Probability Models for Customer Lifetime Value Analysis
Day 1

Peter S. Fader
University of Pennsylvania

Bruce G.S. Hardie
London Business School

Stockholm School of Economics
June 3–4, 2024

PELOTON

Connected Fitness Subscriber Lifetime Value	FY 2019
Monthly subscription price	$39.00
×	
Subscriber LTV months	154
×	
(Subscription Contribution plus content costs for past use) divided by Subscription Revenue[1]	59.8%
LTV per Connected Fitness Subscriber	**$3,593**

Implied by 1 / **0.65%** Average Net Monthly Connected Fitness Churn

https://investor.onepeloton.com/static-files/73e7570a-6aca-47df-b080-a6a069139eed

NETFLIX®

Non-GAAP Estimated Subscriber Lifetime Value*

	Three Months Ended		
	March 31, 2002	March 31, 2003	Calculation Methodology
Monthly subscription charge	$ 19.95	$ 19.95	Standard subscription fee for three out program
Monthly churn	7.2%	5.8%	Reported churn rate
Implied subscriber lifetime (months)	13.9	17.2	Reciprocal of reported churn
Implied lifetime revenue	$ 277	$ 343	Implied subscriber life multiplied by monthly subscription charge
Cost of revenues	137	185	Reported costs of revenue margin multiplied by implied lifetime revenue
Gross profit per subscriber	140	158	
Gross Margin	*50.4%*	*46.1%*	

https://ir.netflix.net/files/doc_news/archive/a12fabbc-a83a-45a0-ae41-7afe37ff31ab.pdf

FORRESTER®

Using The Right Metrics For The Subscription-Based Business Model

The 10 Key Subscription Business Metrics To Track And Manage For Success

by Dan Bieler
November 3, 2020

Customer lifetime value (LTV)	LTV measures the value of a customer over their lifetime.	Average revenue per account (ARPA); gross margin (GM); customer churn (CC) $LTV = \frac{(ARPA * GM)}{CC}$	The average revenue per customer is $120. The gross margin is 80%. The customer churn rate is 20%. The lifetime value of the customer is $480.

PELOTON

Connected Fitness Subscriber Lifetime Value	FY 2019
Monthly subscription price	$39.00
×	
Subscriber LTV months	154
×	
(Subscription Contribution plus content costs for past use) divided by Subscription Revenue[1]	59.8%
LTV per Connected Fitness Subscriber	**$3,593**

Implied by 1 / 0.65% Average Net Monthly Connected Fitness Churn

What is wrong with this calculation?

Deriving the formula

	Month 1	Month 2	Month 3	$\cdots$	Month t
P(still a customer)	1	$1-c$	$(1-c)^2$		$(1-c)^{t-1}$
Contribution margin	$rev \times m$	$rev \times m$	$rev \times m$		$rev \times m$

$$
\begin{aligned}
E(CLV) &= \sum_{t=1}^{\infty} rev \times m \times (1-c)^{t-1} \\
&= rev \times m \times \sum_{s=0}^{\infty} (1-c)^{s} \\
&= rev \times m \times \frac{1}{c}
\end{aligned}
$$

A “correct” approach to the calculation

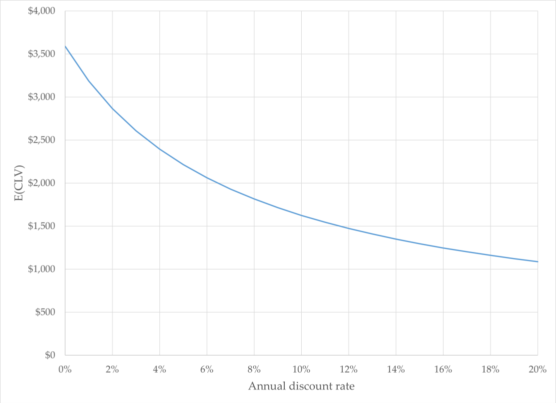
	Month 1	Month 2	Month 3		Month t
P(still a customer)	1	$1-c$	$(1-c)^2$		$(1-c)^{t-1}$
Contribution margin	$rev \times m$	$rev \times m$	$rev \times m$		$rev \times m$
Discount	1	$\frac{1}{1+d}$	$\frac{1}{(1+d)^2}$		$\frac{1}{(1+d)^{t-1}}$

