Using survival analytics to estimate lifetime value

Mike Grigsby
has worked in marketing analytics for nearly 30 years, working at Sprint, Dell, HP and the Gap. He is now an analytic consultant at Targetbase, focusing on the retail industry. Mike has been a frequent speaker at trade/academic conferences. He teaches marketing analytics at UTD’s graduate school. Mike has published in Marketing Insights, Canadian Journal of Marketing Research, Marketing Management and similar journals. His book, Marketing Analytics, was recently published by Kogan Page.

E-mail: m666grigsby@yahoo.com

Abstract Typically, lifetime value (LTV) is merely a calculation using descriptive/historical data. This calculation makes some rather heroic assumptions to project into the future but most importantly gives no insights into why a customer is, for example, lower valued, or how to make a customer higher valued. That is, descriptive techniques offer no insights into predicting, incentivizing or changing customer behavior. Using predictive techniques – in this case survival analysis – can give an indication into what causes purchases to happen. This means marketers get insights – levers – into how to increase LTV. This predictive modeling is strategically lucrative. This paper appeared in a different formant in Marketing Analytics, Kogan Page, June, 2015.

KEYWORDS: lifetime value, LTV, predicting next purchase, time until purchase, survival modeling, retail analytics, predictive modeling, targeting, consumer behavior, financial implications.

LIFETIME VALUE WITH DESCRIPTIVE ANALYSIS
Lifetime value (LTV) is typically done as just a calculation, using past (historical) data. That is, it is only descriptive.

While there are many versions of LTV (depending on data, industry, interest, etc.) the following is conceptually applied to all. LTV, via descriptive analysis:

1. Uses historical data to sum up each customer’s total revenue.

2. This sum then has subtracted from it some costs: typically cost to serve, cost to market, maybe cost of goods sold, etc.

3. This net revenue is then converted into an annual average amount and depicted as a cash flow.

4. These cash flows are assumed to continue into the future and diminish over time (depending on durability, sales cycle, etc.), often decreasing arbitralby by, say, 10 percent each year until they are effectively zero.

5. These (future, diminished) cash flows are then summed up in discounted (usually by weighted average cost of capital) to get their net present value (NPV).

6. This NPV is called LTV. This calculation is applied to each customer.
Thus, each customer has a value associated with it. The typical use is for marketers to find the “high-valued” customers (based on past purchases). These high-valued customers get most of the communications, promotions/discounts, marketing efforts, etc. Descriptive analysis is merely about targeting with already (on average) engaged, much like RFM.

This seems to be a good starting point but, as is usual with descriptive analysis, contributes nothing about why. Why is one customer more valuable? Will they continue to be? Is it possible to extract additional value, but at what cost? Is it possible to garner more revenue from a lower-valued customer because they are more loyal or cost less to serve? What part of the marketing mix is each customer most sensitive to? LTV (as described above) gives no implications for strategy. The only strategy is to offer and promote to the high-valued customers.

**SURVIVAL MODELING: AN INTRODUCTION**

Survival modeling came into bio-statistics in the early 1970s, where the subject studied was an event, death. Survival analysis is about modeling the time until an event. In bio-statistics the event is typically death, but in marketing the event can be response, purchase, churn, etc.

Due to the nature of survival analysis, there are a couple of characteristics that are endemic to this technique. The dependent variable is time-until-an-event, so time is built into the analysis. The second endemic thing to survival analysis is observations that are censored. A censored observation is either an observation that was lost to the study and there is no knowledge of having the event or not – but we do know at some point in time that the observation has not had the event.

In marketing it is common for the event to be a purchase. Imagine scoring a database of customers with time-until-purchase. That is far more actionable than, from logistic regression, probability of purchase.

What can be done about censored observations, that is, those that did not have the event? We could delete them. That would be simple, but depending on how many there are, it might discard a lot of data. Also, these might be the most interesting data of all, so deleting them is probably a bad idea. We could just give the maximum time until an event to all those that have not had the event. This would also be a bad idea, especially if a large portion of the sample is censored, as is often the case. (It can be shown that throwing away a lot of censored data will bias the results.) Thus, we need a technique that can deal with censored data. Survival modeling takes into account time until the event and includes information about those that have not had the event (yet).

