at a fixed time interval, then his purchase history can be expressed by a binary string $(y_1, y_2, ..., y_t, ..., y_n)$ with $y_t = 1$ indicating that a transaction occurs at time $t$ and $n$ the total opportunities. Our model is based on the following four assumptions.

**Assumption 1.** While active, the probability that a customer buys at any transaction opportunity is conditioning on whether he has purchased at the last time point. We define $p_{01} = P(y_t = 1 | \text{alive at } t, y_{t-1} = 0)$ and $p_{11} = P(y_t = 1 | \text{alive at } t, y_{t-1} = 1)$

**Assumption 2.** After any transaction, the customer leaves the market with probability $\theta$. Then the state of whether the customer has left follows a geometric distribution.

**Assumption 3.** Heterogeneity in $p_{01}, p_{11}$ and $\theta$ all follow beta distributions. The density functions are:

\[
\begin{align*}
    f(p_{01} | \alpha, \beta) &= \frac{p_{01}^{\alpha-1}(1-p_{01})^{\beta-1}}{B(\alpha, \beta)}, 0 < p_{01} < 1, \alpha, \beta > 0 \\
    f(p_{11} | \gamma, \delta) &= p_{11}^{\gamma-1}(1-p_{11})^{\delta-1} \frac{1}{B(\gamma, \delta)}, 0 < p_{11} < 1, \gamma, \delta > 0 \\
    f(\theta | \theta, \mu) &= (1-\theta)^{\mu-1} \frac{1}{B(\theta, \mu)}, 0 < \theta < 1, \theta, \mu > 0
\end{align*}
\]

**Assumption 4.** The two transaction probabilities $p_{01}, p_{11}$ and the dropout probability $\theta$ vary independently across customers.

From the above assumption, we show that, compared with BG/BB, the probability that one customer makes purchases is split into two parts, leading to two binomial buying distribution. Therefore, we call our model as BG/BB model.

B. Likelihood Function

For a customer who has $x$ transactions in $n$ transaction opportunities, where the last transaction occurs at $t_x$. Further, we also observe that there are $x_c$ consecutive transactions, which means that there are totally $x_c$ purchases happening immediately after the last purchase. Thus, for a customer with $(x, x_c, t_x, n)$, the likelihood contributed is

\[
L(p_{01}, p_{11}, \theta | x, x_c, t_x, n) = \sum_{j=1}^{n} p_{01}^{x-x_c} (1-p_{01})^{j-x_c} | j > t_x \ast p_{11}^{x_c} \ast (1-p_{11})^{j} (1-\theta)^j
\]

Equation (1) is easy to be explained. Since there are $x_c$ consecutive transactions, there should be $x-x_c$ transactions followed by the opportunities where no purchase is made, leading to two components $p_{01}^{x-x_c}$ and $p_{11}^{x_c}$. And in the interval $(0, t_x)$, because $y_0 = 0$, then times that the customer transfer from $y = 1$ to $y = 0$ is the same with the times that he transfers from $y = 0$ to $y = 1$, leading to the component $(1-p_{11})^{j}$.

Taking expectation of (1) over the distribution of $p_{01}, p_{11}$ and $\theta$, we derive the likelihood function of a randomly chosen customers:

\[
\begin{align*}
    L(\alpha, \beta, \gamma, \delta, \theta, \mu | x, x_c, t_x, n) &= \sum_{j=1}^{n} \frac{B(\alpha + x - x_c, \beta + j - 2x + x_c - \delta_j > t_x)}{B(\alpha, \beta)} \ast \frac{B(y + x_c, \delta + x - x_c + \delta_j > t_x)}{B(y, \delta)} \ast \frac{B(\theta + \delta_j > t_x)}{B(\theta, \delta)} (2)
\end{align*}
\]

There are totally six parameters to be estimated. Suppose the observation is taken on $l$ customers, where customer $i$ is observed with $(x, x_c, t_x, n_i)$. Then the sample log likelihood is given by

\[
\begin{align*}
    LL(\alpha, \beta, \gamma, \delta, \theta, \mu) &= \sum_{i=1}^{l} Ln(L(\alpha, \beta, \gamma, \delta, \theta, \mu | x_i, x_{ci}, t_{xi}, n_i))
\end{align*}
\]

It can be maximized easily with Micro Excel, which is one of the desirable properties that we retained from [3].