$$E(CLV) = \sum_{t=1}^{\infty} rev \times m \times \frac{(1-c)^{t-1}}{(1+d)^{t-1}}$$

$$= rev \times m \times \frac{1+d}{d+c}$$

Impact of the time value of money

Monthly sub	$39.00
CM ratio	59.8%
Churn	0.65%



Exploring the impact of incorrect calculations

Peloton calculation:

$$\text{“}CLV\text{”} = rev \times m \times \underbrace{\frac{1}{c}}_{\text{expected lifetime, } E(L)}$$

“Correct”’ calculation:

$$E(CLV) = rev \times m \times \underbrace{\frac{1+d}{d+c}}_{\text{expected discounted lifetime, } E(DL)}$$

How does the over-estimation vary as a function of d and c?

$E(L)/E(DL)$ as a function of annual d and monthly c

	Annual discount rate																				
Monthly churn rate	0%	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%	11%	12%	13%	14%	15%	16%	17%	18%	19%	20%
0.5%	1.00	1.16	1.33	1.49	1.65	1.81	1.96	2.12	2.27	2.42	2.57	2.72	2.87	3.02	3.16	3.30	3.45	3.59	3.73	3.86	4.00
1.0%	1.00	1.08	1.16	1.24	1.32	1.40	1.48	1.56	1.63	1.71	1.78	1.86	1.93	2.00	2.08	2.15	2.22	2.29	2.36	2.42	2.49
1.5%	1.00	1.05	1.11	1.16	1.21	1.27	1.32	1.37	1.42	1.47	1.52	1.57	1.62	1.67	1.71	1.76	1.81	1.85	1.90	1.95	1.99
2.0%	1.00	1.04	1.08	1.12	1.16	1.20	1.24	1.28	1.31	1.35	1.39	1.42	1.46	1.50	1.53	1.57	1.60	1.64	1.67	1.71	1.74
2.5%	1.00	1.03	1.06	1.10	1.13	1.16	1.19	1.22	1.25	1.28	1.31	1.34	1.37	1.40	1.42	1.45	1.48	1.51	1.53	1.56	1.59
3.0%	1.00	1.03	1.05	1.08	1.11	1.13	1.16	1.18	1.21	1.23	1.26	1.28	1.30	1.33	1.35	1.37	1.40	1.42	1.44	1.47	1.49
3.5%	1.00	1.02	1.05	1.07	1.09	1.11	1.13	1.16	1.18	1.20	1.22	1.24	1.26	1.28	1.30	1.32	1.34	1.36	1.38	1.40	1.42
4.0%	1.00	1.02	1.04	1.06	1.08	1.10	1.12	1.13	1.15	1.17	1.19	1.21	1.23	1.24	1.26	1.28	1.30	1.31	1.33	1.35	1.36
4.5%	1.00	1.02	1.03	1.05	1.07	1.09	1.10	1.12	1.14	1.15	1.17	1.18	1.20	1.22	1.23	1.25	1.26	1.28	1.29	1.31	1.32
5.0%	1.00	1.02	1.03	1.05	1.06	1.08	1.09	1.11	1.12	1.14	1.15	1.16	1.18	1.19	1.21	1.22	1.23	1.25	1.26	1.27	1.29
5.5%	1.00	1.01	1.03	1.04	1.06	1.07	1.08	1.10	1.11	1.12	1.14	1.15	1.16	1.17	1.19	1.20	1.21	1.22	1.24	1.25	1.26
6.0%	1.00	1.01	1.03	1.04	1.05	1.06	1.08	1.09	1.10	1.11	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.23	1.24
6.5%	1.00	1.01	1.02	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.14	1.15	1.16	1.17	1.18	1.19	1.20	1.21	1.22
7.0%	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19	1.20
7.5%	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.16	1.17	1.18	1.19
8.0%	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.06	1.07	1.08	1.09	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.16	1.17	1.17
8.5%	1.00	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08	1.09	1.09	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.15	1.16
9.0%	1.00	1.01	1.02	1.02	1.03	1.04	1.05	1.06	1.06	1.07	1.08	1.09	1.10	1.10	1.11	1.12	1.12	1.13	1.14	1.15	1.15
9.5%	1.00	1.01	1.02	1.02	1.03	1.04	1.05	1.05	1.06	1.07	1.08	1.08	1.09	1.10	1.10	1.11	1.12	1.12	1.13	1.14	1.14
10.0%	1.00	1.01	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.06	1.07	1.08	1.08	1.09	1.10	1.10	1.11	1.12	1.12	1.13	1.14

