Exploring the Distribution of Customer Lifetime Value (in Contractual Settings)

Peter S. Fader
www.petefader.com

Bruce G. S. Hardie
www.brucehardie.com[†]

October 2017

1 Introduction

Many MBA-level introductory marketing courses now cover the concept of customer lifetime value (CLV), including a case or exercise in which students are expected to perform some simple CLV calculations. As discussed in Fader and Hardie (2014b), there are a number of problems with these introductory examinations of the topic, which means they are of limited value to anyone actually interested in computing CLV in practice.[1] But beyond the basic calculation of CLV — and the validity of the methods used to do so — these initial classroom explorations generally fall short in other ways. In this note we introduce a number of ideas that have been ignored in standard discussions of CLV, namely the distribution of CLV and quantifying the uncertainty about the value and future size of a group of customers.

2 Motivating Problem — Part 1

Consider a company with a subscription-based business model that acquires 1000 customers (on annual contracts) at the beginning of Year 1. Table 1 reports the pattern of renewals by this cohort over the subsequent four years. (A cohort is a group of customers acquired at the same time.)

Let us suppose the average net cash flow per customer is $100/year, which is booked at the beginning of the contract period. Let us also assume a discount rate of 10% to reflect the time value of money.

[1] See Bendle and Bagga (2017) for an additional critique of such introductory material.

ID	Year 1	Year 2	Year 3	Year 4	Year 5
0001	1	1	0	0	0
0002	1	0	0	0	0
0003	1	1	1	0	0
0004	1	1	0	0	0
0005	1	1	1	1	1
⋮		⋮		⋮	
0998	1	0	0	0	0
0999	1	1	1	0	0
1000	1	0	0	0	0
	1000	631	468	382	326

Table 1: Pattern of year-on-year renewals for the cohort of 1000 customers acquired at the beginning of Year 1.

Consider the following question:

> Assuming our current prospect pool has the same characteristics as that from which these customers were acquired, what is the expected value of a new customer (ignoring any customer acquisition costs)?

According to these data, the probability of a customer renewing at the end of the first year (and therefore being a customer in the second year) is $631/1000 = 0.631$. The probability that a randomly chosen member of the cohort of 1000 customers is still a customer in the third year is $468/1000 = 0.468$. And so on. Therefore, standing at the beginning of Year 1, the (discounted) expected value of a customer can initially be calculated as

$$\$100 + \$100 \times \frac{0.631}{1.1} + \$100 \times \frac{0.468}{(1.1)^2} + \$100 \times \frac{0.382}{(1.1)^3} + \$100 \times \frac{0.326}{(1.1)^4},$$

which equals \$247.01.

Let us consider an alternative way of performing this calculation. We first note from Table 1 that 369 individuals are customers for just one year, 163 for just two years, 86 for just three years, and 56 for just four years. A total of 326 make their fourth renewal and are therefore customers for five years.

- Anyone who stays with the firm for just one year is worth \$100.
- Anyone who stays with the firm for just two years is worth

$$\$100 + \frac{\$100}{1.1} = \$190.91.$$

- Anyone who stays with the firm for just three years is worth

$$\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} = \$273.55.$$

- Anyone who stays with the firm for just four years is worth

$$\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} + \frac{\$100}{(1.1)^3} = \$348.69.$$

- Anyone who is a customer for five years is worth

$$\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} + \frac{\$100}{(1.1)^3} + \frac{\$100}{(1.1)^4} = \$416.99$$

Therefore, standing at the beginning of Year 1, the expected discounted value of an as-yet-to-be-acquired customer is

$$\begin{aligned}\$100.00 \times \frac{369}{1000} + \$190.91 \times \frac{163}{1000} + \$273.55 \times \frac{86}{1000} \\ + \$348.69 \times \frac{56}{1000} + \$416.99 \times \frac{326}{1000} = \$247.01.\end{aligned}$$

The problem with this answer (whichever way it is computed) is that it ignores any cash flow the customer might generate after Year 5. Who says their maximum "lifetime" as a customer is limited to five years? We would expect that some of the 32.6% of customers who survive to Year 5 survive into Year 6 and are therefore worth more than the \$416.99 associated with their still being a customer in Year 5. To get a true sense of (expected) customer lifetime value, we need to know the probability that the customer is still alive in Year 6, Year 7, and so on.

Let us step back and introduce some notation and terminology. More formally, the survivor function $S(t)$ is the probability that a randomly chosen member of the original cohort of customers survives beyond time t. With reference to Figure 1, a customer is "born" at $t = 0$ (the beginning of Year 1) and therefore, by definition, $S(0) = 1$. The empirical survivor function (i.e., the survivor function computed directly from the data) for this cohort is $S(1) = 0.631$, $S(2) = 0.468$, $S(3) = 0.382$ and $S(4) = 0.326$.

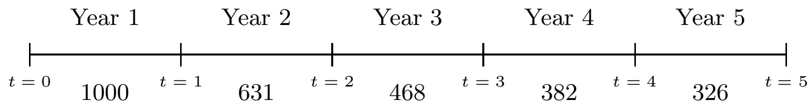


Figure 1: Summary of the cohort's subscription's renewal behavior.