C. Key Result

Here we present several quantities which are of interest to the scholars and managers who are mining the customer behavior data. In the following, we remove the subscript for convenience. To answer the first question in the Introduction, we have to compute the conditional probability that a customer being alive at $n$ with purchase history $(x, x_c, t_x, n)$.
\[
\begin{align*}
P(\text{alive at n}|p_{01}, p_{11}, \theta, x, x_c, t_x, n) &= P(x, x_c, t_x, n, \text{alive at n}|p_{01}, p_{11}, \theta) \\
&= \sum_{j=0}^{n} P(x, x_c, t_x, n, \text{alive at j}|p_{01}, p_{11}, \theta) \\
&= p_{01} x - x_c \ast \left(1 - p_{01}\right)^{n-x_c+\delta_{n>tx}} \ast p_{11} x_c \ast \left(1 - p_{11}\right)^{x - x_c - \delta_{n>tx}} \ast \left(1 - \theta\right)^n \\
&\ast L(p_{01}, p_{11}, \theta|x, x_c, t_x, n)^{-1}
\end{align*}
\]

(3)

Taking the expectation of (3) over the posterior distribution of \(p_{01}, p_{11}\), and \(\theta\), we have:

\[
P(\text{alive at n}|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n) = \frac{B(\alpha + x - x_c, \beta + n - 2x + x_c - \delta_{n>tx})}{B(\alpha, \beta)} \ast \frac{B(\gamma + x_c, \delta + x - x_c + \delta_{n>tx})}{B(\gamma, \delta)} \\
\ast \frac{B(\theta, \mu + n)}{B(\theta, \mu)} \\
\ast L(\alpha, \beta, \gamma, \delta, \theta, \mu|x, x_c, t_x, n)^{-1}
\]

(4)

To answer the second question, we first calculate the conditional probability one customer makes \(x^*\) purchases in the interval \((n, n + n^*)\) (including the transaction opportunity \(n + 1, n + 2, ..., n + n^*\)).

\[
P(X(n, n + n^*) = x^*|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n) = \delta_{x^*=0}P(\text{die before n}|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n) \\
+ P(\text{alive at n}|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n) \\
\ast P(X(n, n + n^*) = x^*|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n, \text{alive at n}) \\
= \delta_{x^*=0}A1 + \delta_{tx=n}A2 + \delta_{tx<n}A3
\]

(5)

where

\[
A1 = 1 - P(\text{alive at n}|\alpha, \beta, \gamma, \delta, \theta, \mu, x, x_c, t_x, n) \\
A2 = L(\alpha, \beta, \gamma, \delta, \theta, \mu|x, x_c, t_x, n)^{-1} \\
\ast \left[ \sum_{x_c = \max(\alpha, \theta)}^{\min(\alpha, \theta)} \sum_{x = \sigma}^{n^*} \sum_{j = 0}^{n^*} \frac{B(\alpha^*, \beta^*)}{B(\alpha, \beta)} \ast \frac{B(\gamma^*, \delta^*)}{B(\gamma, \delta)} \ast \frac{B(\theta^*, \mu^*)}{B(\theta, \mu)} \right] \\
A3 = L(\alpha, \beta, \gamma, \delta, \theta, \mu|x, x_c, t_x, n)^{-1} \\
\ast \left[ \sum_{x_c = \max(\alpha, \theta)}^{\min(\alpha, \theta)} \sum_{x = \sigma}^{n} \sum_{j = 0}^{n} \frac{B(\alpha^*, \beta^*)}{B(\alpha, \beta)} \ast \frac{B(\gamma^*, \delta^*)}{B(\gamma, \delta)} \ast \frac{B(\theta^*, \mu^*)}{B(\theta, \mu)} \right]
\]

where

\[
\sigma = 2x^* - n^*, \\
\alpha^* = \alpha + x - x_c + x^* - x_c, \\
\beta^* = \beta + n - 2x + x_c + j^* - 2x^* + x_c - \delta_{j>tx^*}, \\
\gamma^* = \gamma + x_c + x^*, \\
\delta^* = \delta + x - x_c + x^* - x_c + \delta_{j>tx^*},
\]

The empirical analysis shows that the model fit has been improved, though little, according to the AIC, which means that it is necessary to take into account not only the recency and frequency, but also whether and how the customer concentrate his purchases into a short interval.