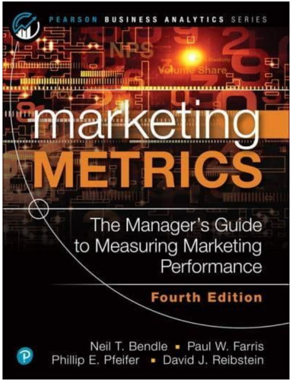
PELOTON

Connected Fitness Subscriber Lifetime Value	FY 2019
Monthly subscription price	$39.00
×	
Subscriber LTV months	154
×	
(Subscription Contribution plus content costs for past use) divided by Subscription Revenue[1]	59.8%
LTV per Connected Fitness Subscriber	**$3,593**

Implied by 1 / 0.65% Average Net Monthly Connected Fitness Churn

What else is wrong with this calculation?

Understanding retention rates



Retention rate:

> The ratio of the number of customers retained to the number of customers at risk.

The complement of retention is attrition or churn.

Understanding retention rates

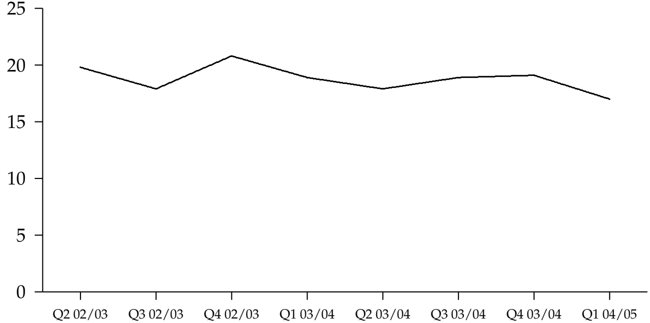
At the cohort level, we (almost) always observe increasing retention rates.

> *Renewal rates at regional magazines vary; generally 30% of subscribers renew at the end of their original subscription, but that figure jumps to 50% for second-time renewals and all the way to 75% for longtime readers.*
>
> Fielding, Michael (2005), "Get Circulation Going: DM Redesign Increases Renewal Rates for Magazines," *Marketing News*, September 1, 9–10.

> *New subscribers are actually more likely to cancel their subscriptions than older subscribers, and therefore, an increase in subscriber age helps overall reductions in churn.*
>
> Netflix (10-K for the fiscal year ended December 31, 2005)

Understanding retention rates

Vodafone Germany: Quarterly Annualized Churn Rate



Source: Vodafone Germany “Vodafone Analyst & Investor Day” presentation (2004-09-27)

Understanding retention rates

I am happy to report that 41% of new members who joined in 2011 renewed their membership in 2012, and that ION has an overall retention of 78%.

ION Newsletter, Winter 2011–2012.

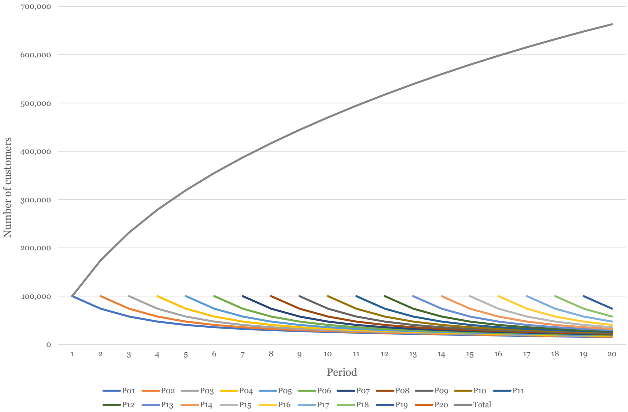
Cohort-level vs. aggregate numbers

Assume a firm acquires 100,000 customers each period, with each cohort being identical in its retention pattern. We observe this firm for 20 periods.