We also note that the proportion of Year 1 customers who are customers in Year 2 is $631/1000 = 0.631$. Similarly, the proportion of Year 2 customers who are customers in Year 3 is $468/631 = 0.742$. This is captured by the notion of the *retention rate*, denoted by $r(t)$, which is the proportion of customers who survived beyond $t-1$ who also survive beyond t. The empirical retention rates for this cohort are $r(1) = 0.631$, $r(2) = 0.742$, $r(3) = 0.816$ and $r(4) = 0.853$. Retention rates can be computed directly from the data (as above) or via the survivor function using the following formula:

$$r(t) = \frac{S(t)}{S(t-1)}\,,\ t = 1, 2, 3, \ldots$$

Equivalently, given knowledge of the retention rates, we can compute the survivor function using the following *forward recursion*:

$$S(t) = \begin{cases} 1 & \text{if } t = 0 \\ r(t) \times S(t-1) & \text{if } t = 1, 2, 3, \ldots \end{cases} \tag{1}$$

Using this notation, our initial expected discounted value of a customer calculation performed above can be written as

$$\$100 \times \sum_{t=0}^{4} \frac{S(t)}{(1.1)^t}\,.$$

If we wish to determine the expected *lifetime* value of a customer, we need to compute

$$E(CLV) = \$100 \times \sum_{t=0}^{\infty} \frac{S(t)}{(1.1)^t}\,. \tag{2}$$

The challenge is how to compute $S(5)$, the probability the customer is still with the firm in year 6, $S(6)$, the probability the customer is still with the firm in year 7, and so on.

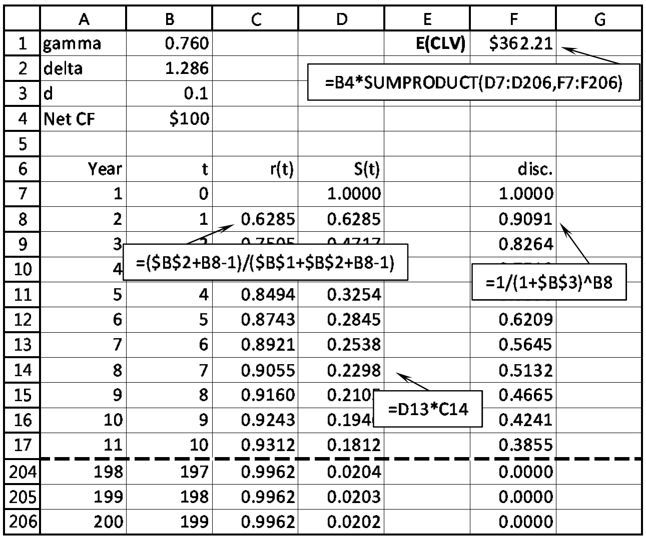
A simple solution is to use the beta-geometric (BG) distribution to project customer retention and survival into the future — see Fader and Hardie (2007, 2014a) for details. Of particular interest is the following expression for the retention rate,

$$r(t \mid \gamma, \delta) = \frac{\delta + t - 1}{\gamma + \delta + t - 1}\,,\ t = 1, 2, 3, \ldots, \tag{3}$$

where γ and δ are the two parameters of this distribution.

Fitting the BG model to the summary data given in Figure 1, Fader and Hardie (2014a) arrive at the following parameter estimates: $\hat{\gamma} = 0.760$ and $\hat{\delta} = 1.286$. Given these parameter estimates, we can use (1) and (3) to compute $S(t)$ as far into the future as desired. We can then compute $E(CLV)$ using (2).

For practical purposes, we treat a 200-year time horizon as being effectively equivalent to infinity. The Excel worksheet we use to do perform these calculations is shown in Figure 2.[2] We note that our estimate of the value of a new customer increases by \$115 to \$362 (ignoring cents for now).



	A	B	C	D	E	F	G
1	gamma	0.760			E(CLV)	$362.21	
2	delta	1.286					
3	d	0.1					
4	Net CF	$100					
5							
6	Year	t	r(t)	S(t)		disc.	
7	1	0		1.0000		1.0000	
8	2	1	0.6285	0.6285		0.9091	
9	3	[illegible]	[illegible]	[illegible]		0.8264	
10	4	[illegible]	[illegible]	[illegible]		[illegible]	
11	5	4	0.8494	0.3254		[illegible]	
12	6	5	0.8743	0.2845		0.6209	
13	7	6	0.8921	0.2538		0.5645	
14	8	7	0.9055	0.2298		0.5132	
15	9	8	0.9160	0.210[illegible]		0.4665	
16	10	9	0.9243	0.194[illegible]		0.4241	
17	11	10	0.9312	0.1812		0.3855	
204	198	197	0.9962	0.0204		0.0000	
205	199	198	0.9962	0.0203		0.0000	
206	200	199	0.9962	0.0202		0.0000	

Figure 2: Screenshot of the worksheet `Calculating E(CLV)`.

3 Moving Beyond a Point Estimate

Let us think about our estimate of customer lifetime value. We say that we would expect a customer to be worth \$362. However, if we reflect on the alternative initial calculation performed above, no single customer will actually be worth that specific amount. We would expect approximately 37% of the acquired customers to remain with us for just one year and therefore each be worth \$100, approximately 16% of the acquired customers to remain with us for just two years and therefore each be worth \$190.91, and so on. Our $E(CLV)$ number is an average across a group of acquired customers but need not be true for any individual customer. Let us therefore consider the *distribution* of customer lifetime value.

With reference to the worksheet `Distribution of CLV (I)` (Figure 3), we compute the distribution in the following manner.