**LIFETIME VALUE WITH PREDICTIVE ANALYSIS (USING SURVIVAL MODELING)**

How would LTV change using predictive analysis instead of descriptive analysis? First note that while LTV is a future-oriented metric, descriptive analysis uses historical (past) data and the entire metric is built on those, with assumptions about the future applied unilaterally to every customer.

Prediction will specifically thrust LTV into the future (where it belongs) by using independent variables to estimate the next time until purchase. Since the major customer behavior driving LTV is timing – and secondly, the amount and number of purchases – a statistical technique needs to be used that predicts time until an event. (Ordinary regression predicting the LTV amount ignores timing.)

Survival analysis is a technique designed specifically to study time-until-event problems. It has timing built into it, and thus a future view is already embedded in
the algorithm. This removes much of the arbitrariness of typical (descriptive) LTV calculations.

So, what about using survival analysis to see which independent variables, say, bring in a purchase? This decreasing time until purchase tends to increase LTV. While survival analysis can predict the next time until purchase, the strategic value of survival analysis is in using the independent variables to change the timing of purchases. That is, descriptive analysis shows what happened; predictive analysis gives a glimpse of what might change in the future.

Strategy using LTV dictates understanding the causes of customer value: why a customer purchases, what increases/decreases the time until purchase, probability of purchasing at future times, etc. Then, when these insights are learned, marketing levers (shown as independent variables) are exploited to extract additional value from each customer. This means knowing that the one customer is, say, sensitive to price and that offering a discount will tend to decrease their time until purchase. That is, they will purchase sooner (maybe purchase large amounts and maybe purchase more often) with a discount. Another customer prefers say, product X and product Y bundled together to increase the probability of purchase and this bundling decreases their time until purchase. This means that merely assuming the past behavior will continue into the future (as descriptive analysis does), with no idea why, is no longer necessary. It is possible for descriptive and predictive analysis to give contradictory answers, which is why “crawling” might be detrimental to “walking”.

If a firm can get a customer to purchase sooner, there is an increased change of adding purchases – depending on the product. But even if the number of purchases is not increased, the firm getting revenue sooner will add to their financial value (time is money).

Also, a business case can be created by showing the trade-off in giving up, say, margin but obtaining revenue faster. This means strategy can revolve around maximization of cost balanced against customer value.

The idea is to model next time until purchase – the baseline – and see how to improve that. How is this carried out? A behaviorally based method would be to segment the customers (based on behavior) and apply a survival mode to each segment and score each individual customer. By behavior is typically meant purchasing (amount, timing, share of products, etc.) metrics and marcom (open and click, direct mail coupons, etc.) responses.

**AN EXAMPLE**

Let us use an example. Table 1 shows two customers from two different behavioral segments. Customer XXX purchases every 88 days with an annual revenue of $43,958, costs of $7,296 for a net revenue of $36,662. Say the second year is exactly the same. So year 1 discounted at 9 percent is an NPV of $33,635 and year 2 discounted at 9 percent for two years is $30,857 for a total LTV of $64,492. Customer YYY has similar calculations for LTV of $87,898. The above (using descriptive analysis) would

<table>
<thead>
<tr>
<th>Customer</th>
<th>Days between purchases</th>
<th>Annual Purchases</th>
<th>Total Revenue</th>
<th>Total Costs</th>
<th>Net Revenue Yr 1</th>
<th>Net Revenue Yr 2</th>
<th>Yr 1 Discounted</th>
<th>Yr 2 Discounted</th>
<th>LTV at 9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXX</td>
<td>88</td>
<td>4.148</td>
<td>43,958</td>
<td>7,296</td>
<td>36,662</td>
<td>36,662</td>
<td>33,635</td>
<td>30,857</td>
<td>64,492</td>
</tr>
<tr>
<td>YYY</td>
<td>58</td>
<td>6.293</td>
<td>62,289</td>
<td>12,322</td>
<td>49,967</td>
<td>49,967</td>
<td>45,842</td>
<td>42,056</td>
<td>87,898</td>
</tr>
</tbody>
</table>
have the marketers targeting customer YYY with > $23,000 value over customer XXX. But do we know anything about why customer XXX is so much lower valued? Is there anything that can be done to make them higher valued?