	Period																			
Cohort	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
P01	100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735	15,895	15,098
P02		100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735	15,895
P03			100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735
P04				100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622
P05					100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560
P06						100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554
P07							100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612
P08								100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745
P09									100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970
P10										100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312
P11											100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815
P12												100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545
P13													100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613
P14														100,000	74,000	57,700	47,255	40,356	35,617	32,203
P15															100,000	74,000	57,700	47,255	40,356	35,617
P16																100,000	74,000	57,700	47,255	40,356
P17																	100,000	74,000	57,700	47,255
P18																		100,000	74,000	57,700
P19																			100,000	74,000
P20																				100,000
Total	100,000	174,000	231,700	278,955	319,311	354,928	387,131	416,744	444,289	470,104	494,416	517,386	539,131	559,743	579,297	597,857	615,479	632,214	648,109	663,207

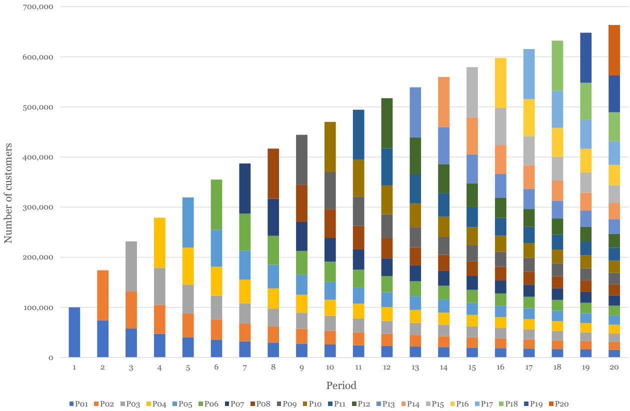
Cohort-level vs. aggregate numbers

Number of customers: Total and by cohort



Cohort-level vs. aggregate numbers

Size of customer base (by cohort)



Cohort-level vs. aggregate numbers

From number of customers ...

		Period																			
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Cohort	P01	100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735	15,895	15,098
	P02		100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735	15,895
	P03			100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622	16,735
	P04				100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560	17,622
	P05					100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554	18,560
	P06						100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612	19,554
	P07							100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745	20,612
	P08								100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970	21,745
	P09									100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312	22,970
	P10										100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815	24,312
	P11											100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545	25,815
	P12												100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613	27,545
	P13													100,000	74,000	57,700	47,255	40,356	35,617	32,203	29,613
	P14														100,000	74,000	57,700	47,255	40,356	35,617	32,203
	P15															100,000	74,000	57,700	47,255	40,356	35,617
	P16																100,000	74,000	57,700	47,255	40,356
	P17																	100,000	74,000	57,700	47,255
	P18																		100,000	74,000	57,700
	P19																			100,000	74,000
	P20																				100,000
	Total	100,000	174,000	231,700	278,955	319,311	354,928	387,131	416,744	444,289	470,104	494,416	517,386	539,131	559,743	579,297	597,857	615,479	632,214	648,109	663,207