B. Conditional Expectations

To show how the model forecast the supporters’ future donations, we compute the expected number of people who make \(t\) donations during year 2003-2006 with equation (5) by setting \(n^* = 4, x^* = t\). As is drawn in Fig 2, both BG/BB and BG/2BB model yield similar results, with BG/2BB performing a little worse. However, they provide a very good fit of the data.
Our model outperforms BG/BB model when we track the repeat donations over time. In particular, we compute the expected number of repeat donations for every single supporters with equation (6) by setting \( n^* = 1, 2, 3, 4 \), and then make a summation over the whole cohort. Results are plotted in Fig 3 with the actual cumulative numbers. It shows that the BG/BB model overestimate the purchase times. As illustrated in Fig 3 and Fig 4, both of the two models accurately track the future donations. Especially, with our model, the estimated total cumulative donation is only 1.81 percent lower. And the most deviation between the two curves in Fig 3 and Fig 4 is only 3.92% and 14.2%, respectively. In addition, our model follows the trend of the yearly donations better. While there is a sharply down in the curve of Actual in Fig 4, the curve of BG/BB seems to be too smooth. What’s more, BG/BB model performs better at first and yet turns to be the worse one at last, which means that our model may generate better prediction. It is further proved by the measure of mean absolute deviation (MAD) reported in Table II.

### Table II

<table>
<thead>
<tr>
<th>MODEL FIT</th>
<th>MAD</th>
<th>Year</th>
<th>Cumulative</th>
<th>Yearly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>2003</td>
<td>52.08</td>
<td>144.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2004</td>
<td>230.16</td>
<td>207.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2005</td>
<td>57.33</td>
<td>273.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2006</td>
<td>183.54</td>
<td>180.81</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>523.11</td>
<td>806.71</td>
</tr>
</tbody>
</table>

Obviously, the MAD for BG/2BB is less than BG/BB in the total number, illustrating that BG/2BB should be the better choice on this data.

### IV. CONCLUSION

Counting the number of active customers and the repeat transactions is important because it assists customer managers to decide which customers are more loyalty, which customer will make more transactions, and how many new customers should be attracted so as to achieve sales target. Today, with the advances of information technology, database marketing and the conceptual development in CRM, firms can accept more complicated statistical models to forecast consumer behavior.

In this paper, we extend the BG/BB model on the consideration that traditional RF-framework may be insufficient to identify customer purchase history. Like BG/BB model, we assume a geometric death process, while assuming two binomial purchase distributions conditional on the behavior in the past transaction opportunity. By doing so, we complement the RF framework with I, i.e., the intensity which measures whether and how customers concentrate their purchases within a short interval.

We show that our model retains the desirable properties of BG/BB model: (1) easiness of implementing the model with Microsoft Excel; (2) analytical expression of several statistics on existing customers, such as the probability to leave, the expected number of transactions and so on.

Our model is shown to perform well in the empirical analysis on a dataset which contains donations over year 1995 to 2006 by the supporters who made their first donations in 1995. With two more parameters, BG/2BB model outperforms BG/BB model in capturing the changing trends on purchase times.

Though RFI framework may be a better choice in discriminating customers by their transaction history, weakness
still exists. Intensity reflects whether purchases are often made intensively, but it fails to show when this intensive buying shows up. For example, in Fig 5, two customers with the same characteristics under RFI, but the former one is more likely to decrease his transaction times, while the other seems to become more and more preferable on this product. Future research would be on this topic to incorporate the tendency that a customer adjusts his purchase behavior.

Another direction is to apply the RFI framework in a continuous-time setting, that is the death time, or the transaction time, or both, could be continuous. Such research can be seen as the extensions of Pareto/NBD model or the BG/NBD model. One may also take into account of the metric of monetary of each transaction, considering whether the customer will spend more and less in the next purchase after a short time interval.

REFERENCES

http://dx.doi.org/10.1287/mnsc.33.1.1

http://dx.doi.org/10.1287/mksc.104.0098

http://dx.doi.org/10.1287/mksc.111.00654


http://dx.doi.org/10.1509/jmkr.44.4.579

http://dx.doi.org/10.1287/mksc.13.1.41

http://dx.doi.org/10.1287/mnsc.1070.0746

http://dx.doi.org/10.1287/mksc.109.0502

http://dx.doi.org/10.1007/s11002-010-9123-0