sas

Segmentation and Lifetime Value Models

Using SAS®

Edward C. Malthouse

Contents

The Simple Retention Model

3

Contents

Your company acquires customers, provides them with a product or service, and makes a certain amount of profit each month until they terminate the relationship forever. How much profit do you expect to make from customers during their lifetimes? How would increasing retention rates affect future profit?

These are very common questions. Contractual service providers such as Internet service providers, health clubs, and media content providers (for example, Netflix, digital content subscribers, and so on), are in exactly this situation. For example, subscribers to Netflix pay a certain amount each month until they cancel. Companies that provide cellular phone service receive monthly payments from their customers until they cancel.

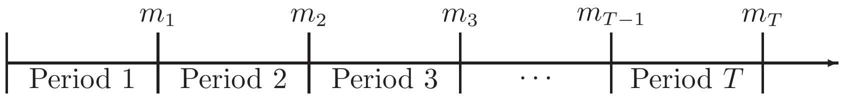
One application is determining how much can be spent to acquire a customer. For example, it may cost a cellular phone company $400 to acquire a new customer and provide a handset; even if he only generates $50 in profit each month this would be a good investment as long as the cellular phone company can retain him sufficiently long to recoup the acquisition cost. Likewise, a company that is considering whether to invest marketing resources in retaining customers longer will need to know CLV.

This chapter and the next show how to estimate the value of such customers in contractual situations. We begin with the case in which customers sign a contract for a certain number of periods and are not allowed to cancel. Next, we present the *simple retention model*, which allows customers to cancel, but assumes that the retention rate is constant over time and across customers, and that cash flows are independent of the cancelation time. The next chapter discusses when retention rates change over time, and when payment amounts depend on the time of cancelation.

3.1 The customer annuity model

Suppose for now that new customers sign a contract to make T payments of amount m in the future and are not allowed to cancel the contract. A book-of-the-month club is an example, where customers receive a series of 12 books on different topics, one each month. Customers allow the publisher to debit a credit card for \$20 after receiving each book.

Customers in this business situation are annuities. The publisher invests a certain amount of money to acquire a customer and, thus, the promise of T future payments of amount m. It is helpful to illustrate this with a diagram, where each m represents a payment and the subscript the payment number:



Notice that the payments come at the end of every period. Suppose further that the discount rate is d. The CLV of a customer is given by the formula for the present value of an *ordinary annuity*:

$$\text{PV}_T = \sum_{t=1}^{T} \frac{m}{(1+d)^t} = m\frac{1-(1+d)^{-T}}{d}. \tag{3.1}$$

More generally, the payment m could be a sum of cash inflows and outflows. An example of a negative cash flow is the cost of marketing to a customer each period. We discuss what makes up m in section 3.4. For example, it is often the gross contribution margin per customer—revenues minus the cost of sales.

EXAMPLE 3.1: Book-of-the-month club

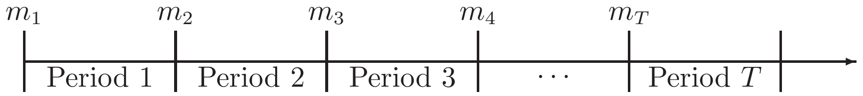
Those who join a monthly book club agree to buy $T = 12$ books, one at the end of each month, each generating a gross margin of $m = \$10$. Find the CLV of a customer, which is the present value of the 12 payments using a monthly discount rate of $d = 1\%$.

Solution

$$\text{CLV} = \text{PV}_{12} = 10\frac{1-(1+.01)^{-12}}{.01} = \$112.55.$$

The Excel function `PV` also gives the answer: −PV(.01, 12, 10).

Companies usually require prepayment for the product or service so that the first payment is received at time 0 rather than at the end of the first period. The T payments are illustrated as follows:



In these situations the formula for an *immediate annuity* is used:

$$\text{PV}_T = \sum_{t=0}^{T-1} \frac{m}{(1+d)^t} = m\frac{(1+d)[1-(1+d)^{-T}]}{d}. \tag{3.2}$$

EXAMPLE 3.2: Book-of-the-month club (continued)

Suppose the book-of-the-month-club requires prepayment each month for the $T = 12$ books and that each one generates a gross margin of $m = \$10$. Find the present value of the 12 payments, assuming a monthly discount rate of $d = 1\%$.

Solution

$$\mathrm{PV}_{12} = 10\frac{(1+.01)[1-(1+.01)^{-12}]}{.01} = \$113.68$$

or $-$PV(.01, 12, 10,,1) in Excel.

To understand where the annuity formulas come from (and other formulas in this chapter), we must remember how to find the sum of a geometric series. A *geometric series* is given by

$$S_T = a + ar + ar^2 + \cdots + ar^{T-1},$$

where $a \neq 0$. This sum can be evaluated by multiplying both sides of the equation by r,

$$S_T r = ar + ar^2 + ar^3 + \cdots + ar^T,$$

subtracting the second expression from the first,

$$S_T - S_T r = S_T(1-r) = a - ar^T,$$

and solving for S_T

$$S_T = \frac{a(1-r^T)}{1-r}. \tag{3.3}$$

The annuity formulas are a special case of this expression, where $a = m$ and $r = 1/(1+d)$ (see Exercise 3.9). When $|r| < 1$,

$$\lim_{T\to\infty} S_T = \frac{a}{1-r}. \tag{3.4}$$

3.2 The simple retention model

The customer annuity model assumes that customers make a pre-determined number of payments and that they never stop making payments before the end of the contract, but these assumptions are usually not realistic. Instead, some customers discontinue the relationship early and others continue using the product or service long after the initial contract has expired. The *simple retention model* (SRM) estimates CLV assuming the following

- The percentage of customers retained each month (the *retention rate*) r is constant over *time* and across *customers*.
- The period cash flow m is unaffected by the cancelation time.
- The event that a customer cancels in time period t is independent of the event that the customer cancels in any other time period.