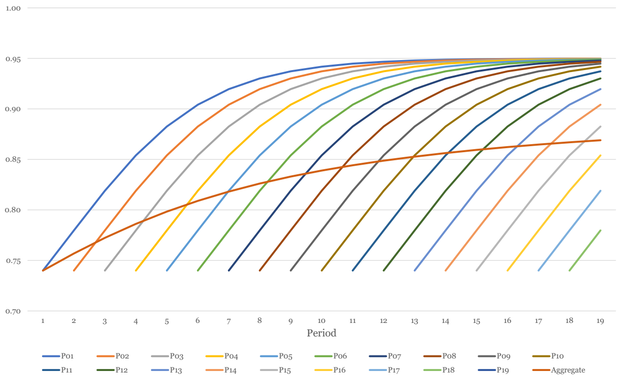
Cohort-level vs. aggregate numbers

... to retention rates

		Period																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Cohort	P01	0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949	0.949	0.949	0.950	0.950	0.950
	P02		0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949	0.949	0.949	0.950	0.950
	P03			0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949	0.949	0.949	0.950
	P04				0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949	0.949	0.949
	P05					0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949	0.949
	P06						0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948	0.949
	P07							0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947	0.948
	P08								0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945	0.947
	P09									0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942	0.945
	P10										0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937	0.942
	P11											0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930	0.937
	P12												0.740	0.780	0.819	0.854	0.883	0.904	0.920	0.930
	P13													0.740	0.780	0.819	0.854	0.883	0.904	0.920
	P14														0.740	0.780	0.819	0.854	0.883	0.904
	P15															0.740	0.780	0.819	0.854	0.883
	P16																0.740	0.780	0.819	0.854
	P17																	0.740	0.780	0.819
	P18																		0.740	0.780
	P19																			0.740
	Total	0.740	0.757	0.772	0.786	0.798	0.809	0.818	0.826	0.833	0.839	0.844	0.849	0.853	0.856	0.859	0.862	0.865	0.867	0.869

Cohort-level vs. aggregate numbers

Retention rates: Aggregate and by cohort



Why do cohort-level retention rates increase over time?

Individual-level time dynamics

- increasing loyalty as the customer gains more experience with the firm
- increasing switching costs with the passage of time

vs.

A sorting effect in a heterogeneous population

The role of heterogeneity

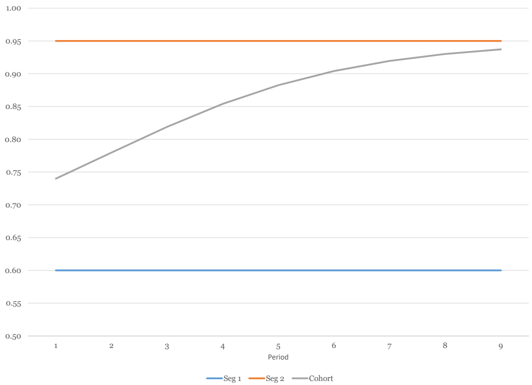
Suppose we acquire a cohort of 100,000 customers.

Each customer belongs to one of two underlying (and *unobservable*) segments:

- Segment 1 comprises 60% of the customers, each with a time-invariant retention probability of 0.60.
- Segment 2 comprises 40% of the customers, each with a time-invariant retention probability of 0.95.

The role of heterogeneity

	# Customers still alive			Retention rate		
Period	Seg 1	Seg 2	Total	Seg 1	Seg 2	Total
1	60,000	40,000	100,000			
2	36,000	38,000	74,000	0.600	0.950	0.740
3	21,600	36,100	57,700	0.600	0.950	0.780
4	12,960	34,295	47,255	0.600	0.950	0.819
5	7,776	32,580	40,356	0.600	0.950	0.854
6	4,666	30,951	35,617	0.600	0.950	0.883
7	2,799	29,404	32,203	0.600	0.950	0.904
8	1,680	27,933	29,613	0.600	0.950	0.920
9	1,008	26,537	27,545	0.600	0.950	0.930
10	605	25,210	25,815	0.600	0.950	0.937



Motivating problem

1000 customers are acquired at the beginning of Year 1 with the following pattern of renewals:

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
0006	1	0	0	0	0
⋮		⋮		⋮	
0998	1	0	0	0	0
0999	1	1	1	0	0
1000	1	0	0	0	0
	1000	631	468	382	326

Motivating problem

Year 1	Year 2	Year 3	Year 4	Year 5
1000	631	468	382	326

Assume:

- Each contract is annual, starting on January 1 and expiring at 11:59pm on December 31
- An average contribution margin of $100/year, which is "booked" at the beginning of the contract period
- A 10% discount rate

Motivating problem

Q1 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)?