- Sticking with the 200-year time horizon used above, an individual's "lifetime" ranges from one year to 200 years (column `F`). (We will reconsider this arguably unrealistic assumption in Section 7.)
- For each possible lifetime, we compute the lifetime value of a customer with that lifetime (column `G`) and the probability that a randomly

[2]This worksheet is part of the workbook `distribution_of_CLV_contractual.xlsx`.

chosen member of the cohort has a lifetime of that length (column `H`).[3]

	A	B	C	D	E	F	G	H	I
1	gamma	0.760							
2	delta	1.286							
3	d	0.1							
4	Net CF	$100							
5									
6	t	r(t)	S(t)	disc.		Lifetime	CLV	P(Lifetime)	
7	0		1.0000	1.0000		1	100.00	0.3715	
8	1	0.6285	0.6285	0.9091		2	190.91	0.1568	
9	2	0.7505	0.4717	0.8264		3	273.55	0.088	=C7-C8
10	3	0.8122	0.3831	0.7513		=G7+B$4*D8	348.69	0.057	
11	4	0.8494	0.3254	0.6830		5	416.99	0.0409	
12	5	0.8743	0.2845	0.6209		6	479.08	0.0307	
13	6	0.8921	0.2538	0.5645		7	535.53	0.0240	
14	7	0.9055	0.2298	0.5132		8	586.84	0.0193	
15	8	0.9160	0.2105	0.4665		9	633.49	0.0159	
16	9	0.9243	0.1946	0.4241		10	675.90	0.0134	
17	10	0.9312	0.1812	0.3855		11	714.46	0.0114	
133	126	0.9940	0.0285	0.0000		127	1099.99	0.0002	
134	127	0.9941	0.0284	0.0000		128	1099.99	0.0002	
135	128	0.9941	0.0282	0.0000		129	1099.99	0.0002	
136	129	0.9942	0.0280	0.0000		130	1100.00	0.0002	
137	130	0.9942	0.0279	0.0000		131	1100.00	0.0002	
138	131	0.9942	0.0277	0.0000		132	1100.00	0.0002	
139	132	0.9943	0.0275	0.0000		133	1100.00	0.0002	
203	196	0.9961	0.0204	0.0000		197	1100.00	0.0001	
204	197	0.9962	0.0204	0.0000		198	1100.00	0.0001	
205	198	0.9962	0.0203	0.0000		199	1100.00	0.0001	
206	199	0.9962	0.0202	0.0000		200	1100.00		

Figure 3: Screenshot of the worksheet `Distribution of CLV (I)`.

We note that beyond a certain point in time, the customer's lifetime value appears to be the same ($1100), regardless of how long they remain a customer. Recall from high school algebra that, for $0 < k < 1$, the infinite geometric series $\sum_{n=0}^{\infty} ak^n$ has the solution $a/(1-k)$. Letting $a = \$100$ and $k = 1/1.1$, it follows that the present value of the income stream of a customer who lives forever is

$$\frac{\$100}{1-\frac{1}{1.1}} = \$1100.$$

The value of a customer who remains with the firm for 200 years is less than 1/1,000th of a cent below this limit.

The final steps in computing the distribution of CLV are undertaken in the worksheet `Distribution of CLV (II)`. We first make a copy of

[3]Quiz: `=SUMPRODUCT(G7:G205,H7:H205)` equals $339.97, which is less than our estimate of $E(CLV)$ computed above. Why is this the case? We are ignoring the approximately 2% of customers who the model predicts will remain with the firm for 200+ years, each of whom has a present value of $1100. This is corrected below. As already noted, the question of the time horizon used in our calculations is considered in Section 7.

`Distribution of CLV (I)`, round the CLV associated with each lifetime to the nearest dollar (column `I`) and then create a pivot table that gives us the probability of seeing each value of CLV. Note that the sum of these probabilities is less than 1.0 (cell `L66`); this is because we are truncating the lifetime distribution at 200 years. However, we know that $1100 is the highest possible CLV and so the probability of seeing a CLV of $1100 is simply one minus the probability of seeing a value less than this — see cell `O65`. The corresponding distribution of CLV is plotted in Figure 4.

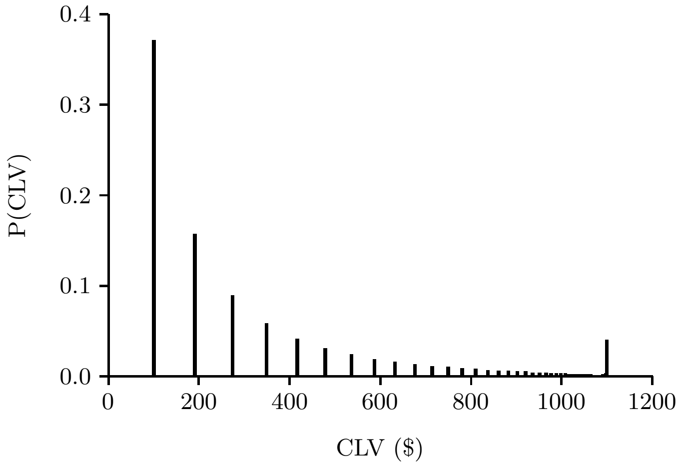


Figure 4: Estimated distribution of CLV

The mean of this distribution is $362.28 and the associated variance is $105,654.45; see cells `O3` and `O4` in the worksheet `Distribution of CLV (II)`. Note that the mean is seven cents higher than our previous estimate. This difference can be attributed to the fact that we have rounded the CLV numbers to the nearest dollar. We repeat this analysis in the worksheet `Distribution of CLV (III)`, this time rounding the CLV numbers to the nearest cent. The mean of the resulting distribution of CLV is $362.21.

4 Uncertainty About a New Cohort

Let us consider the following questions:

> We acquire 1000 customers from our current prospect pool, which is assumed to have the same characteristics as that from which the customers whose behavior is documented in Table 1 were acquired.
>
> i) What is the distribution of the total present value of these customers? What is the probability that it is between $350,000 and $370,000?

ii) What is the distribution of the number of individuals who remain customers of the firm for more than three years? What is the probability that it has 370 or fewer surviving customers?