These assumptions are relaxed in the next chapter. First, we will show how to estimate CLV using a spreadsheet, and then we will derive formulas.

EXAMPLE 3.3: Internet service provider (ISP)

An Internet service provider (ISP) acquires 1,000 customers who will pay \$50 at the end of each month for the service, with a gross margin of \$25. The ISP retains 80% of its customers each month and discontinues service immediately to anyone who fails to make a payment, without attempting to reactivate the customer. Find the CLV of this cohort, assuming a monthly discount rate of 1%.

Solution The answer can be found using the spreadsheet shown below. Immediately after acquisition (time 0) the ISP has 1,000 customers with a cash flow of \$25,000. At the end of the first month (time = 1), 800 (80% of the 1,000) are expected to make payments, generating a total cash flow of \$20,000 (\$800 × \$25) with a present value of \$19,802(\$20,000/1.01). At the end of the second month (time = 2), 640 (80% of the 800) customers are expected to make payments generating a cash flow of \$16,000 with a present value of \$15,685 ($\$16,000/1.01^2$). The answers for periods 0–5 are shown below. The formulas can be easily programmed in a spreadsheet and then copied. After 100 if you sum the discounted payments the CLV = $\$120,238$. (Readers should reproduce this spreadsheet at this time in software of their choice.)

Time	Expected Number of Customers	Raw Cash Flow	Discounted Cash Flow
0	1000.0	25,000	25,000
1	800.0	20,000	19,802
2	640.0	16,000	15,685
3	512.0	12,800	12,424
4	409.6	10,240	9,840
5	327.7	8,192	7,794
⋮	⋮	⋮	⋮

A virtue of this spreadsheet solution is that it is concrete and intuitive, making the problem easy to understand, but it is somewhat cumbersome, and we have not provided any formal rationale for it. We will develop a convenient, closed-form expression for CLV that enables us to avoid evaluating the sums in a spreadsheet and is based on a probabilistic model. Before doing so, we will briefly review and illustrate the necessary concepts from elementary probability theory.

A *random variable* X assigns a number to the outcome of a random experiment. The *probability mass function* (PMF) of a random variable that takes discrete values gives the probability $f(x) = P(X = x)$ that random variable X will take the value x. Upper-case letters will be used to represent random variables and corresponding lower-case letters indicate a realization of the random variable. We are often interested in knowing the *mean* or *expected value* of X, which describes the location of the middle of the distribution of X:

$$E(X) = \sum_x xP(X = x). \tag{3.5}$$

The mean can be interpreted as a weighted average of the x values, with the weights determined by the probabilities. Sometimes we are more interested in some transformation $g(X)$. For example, X could be the time until cancelation, and $g(X)$ could be the CLV of a customer who cancels at time X. The mean of $g(X)$ is

$$E[g(X)] = \sum_x g(x)P(X = x). \tag{3.6}$$

EXAMPLE 3.4: Number of orders

Suppose that on any given day a company receives zero orders with probability .2, one order with probability .4, two orders with probability .3, and three orders with probability .1. The company never receives more than three orders in a day. The fixed costs, which are incurred regardless of the number of orders, are five, the marginal cost per order is two and the marginal revenue per order is eight. Find the expected number of orders and the expected profit on a given day. Graph the PMF.

Solution Let X be a random variable indicating the number of orders on a day and let $g(X) = (8-2)X - 5 = 6X - 5$ be the profit on a day when there are X orders. The means are computed by using equations 3.5 and 3.6 above (Malthouse and Mulhern, 2007). It is convenient to evaluate these formulas using a spreadsheet as shown below. The first column enumerates all possible values of X (sample space) and the second records the PMF. The sum of all the probabilities must equal 1, as shown in the last row. The third column computes the product of the first two columns, and its sum is the mean number of orders in a day $E(X) = 1.3$.

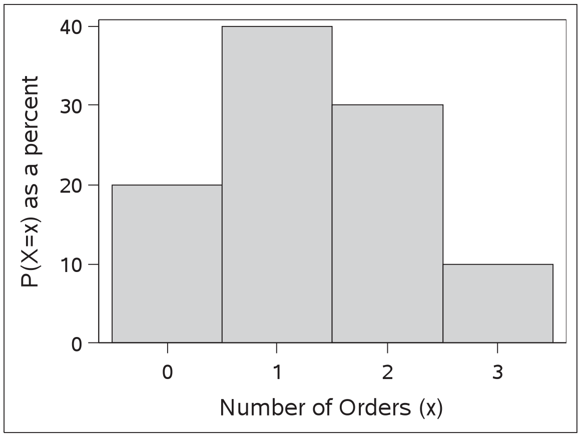
x	$P(X=x)$	$xf(x)$	$g(x)$	$g(x)f(x)$
0	.2	0	-5	-1.0
1	.4	0.4	1	0.4
2	.3	0.6	7	2.1
3	.1	0.3	13	1.3
Sum	1	1.3		2.8

The fourth column transforms the number of orders into profit. For example, the probability that profit will equal 13 is the same as the probability of receiving three orders, $P[g(X) = 13] = P(X = 3) = .1$. The last column evaluates equation 3.6 by computing the product of the second and fourth columns; its sum gives the expected daily profit[1] $E[g(X)] = 2.8$.