Q2 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?

Expected value of a new customer

	Year 1	Year 2	Year 3	Year 4	Year 5
P(still a customer)	1.000	0.631	0.468	0.382	0.326
Contribution margin	\$100	\$100	\$100	\$100	\$100
Discount	1	$\frac{1}{1.1}$	$\frac{1}{(1.1)^2}$	$\frac{1}{(1.1)^3}$	$\frac{1}{(1.1)^4}$

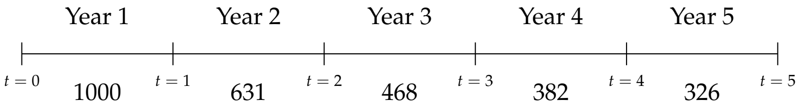
$$\$100 + \$100 \times \frac{0.631}{1.1} + \$100 \times \frac{0.468}{(1.1)^2} + \$100 \times \frac{0.382}{(1.1)^3} + \$100 \times \frac{0.326}{(1.1)^4} = \$247$$

Expected value of a new customer

Problem:

- We are ignoring any contribution margin the customer could possibly generate after Year 5.
- To get a true sense of (expected) customer lifetime value, we need to know the probability that the individual is still a customer in year 6, year 7, and so on.

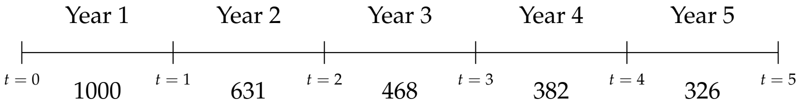
Notation and Terminology



The *survivor function* $S(t)$ is the proportion of the cohort that continue as a customer beyond t.

$$S(0) = ?$$
$$S(1) = ?$$
$$S(2) = ?$$

Notation and terminology



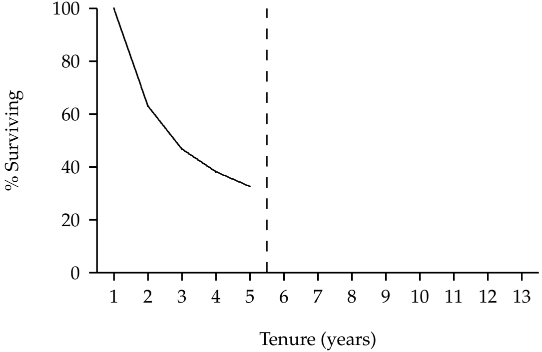
The *retention rate* is the ratio of customers retained to the number at risk.

$$r(1) = ?$$
$$r(2) = ?$$

For survivor function $S(t)$, $r(t) = \dfrac{S(t)}{S(t-1)}$.

Modelling objective

We want to derive a mathematical expression for $S(t)$, which can then be used to generate the desired forecasts.



Model development (I)

At the end of each contract period, a customer makes their renewal decision by tossing a coin: H $\rightarrow$ renew, T $\rightarrow$ don't renew

Length of relationship			
1 period	T		
2 periods	H	T	
3 periods	H	H	T
…			

$$P(t \text{ periods}) = \begin{cases} P(\mathsf{T}) & t = 1 \\ P(\mathsf{H}) \times P(t-1 \text{ periods}) & t = 2, 3, \ldots \end{cases}$$

Model development (I)

i) $P(\mathsf{H}) = 1 - \theta$ is constant and unobserved.

ii) All customers have the same "churn probability" θ.

	A	B	C	D	E	F
1	theta	0.200				
2						
3						
4	Year	t	# Cust.	# Lost	P("die")	S(t)
5	1	0	1000			1.0000
6	2	1	631	369	0.2000	0.8000
7	3	2	468	163	0.1600	0.6400
8	4	3	382	86	0.1280	0.5120
9	5	4	326	56	0.1024	0.4096
10						
11	E6	=B1				
12	E7	=E6*(1-$B$1)				
13	F6	=F5-E6				

Model development (I)

More formally:

- Let the random variable T denote the duration of the customer's relationship with the firm.
- We assume that the random variable T is distributed geometric with parameter θ:

$$\begin{aligned} P(T = t \mid \theta) &= \theta(1-\theta)^{t-1}, \quad t = 1, 2, 3, \ldots \\ S(t \mid \theta) &= P(T > t \mid \theta) \\ &= (1-\theta)^{t}, \quad t = 0, 1, 2, 3, \ldots \end{aligned}$$

Estimating model parameters

Assuming

i) the observed data were generated according to the "coin flipping" story of contract renewal, and

ii) we know $P(\mathsf{T}) = \theta$,

the probability of the observed pattern of renewals is:

$$\begin{aligned}
&[P(T=1\,|\,\theta)]^{369}[P(T=2\,|\,\theta)]^{163}[P(T=3\,|\,\theta)]^{86} \\
&\qquad \times [P(T=4\,|\,\theta)]^{56}[S(t\,|\,\theta)]^{326} \\
&= [\theta]^{369}[\theta(1-\theta)]^{163}[\theta(1-\theta)^2]^{86} \\
&\qquad \times [\theta(1-\theta)^3]^{56}[(1-\theta)^4]^{326}
\end{aligned}$$

Estimating model parameters

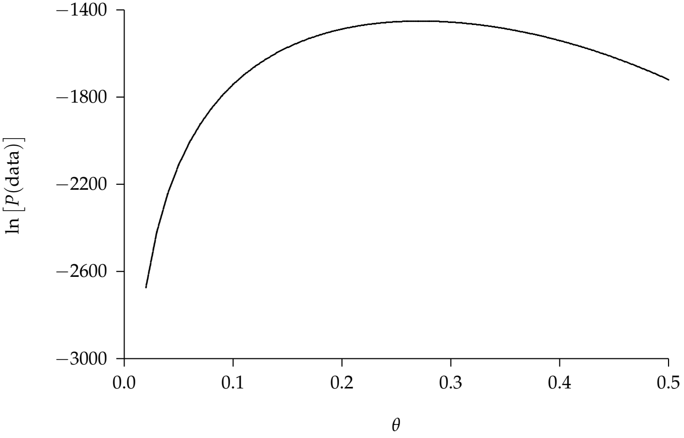
Suppose we have two candidate coins:

$$\begin{aligned} &\text{Coin A:} \quad \theta = 0.2 \\ &\text{Coin B:} \quad \theta = 0.5 \end{aligned}$$

Which coin is more likely to have generated the observed pattern of renewals across this set of 1000 customers?

θ	$P(\text{data} \mid \theta)$	$\ln\left[P(\text{data} \mid \theta)\right]$
0.2	6.00×10^{-647}	-1488.0
0.5	1.40×10^{-747}	-1719.7

Estimating model parameters



Estimating model parameters

We estimate the model parameters using the method of *maximum likelihood*:

– The likelihood function is defined as the probability of observing the data for a given set of the (unknown) model parameters.

– It is computed using the model and is viewed as a function of the model parameters:

$$L(\text{parameters} \,|\, \text{data}) = p(\text{data} \,|\, \text{parameters})\,.$$

– For a given dataset, the maximum likelihood estimates of the model parameters are those values that maximize $L(\cdot)$.

– It is typically more convenient to use the natural logarithm of the likelihood function — the log-likelihood function.

Estimating model parameters

The log-likelihood function is given by:

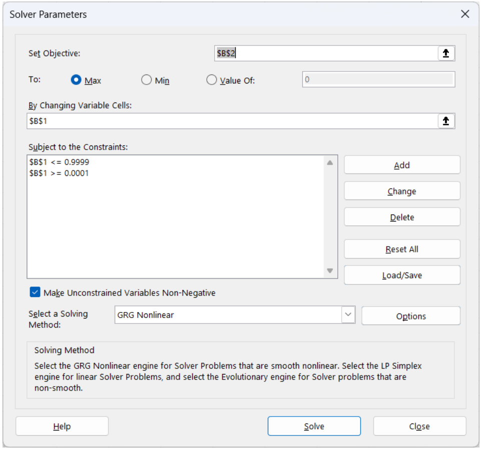
$$\begin{aligned} LL(\theta \mid \text{data}) = {} & 369 \times \ln[P(T = 1 \mid \theta)] + \\ & 163 \times \ln[P(T = 2 \mid \theta)] + \\ & 86 \times \ln[P(T = 3 \mid \theta)] + \\ & 56 \times \ln[P(T = 4 \mid \theta)] + \\ & 326 \times \ln[S(4 \mid \theta)] \end{aligned}$$

The maximum value of the log-likelihood function is $LL = -1451.2$, which occurs at $\hat{\theta} = 0.272$.