Let us denote the value of a new cohort of 1000 customers by CoV (cohort value). The expected value of this cohort is obviously $1000 \times \$362.28 = \$362,280$. But what is the distribution around this mean?

- What is the lowest possible value of this group of customers? The answer is $1000 \times \$100 = \$100,000$, which occurs when each and every new customer only survives for one year. The probability of this occurring is extremely small: $(0.371)^{1000} = 2.37 \times 10^{-431}$.
- What is the highest possible value of this group of customers? The answer is $1000 \times \$1{,}100 = \$1{,}100,000$, which would be realized if everyone remains a customer for a very (and unrealistically) long time. The probability of this occurring is even smaller: $(0.040)^{1000} = 1.15 \times 10^{-1398}$.

We compute the distribution of CoV between these limits using simulation. For 1000 simulated customers, we determine their individual CLVs and sum up the values to arrive at our estimate of CoV. We repeat this 1000 times (i.e., simulate the value of 1000 different cohorts, each containing 1000 simulated customers) and use these 1000 simulated values of CoV to create our distribution.[4]

The first step is to work out how to simulate the CLV of an individual customer. We will do so using the inverse transform method, also known as the inverse CDF method.

- With reference to the worksheet `Distribution of CoV (I)`, the first thing we do is copy across the estimated distribution of CLV (cells `A1:B60`). We can now compute the cumulative distribution function (CDF) of CLV. With reference to cells `C2:C60`, the probability that CLV is $\leq \$100$ is 0.371, the probability that CLV is $\leq \$191$ is 0.528, the probability that CLV is $\leq \$274$ is 0.617, and so on.
- The inverse sampling method sees us simulating CLV by drawing a uniform random number (using the `=RAND` function in Excel). If the value of this draw is between 0 and 0.371, we say that the individual's CLV is \$100. If the value of this draw is between 0.371 and 0.528,

[4] We could do this using the @RISK software, which performs Monte Carlo simulation in Excel, defining a discrete distribution (`RiskDiscrete()` using the discrete distribution created in `Distribution of CLV (II)`. For a problem such as this, we feel that the approach outlined here is actually simpler.

we say that the individual's CLV is $191. If the value of this draw is between 0.528 and 0.617, we say that the individual's CLV is $274. And so on. (What if the value of the draw is 0.720? We would say that the individual's CLV is $479.)

The next step is to simulate the CLVs for 1000 simulated customers, for each of the 1000 simulated cohorts. We do this in cells `F6:ALQ1005`, making use of the `=LOOKUP` Excel function to perform the lookup associated with the inverse sampling method. The resulting simulated CoV numbers are given in cells `F3:ALQ3`. We compute the resulting distribution of RCoV in the following manner.

- We copy the simulated CoV numbers into the worksheet `Distribution of CoV (II)`. Note that the numbers are not the same as those given in `Distribution of CoV (I)` because new random numbers are drawn (and we therefore have new simulated CoV numbers) whenever Excel recalculates the workbook.
- We use the Histogram tool in the Excel Analysis ToolPak to determine the number of draws that fall into the ten $5,000-wide bins between $325,000 and $400,000, and plot the resulting distribution in Figure 5.
- We note that the average of these 1000 draws is $362,278, which is very close to the expected value of $362,280 computed above. The sample variance is $105,759,002, which implies the standard deviation is $10,284.

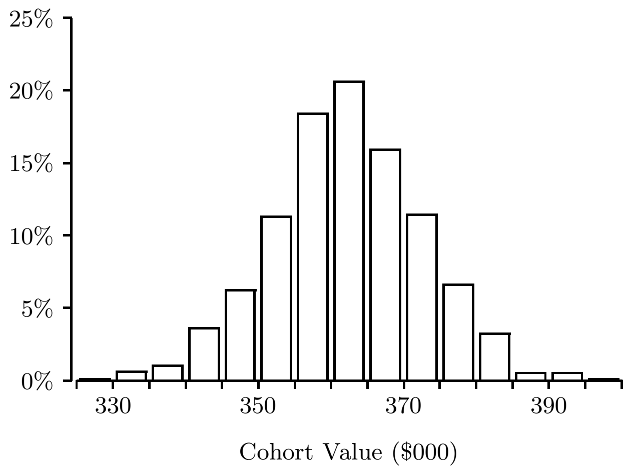


Figure 5: Simulated distribution of the value of a new cohort (CoV)

Looking at Figure 5, we see that the distribution of cohort value looks almost bell-shaped, suggesting it could be characterized by the normal distribution, even though the distribution of individual-level CLV (Figure 4) is most definitely not bell-shaped . Indeed, we can demonstrate this equivalence more formally.

- Recall that $E(CLV) = \$362.28$ and $\text{var}(CLV) = \$105{,}654.45$.
- As we computed above, the expected value of the group of 1000 new customers is $E(CoV) = \$362{,}280$.
- Assuming independence across customers (i.e., an individual's decision to cancel their contract is independent of whether or not other customers have done so), we know from introductory statistics that $\text{var}(CoV) = 1000 \times \text{var}(CLV) = \$105{,}654{,}450$, which implies the standard deviation is \$10,279.
- It follows from the Central Limit Theorem that the distribution of CoV is approximately normal with mean \$362,280 and standard deviation \$10,279.
- We overlay this distribution on the simulated distribution of CoV in Figure 6.