The graph of the PMF is shown below. The mean, 1.3, is an indication of the center of this distribution.

```
DATA pdf;
  INPUT x p @@;
DATALINES;
0 2     1 4     2 3     3 1
RUN;

PROC SGPLOT DATA=pdf;
  HISTOGRAM X / FREQ=p NBINS=4;
  XAXIS LABEL="Number of Orders (x)";
  YAXIS LABEL="P(X=x) as a percent";
RUN;
```



We can now derive a probabilistic model for CLV. Assume that all customers in some group are retained each period with probability r (the retention rate) for all periods and that the event a customer cancels during some period is independent of the event of cancelation during any other period. Let T be a random variable indicating the time of cancelation. Under these assumptions, T has a *geometric distribution*. Probabilities of a geometric distribution are given by PMF

$$f(t) = P(T = t) = r^{t-1}(1 - r). \tag{3.7}$$

The rationale for this formula is that a customer must be retained for $t - 1$ periods and then defect. Because defaulting is assumed to be independent across time periods, the retention probabilities can be multiplied, so that r^{t-1} is the probability of retaining a customer for $t - 1$ periods and $(1 - r)$ is the probability of defaulting (in the last period).

In addition to knowing the probabilities of certain cancelation times in equation 3.7, we sometimes want to know the probability that a customer has survived until the beginning of period t, which is called the *survival function*,

$$S(t) = P(T \geq t) = r^{t-1}. \tag{3.8}$$

We can equivalently think of the survival function as giving the probability that the customer cancels at time t or later, or that the customer survives the first $t - 1$ periods.

The survival function can be used to find quantiles of T. The α quantile[2] of random variable T, call it P_α, divides a distribution so that α percent of the distribution has $T \leq P_\alpha$ and $1 - \alpha$ percent of the distribution has $T \geq P_\alpha$, that is, $P(T \leq P_\alpha) = \alpha$ and $P(T \geq P_\alpha) = 1 - \alpha$. We can find the α quantile of the cancelation time under the SRM by solving

$$S(t) = P(T \geq P_\alpha) = r^{P_\alpha - 1} = 1 - \alpha.$$

Taking logs of both sides and solving for P_α we find that

$$P_\alpha = 1 + \frac{\log(1 - \alpha)}{\log r}. \tag{3.9}$$

For example, the *median time until cancelation* is found by substituting $\alpha = .5$ into the equation. Because events occur only at discrete time, the quantile is the greatest integer of equation 3.9.

The mean of a geometric distribution is given by (see Example 4.2 and Exercise 3.10):

$$E(T) = 1/(1 - r). \tag{3.10}$$

EXAMPLE 3.5: ISP: distribution of cancelation time

Graph the PMF and survival function of cancelation time for the ISP having a retention rate of 80%. Find and interpret the mean and median time of cancelation. Suppose the ISP implements a new retention campaign and is able to increase the retention rate to 90%. Find the mean and median under this new retention rate.

Solution The probabilities are computed using $P(t) = .8^{t-1}(1 - .8)$ and plotted in Figure 3.1 (the reader should replicate the probabilities and graph in a spreadsheet program). The distribution is skewed to the right. As we saw in Example 3.3, 20% of the customers are expected to cancel before the payment at time 1 (that is, the customer cancels at $T = 1$) while others will remain much longer. The expected time until attrition is $E(T) = 1/(1 - .8) = 5$ months. This ISP therefore expects 4 payments from each customer if the payments come at the end of a period. The median cancelation time is $P_{.5} = 1 + \log(1 - .5)/\log(.8) = 4.106 \approx 4$. At least half of the customers will survive until period 4.

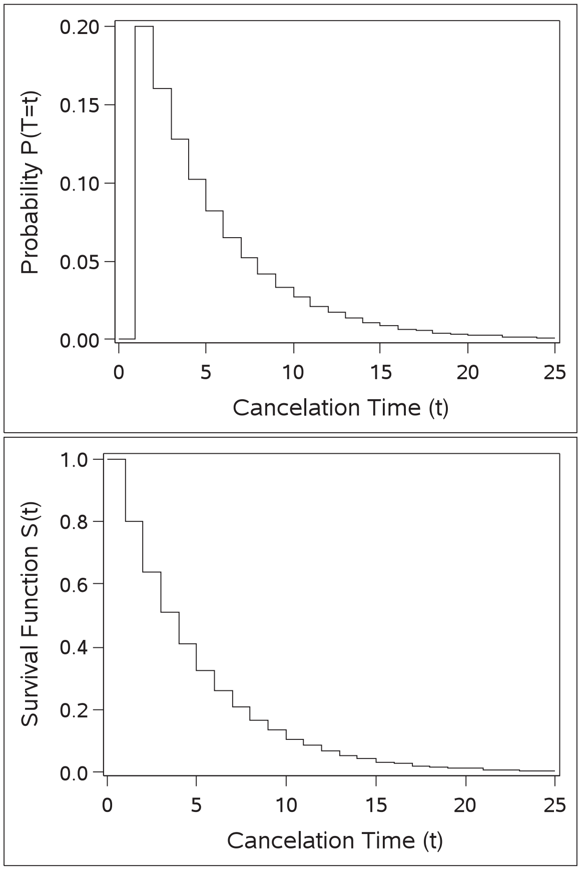
The survival function gives the probability that a customer has survived until the beginning of time t. All of the customers (100%) have survived at time 0 and 80% have

Figure 3.1 The geometric distribution ($r = .8$)

```
DATA geo;
  t=0; p=0;  St=1;   OUTPUT;
  t=1; p=.2; St=1-p; OUTPUT;
  DO t = 2 TO 25;
    p=p*.8;
    St=St-p;
    OUTPUT;
  END;
RUN;

PROC SGPLOT DATA=geo;
  STEP X=t Y=p;
  XAXIS LABEL = "Cancelation Time (t)";
  YAXIS LABEL = "Probability P(T=t)";
RUN;

PROC SGPLOT DATA=geo;
  STEP X=t Y=St;
  XAXIS LABEL = "Cancelation Time (t)";
  YAXIS LABEL = "Survival Function S(t)";
RUN;
```



survived at time 1. Only $.8^5 \approx 33\%$ have survived until period 5 and $.8^{10} = 10\%$ have survived until period 10.