Applying a survival model to each segment outputs independent variables and shows their effect on the dependent variable. In this case the dependent variable is (average) time until purchase. Say the independent variables (which defined the behavioral segments) are things such as price discounts, product bundling, seasonal messages, adding additional direct mail catalogues, offering online exclusives, etc. The segmentation should separate customers based on behavior and the survival models should show how different levels of independent variables drive different strategies. Table 2 shows the results of survival modeling on the two different customers that come from two different segments. The independent variables are price discounts of 10 percent, product bundling, etc. The TTE is time until event shows what happens to time until purchase based on changing one of the independent variable. For example, for customer XXX, giving a price discount of 10 percent on average decreases their time until they purchase by 14 days. Giving YYY a 10 percent discount decreases their time until purchase by only two days. This means XXX is far more sensitive to price then YYY—which would not be known by descriptive analysis alone. Likewise, giving XXX more direct mail catalogues pushes out their TTE but pulls in YYY by two days. Note also that very little of the marketing levers affect YYY very much. We are already getting nearly all from YYY that we can — no marketing effort does very much to impact the TTE. However, with XXX there are several things that can be done to bring in their purchases. Again, none of these would be known without survival modeling on each behavioral segment.

Table 2: Model Results

<table>
<thead>
<tr>
<th>Variables</th>
<th>XXX</th>
<th>YYY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price Discount 10%</td>
<td>-14</td>
<td>-2</td>
</tr>
<tr>
<td>Product bundling</td>
<td>-4</td>
<td>12</td>
</tr>
<tr>
<td>Seasonal message</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Five more catalogues</td>
<td>11</td>
<td>-2</td>
</tr>
<tr>
<td>Online exclusive</td>
<td>-11</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 3: After modeling implementation

<table>
<thead>
<tr>
<th>Customer</th>
<th>Days between purchases</th>
<th>Annual Purchases</th>
<th>Total Revenue</th>
<th>Total Costs</th>
<th>Net Revenue Yr 1</th>
<th>Net Revenue Yr 2</th>
<th>Yr 1 Discounted</th>
<th>Yr 2 Discounted</th>
<th>LTV at 9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>XXX</td>
<td>64</td>
<td>5.703</td>
<td>60,442</td>
<td>10,032</td>
<td>50,410</td>
<td>50,410</td>
<td>33,635</td>
<td>30,857</td>
<td>88,677</td>
</tr>
<tr>
<td>YYY</td>
<td>58</td>
<td>6.293</td>
<td>62,289</td>
<td>12,322</td>
<td>49,967</td>
<td>49,967</td>
<td>45,842</td>
<td>42,056</td>
<td>87,898</td>
</tr>
</tbody>
</table>
CONCLUSION
While it took some time for survival analysis to move from bio-statistics into marketing, it has now become an important tool in marketing analytics. Much of this is because when an event happens it is typically more important to marketers than the probability of an event happening. Marketing is about choice and that choice happens in time.

In terms of LTV, survival analysis has several advantages:

1. It is statistical, rather than merely mathematical. This means that it will not suffer as much from a recent change in a customer’s purchasing behavior. That is, if a customer, say, severely decreases their purchasing in recent time periods, predictive analytics will catch/adjust for this far faster/better than descriptive analysis (which tends to use past/historical data). The converse is also true.

2. As mentioned, it is predictive. That is, LTV can be estimated by several cause-and-effect (independent) variables.

3. Because it is predictive, these independent variables (especially marketing levers like price/discount offers, marcom vehicle/channel, product bundling, etc.) provide insight into how to change (increase) LTV.

4. Because of the above, a financial business case can be made. That is, if it, say, costs 5 percent more in discounts to increase a customer’s LTV, the model can show that the increase in discounts brings about only a 2 percent increase in LTV.

5. It allows the generation of a marketing strategy, rather than merely targeting high-valued customers.

Because of the above, survival analysis should be incorporated into all marketing analysts’ toolboxes. The effort is worth it.