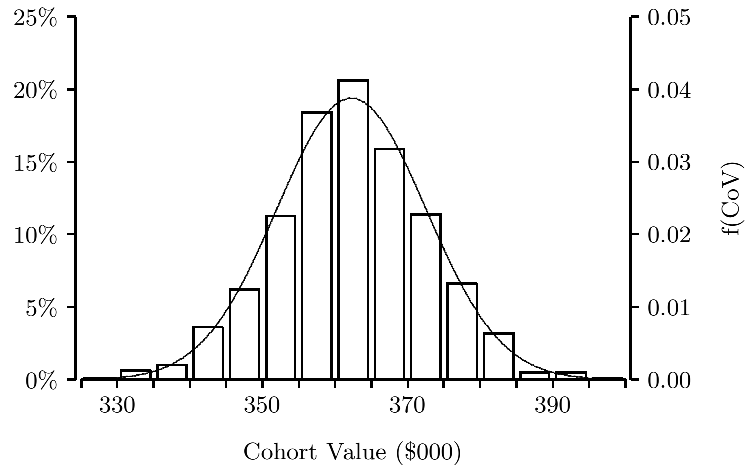


Figure 6: Comparing the simulated versus normal approximation of the distribution of the value of a new cohort (CoV)

Given the distribution of CoV (be it the simulated or normal approximation), we can now compute the probability that the value of the new cohort lies between \$350,000 and \$370,000.

- 662 of the 1,000 simulated values of total cohort value fall in this interval, which would gives us a probability of 0.662.

- Considering the normal approximation, the z-score associated with \$350,000 is $(350{,}000 - 362{,}280)/10{,}279 = -1.195$ while the z-score associated with \$370,000 is $(370{,}000 - 362{,}280)/10{,}279 = 0.751$. This implies the probability that the new cohort value falls in this interval is 0.658. This is computed in Excel as

```
=NORM.S.DIST(0.751,TRUE)-NORM.S.DIST(-1.195,TRUE)
```

Turning to the second question, let the random variable X be the numbers of customers who remain customers of the firm for more than three years. The (model-based) probability that a newly acquired customer remains with the firm for more than three years is $S(3) = 0.383$. The distribution of the number of customers (out of the original 1000) who are still around after three years is therefore binomial with $n = 1000, p = 0.383$. The mean and standard deviation of the binomial distribution are np $(= 383)$ and $\sqrt{np(1-p)}$ $(= 15.372)$, respectively.

For such a large n, it is usually easier to perform the calculations of interest using the normal approximation of binomial distribution. Following the Central Limit Theorem, we could simply assume that X is normally distributed with mean and standard deviation as given above. A more accurate approximation is the so-called normal approximation with continuity correction:[5]

$$P(X \leq x) \approx \Phi\left(\frac{x + 0.5 - np}{\sqrt{np(1-p)}}\right),$$

where Φ is the the standard normal CDF (i.e., $P(Z \leq z)$ for normally distributed variable Z with mean 0 and variance 1), which can be computed in Excel using the formula `=NORM.S.DIST(z,TRUE)`. (See the worksheet `Dist. of Number of Survivors`.) This distribution is plotted in Figure 7. (We have to go to the fourth decimal place to find a difference between the number obtained using the normal approximation and the true value computed using the binomial distribution.) The probability that the firm has 370 or fewer remaining customers after three years is 0.208.

5 Motivating Problem — Part 2

Consider the following question:

> We note that 326 of the original cohort of 1000 customers are still with the firm in Year 5. What is the expected *residual* value of this group of customers at the end of Year 5?

[5] Even more accurate approximations exist; see, for example, Lesch and Jeske (2009).

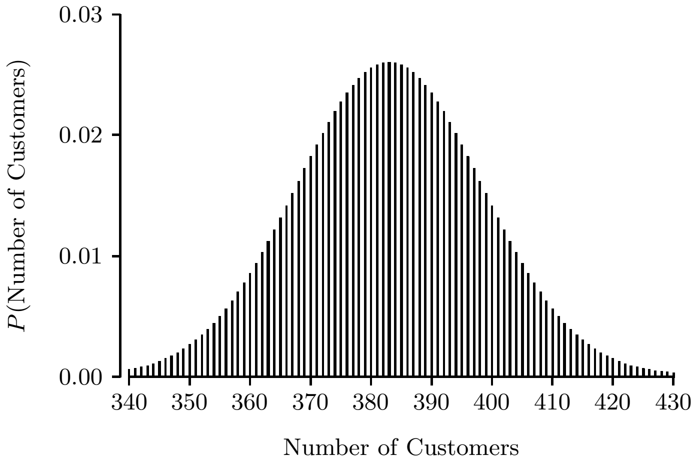


Figure 7: Distribution of the numbers of customers surviving beyond three years.

Let us first consider the task of computing the expected residual (i.e., remaining) lifetime value of an individual chosen at random from this group of 326 Year 5 customers.

- If they renew their contract, they will be worth \$100. If we assume that a dollar received tomorrow is worth a dollar today, the discount factor is 1. The probability that they are customer in Year 6 is the simply the probability that they renew at the end of Year 5, $r(5)$.
- If they renew their contract at the beginning of Year 7, the present value (at the end of Year 5) of \$100 received is 100/1.1. The probability that they are customer in Year 7 is the simply the probability that a Year 5 customer renews at the end of Year 5, $r(5)$, multiplied by the probability that a Year 6 customer renews at the end of Year 6, $r(6)$.
- And so on.

More generally, the probability that one of the original 1000 customers who is still with the firm in Year 5 (i.e., an individual who has renewed 4 times), survives beyond t (i.e., is still a customer in Year $t+1$) is $r(5) \times r(6) \times \cdots \times r(t)$. Denoting this quantity by $S(t \,|\, T > 4)$, we have

$$\begin{aligned} S(t \,|\, T > 4) &= r(5) \times r(6) \times \cdots \times r(t) \\ &= \prod_{i=5}^{t} r(i) \\ &= \prod_{i=1}^{t} r(i) \Big/ \prod_{i=1}^{4} r(i) \end{aligned}$$

$$= S(t)/S(4)$$

This is called the conditional survivor function.