After implementing the new retention campaign, the expected time until attrition increases to $E(T) = 1/(1 - .9) = 10$ months. The ISP now expects 9 payments from each customer.

CLV is the sum of the present values of future cash flows. When a customer cancels during period t, there will be t cash flows if they occur at the beginning of a period and $t - 1$ cash flows if they come at the end. For a specific cancelation time, the present value of these cash flows can be found with the annuity formulas (Equation 3.1 or 3.2). But the cancelation time T is random, and CLV will thus have a distribution. Those customers who have larger T have a larger CLV. We can summarize the distribution of CLV with its mean. When cash flows occur at the beginning of the period (immediate annuity) the expected value is

$$E(\text{CLV}) = \frac{m(1+d)}{1+d-r}. \tag{3.11}$$

When cash flows come at the end (ordinary),

$$E(\text{CLV}) = \frac{mr}{1+d-r}. \tag{3.12}$$

Proof We show the result for ordinary annuities using equation 3.6 with $g(t)$ given by equation 3.1:

$$\begin{aligned}
E(\text{CLV}) &= \sum_{t=1}^{\infty} r^{t-1}(1-r)PV_{t-1} = \sum_{t=0}^{\infty} r^t(1-r)PV_t \\
&= \sum_{t=0}^{\infty} r^t(1-r)m\frac{1-(1+d)^{-t}}{d} \\
&= \frac{m(1-r)}{d}\left[\sum_{t=0}^{\infty} r^t - \sum_{t=0}^{\infty}\left(\frac{r}{1+d}\right)^t\right] \\
&= \frac{m}{d} - \frac{m(1-r)}{d}\cdot\frac{1}{1-r/(1+d)} \\
&= \frac{mr}{1+d-r}.
\end{aligned}$$

When cash flows come at the beginning of the period the company receives an additional (non-random) payment of m at time 0:

$$E(\text{CLV}) = m + \frac{mr}{1+d-r} = \frac{m+md-mr+mr}{1+d-r} = \frac{m(1+d)}{1+d-r}.$$

EXAMPLE 3.6: ISP: lifetime value

Find the expected time until attrition and expected CLV of a single ISP customer assuming a monthly discount rate of 1%, monthly cash flows of \$25 at the beginning of each month, and retention rates of $70\%, 75\%, \ldots 95\%$, and 98%. Plot CLV against the retention rate.

Solution We first illustrate using equation 3.11 to compute CLV when $r = .8$, and confirm the answer found with the spreadsheet method in used Example 3.3.

$$E(\text{CLV}) = \frac{\$25(1+.01)}{1+.01-.8} = \$120.24.$$

CLV for the other attrition rates are computed with the same formula and shown in the table below.

The plot illustrates the importance of nurturing long-term relationships with customers. CLV changes at an increasing rate as the retention rate increases. For example, if a company could improve the retention rate from 70% to 75%, CLV would increase from \$81 to \$97. The same 5% improvement in retention rate from 90% to 95%, however, nearly doubles CLV from \$230 to 421\$. CLV is doubled again by increasing the retention rate from 95% to 98%. High customer loyalty pays good dividends.

Gupta and Lehmann (2003) calls $r/(1+d-r)$ the *margin multiple*. This multiple tells how the period gross margin for a customer (m) is related to CLV. The margin multiples for different retention rates are computed in Table 3.1. Assuming a 15% discount rate, the period gross margin is multiplied by 3.6 with a 90% retention rate, 4.75 for a 95% retention rate, and so on.

EXAMPLE 3.7: Insurance company

(Based on by Berger and Nasr, 1998, Case 1) An insurance company wants to estimate CLV. The company pays, on average, \$50 per customer yearly on promotional expenses. The yearly retention rate is 75%. The yearly gross contribution per customer is expected to amount to \$260. An appropriate annual discount rate is 20%. Find CLV, assuming that

```
DATA clv;
  DO r = .7 to .98 by .01;
    Et = 1/(1-r);
    Eclv = 25*1.01/(1.01-r);
    OUTPUT;
  END;
  FORMAT clv DOLLAR8.2
    r PERCENT5.0 Et 5.2;
  LABEL
    Et = "E(T)"
    Eclv = "E(CLV)";
RUN;

PROC PRINT DATA=clv NOOBS LABEL;
  VAR r Et Eclv;
  WHERE MOD(100*r,5)=0 OR r=.98;
RUN;

PROC SGPLOT DATA=clv;
  SERIES X=r Y=Eclv;
  XAXIS LABEL = "Retention Rate (r)";
RUN;
```

r	*E(T)*	*E(CLV)*
70%	3.33	$81.45
75%	4	$97.12
80%	5	$120.24
85%	6.67	$157.81
90%	10	$229.55
95%	20	$420.83
98%	50	$841.67