It follows that we can compute the expected residual lifetime value of a randomly chosen Year 5 customer by evaluating

$$E(RLV) = \$100 \times \sum_{t=5}^{\infty} \frac{S(t \,|\, T > 4)}{(1.1)^{t-5}} \,. \tag{4}$$

As before, we treat a 200-year time horizon as infinity. We perform this calculation in the worksheet `Calculating E(RLV)` — see Figure 8. Our estimate of $E(RLV)$ is \$569.51.[6] This implies that, standing at the end of Year 5, the expected residual value of this group of customers is 326 $\times$\$569.51 = \$185,660.

	A	B	C	D	E	F	G
1	gamma	0.760			E(RLV)	\$569.51	
2	delta	1.286					
3	d	0.1					
4	Net CF	\$100					
5							
6	Year	t	r(t)		S(t|T>4)	disc.	
7	1	0					
8	2	1	0.6285				
9	3	2	0.7505				
10	4	3	0.8122				
11	5	4	0.8494				
12	6	5	0.8743		0.8743	1.0000	
13	7	6	0.8921		0.7800	0.9091	
14	8	7	0.9055		0.7063	0.8264	
15	9	8	0.9160		0.6470	0.7513	
16	10	9	0.9243		[illegible]	[illegible]	
17	11	10	0.9312		0.5569	0.6209	
209	203	202	0.9963		0.0614	0.0000	
210	204	203	0.9963		0.0612	0.0000	
211	205	204	0.9963		0.0610	0.0000	

=B4*SUMPRODUCT(E12:E211,F12:F211)

=C12

=E12*C13

=1/(1+\$B\$3)^(B13-5)

Figure 8: Screenshot of the worksheet `Calculating E(RLV)`.

6 The Distribution of RCoV

As a final exercise, we estimate the distribution of the residual value of the group of 326 cohort members who are still customers of the firm in Year 5.

[6] Why is this estimate of the residual lifetime value of a Year 5 customer so much higher than the expected value of an as-yet-to-be-acquired customer? The fact that an individual has survived into Year 5 means that they have a low underlying churn propensity; the individuals with high underlying churn propensities have left the company by the time we get to Year 5. This remaining pool of customers who have high retention rates will (on average) stay around for a long time. We do not have a large proportion of the pool of customers disappearing in the first few years of the calculation (as is the case when computing $E(CLV)$), and this results in a higher (residual) lifetime value.

(More formally, we are interested in the residual value of the cohort, which we denote by RCoV. We wish to compute the distribution of RCoV.) The logic is the same as that used to compute the distribution of the value of a new cohort of 1000 customers (i.e., the distribution of CoV).

As a first step, we need to compute the distribution of RLV for such customers (i.e., those who have made four renewals). With reference to the worksheet `Distribution of RLV (I)` (Figure 9), we compute the distribution in the following manner. Sticking with the 200-year time horizon used above, an individual's residual "lifetime" ranges from zero years to 200 years (column `F`). For each possible lifetime, we compute the residual lifetime value of a customer with that lifetime (column `G`) and the probability that a randomly chosen Year 5 customer has a residual lifetime of that length (column `H`).

	A	B	C	D	E	F	G	H	I
1	gamma	0.760							
2	delta	1.286							
3	d	0.1							
4	Net CF	$100							
5									
6	t	r(t)	S(t|T>4)	disc.		RL	RLV	P(RL)	
7	0								
8	1	0.6285						=1-C12	
9	2	0.7505							
10	3	0.8122							
11	4	0.8494				0	0.00	0.1257	
12	5	0.8743	0.8743	1.0000		1	100.00	0.0943	
13	6	0.8921	0.7800	0.9091		=G11+B$4*D12	190.91	0.0737	
14	7	0.9055	0.7063	0.8264			273.55	0.0593	
15	8	0.9160	0.6470	0.7513		4	348.69	[illegible]	
16	9	0.9243	0.5980	0.6830		5	416.99	[illegible]	=C13-C14
17	10	0.9312	0.5569	0.6209		6	479.08	0.0351	
138	131	0.9942	0.0851	0.0000		127	1099.99	0.0005	
139	132	0.9943	0.0847	0.0000		128	1099.99	0.0005	
140	133	0.9943	0.0842	0.0000		129	1099.99	0.0005	
141	134	0.9944	0.0837	0.0000		130	1100.00	0.0005	
142	135	0.9944	0.0832	0.0000		131	1100.00	0.0005	
143	136	0.9945	0.0828	0.0000		132	1100.00	0.0005	
144	137	0.9945	0.0823	0.0000		133	1100.00	0.0004	
209	202	0.9963	0.0614	0.0000		198	1100.00	0.0002	
210	203	0.9963	0.0612	0.0000		199	1100.00	0.0002	
211	204	0.9963	0.0610	0.0000		200	1100.00		

Figure 9: Screenshot of the worksheet `Distribution of RLV (I)`.

As was the case when computing the distribution of CLV, we note that beyond a certain point in time, the customer's residual lifetime value appears to be the same ($1100), regardless of how long they remain a customer.