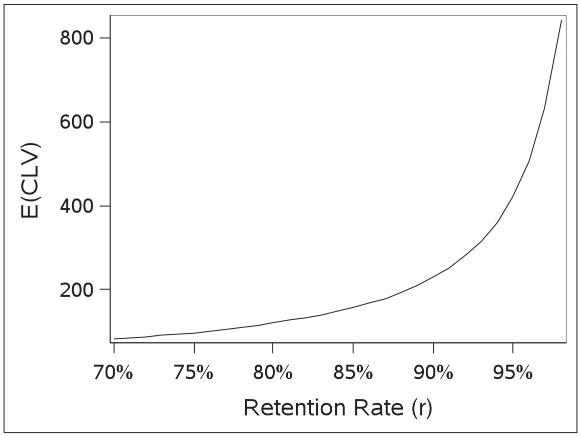


Table 3.1 Margin multiples for different retention and discount rates

Retention	Discount Rate (d)		
Rate (r)	.1	.15	.2
.6	1.20	1.09	1.00
.7	1.75	1.56	1.40
.8	2.67	2.29	2.00
.9	4.50	3.60	3.00
.95	6.33	4.75	3.80
.97	7.46	5.39	4.22
.99	9.00	6.19	4.71

you have just acquired a customer, receive the first payment at time 0 (do not worry about acquisition costs, because you would want to compare your answer with the acquisition costs to evaluate customer profitability), and will incur marketing costs at the end of each year, so that revenues are immediate while costs are ordinary.

Solution We use equation 3.11 to compute lifetime revenue because the first payment comes at the beginning of the relationship. For costs, equation 3.12 is used because the first costs will not be incurred until the end of the first year.

$$\frac{\$260(1+.2)}{1+.2-.75} - \frac{\$50(.75)}{1+.2-.75} = \$610.$$

3.3 Estimating retention rates

The previous section assumed that the retention rate r was known, but in practice it usually must be estimated from data. An organization observes when customers are acquired and can trace the history of payments. Some but not all of these customers will cancel. A customer who has not yet canceled is said to be *censored* and the organization will not have observed this customer's cancelation time yet. This section shows how to estimate retention rates from such data, allowing for some customers to be censored.

We want to estimate the retention rate for some group of $n_0 + n_1$ customers that were acquired in the past, where n_1 of the customers have already canceled (so that defection time t has been observed), while n_0 others are still active (censored). Let T denote a random variable and t be an observed cancelation time, that is, a realization of T. If customer i has already canceled, let t_i be the observed times of defection, so that customer i has been active $t_i - 1$ periods and cancels at time t_i. For those still active, let c_i be the time of censoring so that customer i has been active c_i periods and the company knows that the time of cancelation for this customer is $T_i > c_i$.

EXAMPLE 3.8: Educational service provider

A company provides a supplementary educational service to junior high-school students in Asia. Those who enroll in the service may cancel at any time. Each month the company sends the student a booklet of reading materials and exercises to supplement their lessons at school. During the month, the student completes the exercise book and returns the materials to the company for grading and feedback. This example illustrates the type of data that such an organization has.

The table below shows one year of transaction history for eight customers who were acquired during the year. An R indicates that the customer was retained and a C indicates that the customer canceled. Customers enroll and make their first payment during the period before the first R. Customers do not make a payment during the period in which they cancel. Customer id = 1 subscribed to the service in month 4, continued to use the service during month 7, and then canceled in month 8, so that $t_1 = 3$ indicates a cancelation during the third month of life.

	Month number													
Id	1	2	3	4	5	6	7	8	9	10	11	12	t	c
1						R	R	C					3	
2		R	R	R	R	R	R	R	R	R	R	C	11	
3	R	R	R	R	R	R	R	R	R	R	R	R		12
4						R	R	R	R	C			5	
5	R	R	R	R	R	R	R	R	R	R	R	R		12
6					R	R	R	R	R	R	R	R		8
7					R	R	R	R	R	R	R	C	8	
8	R	R	C										3	

Customer 2 joined during month 1 and canceled during month 12, which was month $t_2 = 11$ of life. Customer 3 is *censored*, having joined during month 1 and never canceled. The company only knows that the time of cancelation for this customer is greater than 12. Estimates that do not account for censoring underestimate the true retention rate. Customers 5 and 6 are also censored. Note that there are $n_1 = 5$ customers who have canceled in this study and $n_0 = 3$ customers who are censored.

We now derive the maximum likelihood estimate for the retention rate (r). Using equations 3.7 and 3.8, the likelihood function for r is

$$L(r) = \prod_{i=1}^{n_1} P(T = t_i) \prod_{i=1}^{n_0} S(c_i + 1) = \prod_{i=1}^{n_1} (1 - r) r^{t_i - 1} \prod_{i=1}^{n_0} r^{c_i}. \tag{3.13}$$

This is maximized by taking logs, differentiating with respect to r, equating the derivative to 0, and solving for r:

$$\hat{r} = \frac{\sum t_t + \sum c_t - n_1}{\sum t_t + \sum c_t} = 1 - \frac{n_1}{\sum t_t + \sum c_t}. \tag{3.14}$$

The caret or hat over the $\hat{r}$ indicates that it is an estimate of parameter r. This equation has an intuitive interpretation. The denominator of the second term gives the total number of periods in which customers can cancel (opportunities to cancel) and numerator is number of cancelations. Thus, the second term estimates the default rate.

EXAMPLE 3.9: Educational service provider (continued)

Estimate the retention rate for the educational service provider using the eight customers discussed in Example 3.8.