The final steps in computing the distribution of RLV are undertaken in the worksheet `Distribution of RLV (II)`. We first make a copy of `Distribution of RLV (I)`, round the RLV associated with each lifetime to the nearest dollar (column `I`) and then create a pivot table that gives us the probability of seeing each value of RLV. The corresponding distribution of RLV is plotted in Figure 10.[7]

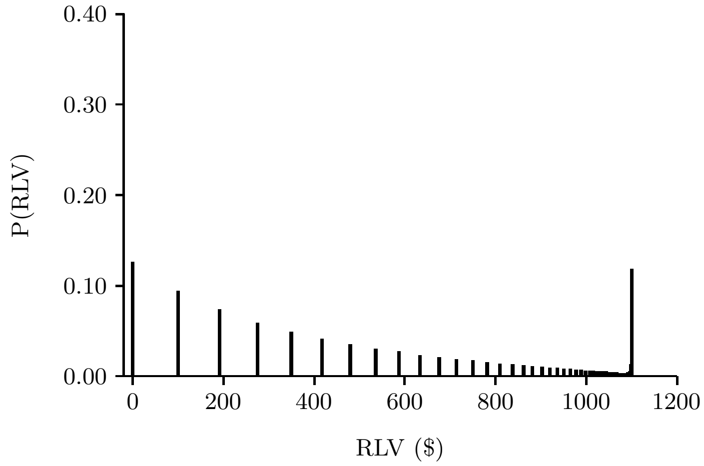


Figure 10: Estimated distribution of RLV

The mean of this distribution (cell `O3`) is \$569.56, and the variance (cell `O4`) is \$163,735.83. Note that this estimate of the mean is five cents higher than our previous estimate. This difference can be attributed to the fact that we have rounded the RLV numbers to the nearest dollar.

The distribution of RCoV can be computed in one of two ways:

- The easiest way of computing this is to use the normal approximation. The expected value of the group of 326 customers who are still with the firm in Year 5 is $326 \times E(RLV) = \$185{,}677$ and the associated variance $326 \times \text{var}(RLV) = \$53{,}377{,}881$, which implies the standard deviation is \$7,306.

- Using the same approach as for the value of a new cohort, we simulate the distribution in the worksheets `Distribution of RCoV (I)` and `Distribution of RCoV (II)`. The average of these 1000 draws is \$185,295 and the sample variance is \$54,291,447, which implies the standard deviation is \$7,368.

The associated distributions are plotted in Figure 11.

Given the distribution in RCoV (be it the simulated or normal approximation), we can make probabilistic statements about RCoV, which can be a useful input to various financial calculations. For example,

[7] Compare this distribution to that shown in Figure 4: the differences in the distribution of value for long-tenured customers (i.e., those active in Year 5) versus as-yet-to-be-acquired ones is quite dramatic. In particular, the fraction of customers in the \$1100 group rises from 4% for the latter group to 11.8% for the former. (See footnote 6 for a discussion of what lies behind this difference.)

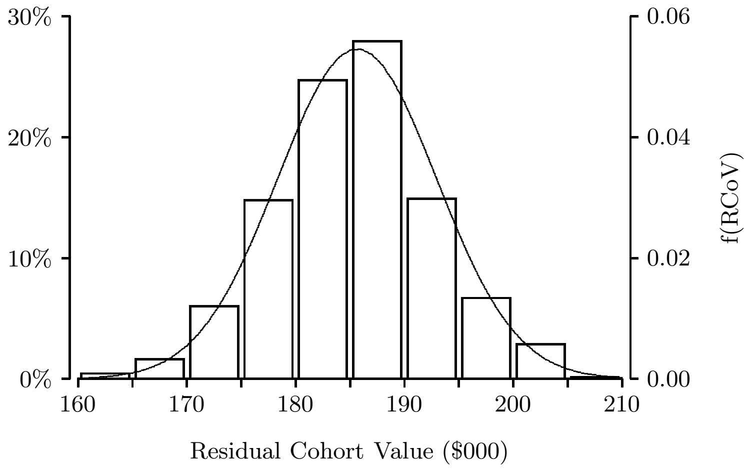


Figure 11: Comparing the simulated versus normal approximation of the distribution of the residual value of the cohort (RCoV).

- What is the probability that the RCoV will be less than $180,000?
  - 228 of the 1,000 simulated values of RCoV are less than $180,000, which gives us a probability of 0.228.
  - Using the normal approximation, $z = -0.752$, which implies a probability of 0.226.
- What is the probability that the RCoV will be between $175,000 and $195,000?
  - 823 of the 1,000 simulated values of RCoV fall in this interval, which would gives us a probability of 0.823.
  - Using the normal approximation, the z-score associated with $175,000 is -1.436 while the z-score associated with $195,000 is 1.301. This implies the probability that the RCoV falls in this interval is 0.828.

7 Discussion

In this note we have explored how to compute CLV, RLV, and the value of a cohort of customers using a simple and robust model for the duration of a customer's relationship with the firm in contractual settings.[8] Most

[8] In contractual settings, the existence of a contract between the customer and the firm means the firm will know when the customer has ended their relationship with the firm. In contrast, customers in noncontractual settings do not explicitly signal to the firm that they wish to terminate their relationship; as a result, the firm does not observe churn, only an absence of behavior. The equivalent calculations are quite different in that setting.

discussions of these concepts focus on the mean of the quantity of interest. We have gone further, demonstrating how to compute the distribution of these quantities.

While a lot more realistic than the typical problem considered in MBA-level marketing courses, our introductory problem is still quite simple. We now discuss a number of issues we have glossed over or ignored as we developed our solution.

i) When computing CLV we have performed our calculations using a 200-year time horizon as an approximation of infinity. Given the data in Figure 1, our model predicts that there is a 2% chance that a customer survives at least 200 years, which is clearly unrealistic. We have also noted that the maximum possible lifetime value of a customer is \$1100. Looking at the worksheet `Distribution of CLV (II)`, we see that a customer would have to remain a customer for 70 years to get within \$1 of that limit. Again, this seems unrealistic. The natural reaction is to suggest that we terminate the calculation at some earlier point in time. The problem then becomes what cutoff point to choose? If we compute (2) using a cutoff of 100 years, the maximum possible lifetime value is within eight cents of the maximum and our estimate $E(CLV)$ is \$362.20. (Recall that the original value was \$362.21.) If we choose a cutoff of 50 years, the maximum possible lifetime value is reduced to \$1091 and our estimate $E(CLV)$ is \$361.73. A cutoff of 30 years sees our estimate $E(CLV)$ dropping to \$357.82. Our take on this is that the overestimation of CLV associated with going out to "infinity" is minimal, and that rather than wasting time discussing an appropriate medium-term cutoff point, simply take the calculation out to "infinity."

ii) If we are uncomfortable with such an approach, there is nothing wrong with using a shorter-term period (e.g., five years). However, it is not appropriate to call the resulting quantity "lifetime" value as it ignores the value of customers who "live" beyond that time horizon.

iii) We have assumed a constant net cash flow per customer of \$100/year. In practice, we would expect it to increase over time. If it grows at annual rate of $g \times 100\%$, and assuming an annual discount rate of $d \times 100\%$, we can rewrite (2) as

$$E(CLV) = \$100 \times \sum_{t=0}^{\infty} S(t)\Big(\frac{1+g}{1+d}\Big)^t ,$$

which is equivalent to using a annual discount rate of $(d-g)/(1+g) \times 100\%$ in our calculations.

iv) We have also assumed that there is no variation in net cash flow across customers. It many situations this will not be a valid assumption. While using an average net cash flow per customer per period may be fine if we are simply interested in $E(CLV)$, it not appropriate when we are interested in the distribution of CLV. The gamma-gamma model of spend (Fader and Hardie 2013) is a natural starting point if the distribution of net cash flow per period can be characterized by a continuous unimodal distribution. If the nature of the contracts/subscriptions is such that there are a small number of "packages" each with a fixed value, the distribution of spend could be characterized by a finite mixture of multinomials or the Dirichlet-multinomial distribution. Given a distribution for net cash flow per period, the easiest way compute the distribution of CLV is by simulation. (The implementation of such models requires a level of analytical skill beyond that assumed for the analysis presented in this note.)

v) When we allow for variation in net cash flow across customers, the issue of a correlation between net cash flow per period and customer lifetime arises. In some cases, the menu of contracts can result in a negative correlation (e.g., when discounts are offered to those taking out a multi-year subscription). On the other hand, some analysts argue that customers who stay around for a long time spend more (although the evidence is mixed). The models for net cash flow per period suggested above assume that it is independent of the length of the customer's relationship with the firm. Extending these models to allow for dependence requires an even higher level of analytical skill beyond that assumed for the analysis presented in this note.

vi) We have talked about net cash flow, ignoring both spend on acquisition and retention activities, as well the impact of the firm's acquisition activities on retention.

    Our use of data on the behavior of previously acquired customers to predict the behavior of as-yet-to-be-acquired customers (as in our CLV calculations above) is based on the assumptions that i) the prospect pool is the same, ii) there have been no major economic or technological shocks in the intervening time period, and iii) there have been no major changes in the firm's acquisition activities (relative to its competition).

    It is important to appreciate how a shift in the firm's acquisition activities can affect the lifetime value of the acquired customers. Money spent on large discounts may attract the "wrong" type of customers, those who are more likely to churn in search of the next "good deal." Alternatively, money spent on carefully targeted acquisition activities will hopefully result in the acquisition of the "right" type of customers, those whose needs better match the firm's offerings and are therefore less likely to churn.

vii) Finally, we have considered the problem of computing lifetime value in contractual settings. As noted in footnote 8, the difference between contractual and noncontractual settings is that the loss of a customer is observed directly in the former setting but not in the latter. The fact that the firm does not observe churn means we cannot use the approaches explored above to compute lifetime value. Researchers have developed analytical tools for characterizing customer behavior and computing customer lifetime value in noncontractual settings (e.g., Fader et al. 2005, Fader et al. 2010).

References

Bendle, Neil T., and Charan K. Bagga (2017), "The Confusion About CLV in Case-Based Teaching Materials," *Marketing Education Review*, **27** (1), 27–38.

Fader, Peter S. and Bruce G.S. Hardie (2007), "How to Project Customer Retention," *Journal of Interactive Marketing*, **21** (Winter), 76–90.

Fader, Peter S. and Bruce G.S. Hardie (2013), "The Gamma-Gamma Model of Monetary Value." <`http://brucehardie.com/notes/025/`>

Fader, Peter S. and Bruce G.S. Hardie (2014a), "A Spreadsheet-Literate Non-Statistician's Guide to the Beta-Geometric Model." <`http://brucehardie.com/notes/032/`>

Fader, Peter S. and Bruce G.S. Hardie (2014b), "What's Wrong With This CLV Formula?" <`http://brucehardie.com/notes/033/`>

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005), "RFM and CLV: Using Iso-value Curves for Customer Base Analysis," *Journal of Marketing Research*, **42** (November), 415–430.

Fader, Peter S., Bruce G. S. Hardie, and Jen Shang (2010), "Customer-Base Analysis in a Discrete-Time Noncontractual Setting," *Marketing Science*, **29** (November–December), 1086–1108.

Lesch, Scott M. and Daniel R. Jeske (2009), "Suggestions for Teaching About Normal Approximations to Poisson and Binomial Distribution Functions," *The American Statistician*, **63** (3), 274–277.