Solution The cancelation time was observed for $n_1 = 5$ customers and $n_0 = 3$ were censored. The sum of defection times is $\sum t_i = 3 + 11 + 5 + 8 + 3 = 30$ and the sum of censoring times is $\sum c_i = 12 + 12 + 8 = 32$. The retention rate is estimated as

$$\hat{r} = 1 - \frac{5}{30 + 32} \approx 92\%.$$

We can think of this as a series of coin flips. Each period the customer flips a coin, where heads indicates canceling, and tails indicates not canceling. Each customer flips the coin until heads comes up, at then stops flipping the coin. In total, these eight customers have flipped the coin 62 times, and only five heads were observed (cancelations). We estimate the defection rate as 5/62.

EXAMPLE 3.10: Educational service provider: one-year

The educational service provider in Example 3.8 acquired more than just eight customers. This problem examines a larger random sample of 671 customers who joined for the first time during this year. The table below gives a crosstab of customer status at the end of the study period (censored or canceled) versus the time of cancelation. For example, four customers canceled during their second month, 16 canceled during their third month, and three customers were censored at during month 1. There were $n_1 = 245$ customers who canceled the service during this year (and thus have observed cancelation times) and $n_0 = 426$ censored cases. Estimate the retention rate.

Table 3.2 Data for the one-year educational service provider example

	Time of Cancelation/Censoring												
Status	1	2	3	4	5	6	7	8	9	10	11	12	Total
Canceled	0	4	16	20	37	28	61	24	19	13	10	13	245
Censored	3	0	2	1	7	33	49	63	30	16	34	188	426
Total	3	4	18	21	44	61	110	87	49	29	44	201	671

Solution Using equation 3.14 requires computing the sum of cancelation and censoring times. For the three customers censored at time 1 the sum is three. The four customers who canceled or censored at time 2 have a sum of 4 × 2 = 8. The 18 who canceled or were censored at time 3 have a sum of 3 × 18 = 54, and so on. The sums are therefore $\sum t_i + \sum c_i = 1(3) + 2(4) + 3(18) + \cdots + 12(201) = 5,828$ and the estimated retention rate is

$$\hat{r} = 1 - \frac{245}{5828} = 95.8\%.$$

The numbers substituted into this equation can be produced using PROC MEANS. Here is the SAS program to read the data in and compute the numbers. The bigT variable gives the time of cancelation or censoring, cancel is a dummy variable that equals 1 for canceling and 0 for censoring, and count is the number of customers.

Program 3.1 Reading data for one-year educational service provider

```
DATA service1yr;
  INPUT bigT cancel count @@;
  LABEL bigT = "Cancelation Time (T)"
    cancel = "Dummy 1=cancel, 0=censored";
DATALINES;
2 1 4     3 1 16     4 1 20     5 1 37     6 1 28     7 1 61     8 1 24     9 1 19
10 1 13   11 1 10    12 1 13    1 0 3      3 0 2      4 0 1      5 0 7      6 0 33
7 0 49    8 0 63     9 0 30     10 0 16    11 0 34    12 0 188
RUN;
```

```
PROC MEANS DATA=service1yr SUM MAXDEC=0;
  VAR cancel bigT;
  WEIGHT count;
  OUTPUT OUT=answer SUM=;
RUN;
```

Variable	Label	Sum
cancel	Dummy 1=cancel, 0=censored	245
bigT	Cancelation Time (T)	5828

We can have SAS compute the retention rate and both the expected and median time until cancelation (ET) with a single SQL statement:

Program 3.2 PROC SQL program computing retention rates

```
PROC SQL;
  SELECT
    cancels LABEL="Number Cancels",
    flips LABEL="Opportunities to Cancel",
    Rhat LABEL="Retention Rate (r)" FORMAT=6.4,
    1/(1-Rhat) AS ET LABEL="E(T)" FORMAT=5.1,
    1+LOG(.5)/LOG(Rhat) AS median LABEL="Median(T)" FORMAT=3.0
  FROM (SELECT SUM(cancel) AS cancels,
      SUM(bigT) as flips,
      1-SUM(cancel)/SUM(bigT) AS Rhat
    FROM answer);
```

Number Cancels	Opportunities to Cancel	Retention Rate (r)	E(T)	Median(T)
245	5828	0.9580	23.8	17

3.4 Per-period cash flows m

The formulas for estimating CLV in Equations (3.12) and (3.11) depend on knowing m, the cash flows realized each period. This section addresses two sets of issues around m: which cash flows should be included, and whether m is constant over time and customers.

3.4.1 Deciding which cash flows to include

Our goal is to arrive at a period contribution for each customer, which is revenues less variable costs, but deciding which cash flows to include in m can be complicated. Following Pfeifer et al. (2005), our definitions of both CP and CLV do not specify exactly which costs and revenues are included, and there can be different implementations depending on the decision being made (Roberts and Berger, 1999; Blattberg et al., 2008). A simple example of how this can become complicated is acquisition costs. If we are deciding whether to spend marketing resources to acquire a customer, then clearly the cost of acquiring the customer should be considered. If a customer has already been acquired and we are deciding, for example, whether to increase the retention marketing efforts (for example, by increasing the number of sales calls), then acquisition and other sunk costs are not relevant.

Roberts and Berger (1999, p. 184) give two general guidelines for making decisions about which costs to include:

- All relevant costs of servicing the customer should be included.
- Exclude fixed costs as appropriate.

In addition, Blattberg et al. (2009, p. 164) emphasize the importance of making cost assumptions explicit and justifying which are included or excluded.

Which costs are relevant depend on the nature of the business and decision being made. Shepard (1999) gives examples of different profit-and-loss statements for various businesses. For example, in the case of a direct-to-consumer retailer the revenues would include gross sales and other fees such as shipping and handling. Deducting returns gives *net sales*. Relevant costs of sales can include the cost of goods sold, costs of lost or damaged items, bad debt, order processing costs, shipping costs, and return processing costs. Promotional expenses can include components such as creative development, media (for example, paying a search engine for keywords, paying a web site to display a banner, list rentals, postage, and so on), paper and printing. For some applications a contribution to overhead is also deducted.

We have stressed the importance of quantifying the effect of marketing actions on *incremental* CLV—what is the *difference* in CLV between when the organization takes the marketing action and when it does not? When incremental CLV is the relevant outcome, decisions about which costs to include are somewhat simpler. Because incremental CLV is the *difference* between CLV with and without the contact or action, fixed costs cancel out. For example, if the decision is whether or not to increase the number of marketing contact points, fixed costs such as executive salaries, rent and depreciation are the same whether or not the number of contacts is increased and therefore have no impact on incremental CLV. All costs associated with the increase in marketing contact points are relevant, such as the cost of developing additional creative content, and purchasing additional media, paper, printing, and so on.

EXAMPLE 3.11: Full or marginal costing

This example is based on Blattberg et al. (2009, Issue 16). Suppose, for the sake of simplicity, that a firm has only two customers. Customer 1 generates net revenues (net of variable costs) of $500 and Customer 2 generates only $150. The company has fixed costs of $400, regardless of the number of customers. These fixed costs pay for executive salaries, a call center, and so on. Averaging these fixed costs between both customers assigns $200 to each. CLV for the two customers is computed the table below.

Customer	Marginal Costing	Full Costing
1	CLV = \$500	CLV = \$500 − \$200 = \$300
2	CLV = \$150	CLV = \$150 − \$200 = −\$50

Under marginal costing both customers have positive CLV, while under full costing Customer 2 has a negative CLV. It seems that the firm would be better of without this customer, and the firm should allow Customer 2 to churn. The company would then be left with the following profits:

	With Customer 1	With Customers 1 and 2
Net Revenues	\$500	\$500 + \$150 = \$650
Fixed Costs	\$400	\$400
Total Profits	\$100	\$250

Clearly the firm is better off retaining Customer 2. Again, we should exclude fixed costs as appropriate. See Roberts and Berger (1999, pp. 184–6) for a similar example.

Sometimes costs are semi-variable, where they are constant within a given range, but then jump once a threshold is crossed Blattberg et al. (2009, p. 164). For example, existing call centers, web services, and fulfillment centers can service some range of customers. If the decision is whether to acquire customers, the fixed costs are not relevant as long as acquiring customers does not necessitate adding additional capacity. At some point, however, a threshold will be crossed and the existing infrastructure will be incapable of serving the customer base in an efficient way. At this point the cost of, for example, an additional fulfillment center becomes relevant.

3.4.2 The constant cash-flow assumption

The SRM assumes that m is constant over time and customers. In some situations this is a reasonable assumption. For example, the amount of revenue that an ISP receives from its customers each month is usually constant and, if we can assume the marginal costs of serving the customer are also constant over time and customers, then profit is also constant. We now give several examples of when profits are not constant and show how to estimate CLV in these situations.

EXAMPLE 3.12: Educational service provider

All subscribers to the educational service provider in Example 3.8 pay the same monthly fee, so revenue is constant over time and customers. Costs, however, vary. All customers receive the same materials, which have marginal costs for the paper, printing and postage. Some students will return the materials for grading, generating additional marginal costs for return postage, labor for grading, and an additional postage for returning the materials to the student. Thus, there are two possible costs, depending on whether the materials are returned.

EXAMPLE 3.13: Movie service

Consider a company such as Netflix that mails movies to its members. Suppose a family pays m per month for a service where the company mails the family a single movie, the family watches the movie and mails it back, the company sends another movie, and so on. Each time the family receives and returns a movie the company encounters marginal costs selecting the movie, packaging it, mailing it to the customer, paying the return postage, and placing the movie back into inventory. While the marginal revenue per month is constant over time, the marginal costs are proportional to the number of movies.

EXAMPLE 3.14: Electronic coupon company

An Internet coupon company acquires members who give their e-mail address and agree to receive coupons for products that interest them. The company sends coupons to their members until they opt out. Each time a member clicks on a coupon the company receives a payment from the coupon sponsor. The marginal revenue depends on the number of coupons that a member decides to click on and thus is not constant over time.

EXAMPLE 3.15: Cell phone provider

Cellular phone customers often have contracts where they pay some (constant) monthly fee for a certain number of minutes that can be used within a certain territory. Customers who use more than this number of minutes or roam outside the territory, however, pay a penalty, so revenue is not constant over time.

When the period payment fluctuates over time we represent it as a random variable M. We use the lower-case m when it is an observed value or a constant. We can conveniently substitute an average into equation 3.12 or 3.11 whenever M and T are statistically *independent*; otherwise, the problem is more complicated, and we have to use the methods discussed in the next chapter.

Before proving this, we review what it means to be independent and when this condition is or is not satisfied. Estimating CLV when M is not constant requires averaging over two sources of randomness, the time until cancelation T and the payment amount M. Loosely speaking, M and T are independent if one does not affect the other, that is, if information about the defection time reveals nothing about the likely payment amounts. For example, if customers must pay a penalty fee for canceling before a contract has ended, cancelation time and payment amount are dependent because a customer who does not cancel generates the normal cash flow, whereas one who does generates the normal cash flow plus the penalty. Information about the cancelation time thus changes the payment amount. This will be discussed further in Example 4.2.

We now discuss the rationale for needing independence. Suppose that the discount rate is 0 and T and M are independent. Recall from probability theory that the covariance between T and M is $\text{Cov}(T, M) = E(TM) - E(T)E(M)$, so that under independence the covariance is 0, $E(TM) = E(T)E(M)$, and

$$E(\text{CLV}) = E((T-1)M) = E(TM - M) = E(T)E(M) - E(M).$$

One can estimate the average monthly payment $E(M)$ and *separately* estimate the mean time of attrition $E(T)$. The same is true when the discount factor is positive, because discounting amounts to computing a function of random variable T, and the expectation of product of functions of random variables equals the product of the expectations of the functions.

EXAMPLE 3.16: Educational service provider: different costs

Suppose that the profit for a customer who returns the materials for grading is \$20 per month, the profit is \$28 per month for a customer who does *not* return the materials, and that 60% of customers return. Payments occur at the beginning of a period. Find CLV using a monthly discount rate of 1% and a retention rate of 95.8%.

Solution The profit is \$20 for 60% of the customers and \$28 for the remaining 40%, so the average profit per customer is

$$E(M) = \$20(.6) + \$28(.4) = \$23.20.$$

We can substitute this into equation 3.11:

$$E(\text{CLV}) = \frac{\$23.20(1+.01)}{1+.01-.958} \approx \$451.$$

3.5 Chapter summary

The retention model can be used by organizations that acquire customers who generate some cash flow m (the payment), each period until the customer cancels the service in period T (the cancelation time). After canceling, customers do not return. Each period the organization retains some percentage r, called the *retention rate*, of its remaining customers. The retention rate is assumed to be constant over time and customers.

The cancelation time T is a random variable with a *geometric distribution*. The probability that a customer is retained for $t-1$ periods and then cancels at time t is $P(T=t)=r^{t-1}(1-r)$. The *survival function*, which is the probability that a customer survives at least until the beginning of period t, is $P(T \geq t)=r^{t-1}$. The expected time of cancelation is $E(T)=1/(1-r)$. The α quantile is given by $[P_\alpha = 1+\log(1-\alpha)/\log r]$, where $[\cdot]$ is the greatest integer function.

CLV is the expected sum of discounted future cash flows from a customer. Suppose the discount rate is d. When customers agree to make n payments in the future and are not allowed to cancel (that is, $r=100\%$), CLV is computed using the formulas for *annuities*, PV_n, given in the table below. *Ordinary annuities* assume that payments are come at the end of a period while *immediate annuities* assume payments come at the beginning of a period. When customers are allowed to cancel ($r<100\%$), the formulas for expected CLV are also provided below.

Quantity	Cash Flow Comes At	
	End of Period	Beginning of Period
PV_n	$\sum_{t=1}^{n} \frac{m}{(1+d)^t} = m\frac{1-(1+d)^{-n}}{d}$	$\sum_{t=0}^{n-1} \frac{m}{(1+d)^t} = m\frac{(1+d)[1-(1+d)^{-n}]}{d}$
$E(\text{CLV})$	$\frac{mr}{1+d-r}$	$\frac{m(1+d)}{1+d-r}$

Estimating r in practice is complicated by the fact not all customers have canceled. Those who have not canceled are said to be *censored*. To estimate retention rate r from a group of customers, suppose that n_1 have canceled with cancelation times $t_1,\ldots,t_{n_1}$, and n_0 have not yet canceled with censoring times $c_1,\ldots,c_{n_0}$. Then the retention rate can be estimated by

$$\hat{r} = 1 - \frac{n_1}{\sum d_t + \sum c_t}.$$

In deciding which revenues and costs to include the per-period cash flows, all relevant costs of servicing the customer should be included and fixed costs should be excluded as appropriate. When the period payment M fluctuates, the average of M can be substituted into the above formulas provided that M is independent of T.

Notes

[1] Some readers may recall that there is a special formula for evaluating the expected value of a *linear transformation*, $E(aX+b)=aE(X)+b$. For this problem, $E(6X-2)=6\times 1.3-5=2.8$, which matches the answer given in the table above. We will, however, be evaluating the expected value of *nonlinear* functions of a random variable where this formula will not work. For this reason, we illustrate how to evaluate equation 3.6 directly from the definition in this example.

[2] This definition of a quantile is for continuous distributions and serves our purpose. Discrete random variables are more complicated. See Hyndman and Fan (1996) for discussion.

About The Author

Edward C. Malthouse is the Theodore R. and Annie Laurie Sills Professor of Integrated Marketing Communications and Industrial Engineering at Northwestern University and the Director of Research for the Spiegel Initiative on Database and Digital Marketing. He was the co-editor of the *Journal of Interactive Marketing* from 2005–2011. He obtained his PhD in 1995 in computational statistics from Northwestern University and completed his postdoctoral fellowship at the Kellogg School of Management at Northwestern.

Malthouse's research interests center on media marketing, database marketing, advertising, new media, and integrated marketing communications. He is the co-editor of *Medill on Media Engagement*, and he has published articles in numerous journals, including the *Journal of Consumer Psychology, Journal of Interactive Marketing, Data Mining and Knowledge Discovery, Journal of Broadcasting and Electronic Media, International Journal of Market Research,* and *Journal of Media Business Studies*. He teaches undergraduates, graduates, and executives, and he has been a visiting professor at universities in Japan, China, and Europe.

Learn more about this author by visiting his author page at support.sas.com/malthouse. There you can download free chapters, access example code and data, read the latest reviews, get updates, and more.