Estimating model parameters

	A	B	C	D	E	F	G
1	theta	0.200					
2	LL	-1487.98					
3							
4	Year	t	# Cust.	# Lost	P("die")	S(t)	
5	1	0	1000			1.0000	
6	2	1	631	369	0.2000	0.8000	-593.88
7	3	2	468	163	0.1600	0.6400	-298.71
8	4	3	382	86	0.1280	0.5120	-176.79
9	5	4	326	56	0.1024	0.4096	-127.62
10							-290.98
11	G6	=D6*LN(E6)					
12	G10	=C9*LN(F9)					
13	B2	=SUM(G6:G10)					

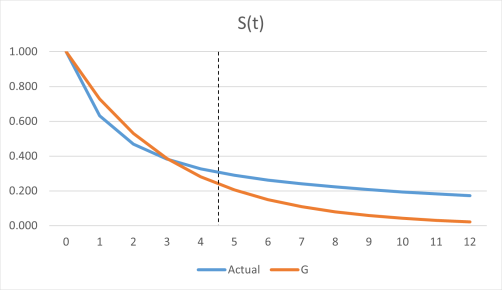
Estimating model parameters

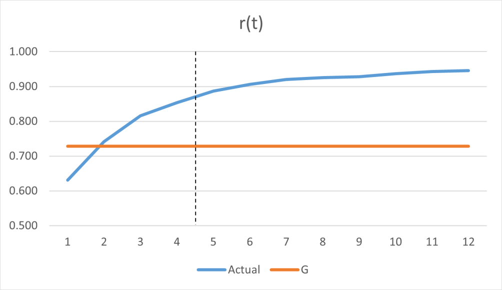


Model performance

	A	B	C	D	E	F	G
1	theta	0.272					
2	LL	-1451.16					
3							
4	Year	t	# Cust.	# Lost	P("die")	S(t)	
5	1	0	1000			1.0000	
6	2	1	631	369	0.2717	0.7283	-480.88
7	3	2	468	163	0.1979	0.5305	-264.09
8	4	3	382	86	0.1441	0.3864	-166.60
9	5	4	326	56	0.1050	0.2814	-126.23
10	6	5			0.0764	0.2050	-413.36
11	7	6			0.0557	0.1493	
12	8	7			0.0406	0.1087	
13	9	8			0.0295	0.0792	
14	10	9			0.0215	0.0577	
15	11	10			0.0157	0.0420	
16	12	11			0.0114	0.0306	
17	13	12			0.0083	0.0223	

Model performance





Model development (II)

Consider the following story of customer behavior:

i) At the end of each period, an individual renews his contract with (constant and unobserved) probability $1 - \theta$.

ii) "Churn probabilities" vary across customers.

Since we don't know any given customer's true value of θ, we treat it as a realization of a random variable (Θ).

We need to specify a probability distribution that captures how θ varies across customers (by giving us the probability of each possible value of θ).

Model development (II)

Segment	Θ	$P(\Theta = \theta_i)$
1	θ_1	π_1
2	θ_2	π_2

For a randomly chosen customer,

$$\begin{aligned} P(T = t) &= P(T = t \mid \Theta = \theta_1) P(\Theta = \theta_1) \\ &\qquad + P(T = t \mid \Theta = \theta_2) P(\Theta = \theta_2) \\ &= \theta_1 (1 - \theta_1)^{t-1} \pi_1 + \theta_2 (1 - \theta_2)^{t-1} \pi_2 \end{aligned}$$

$$\begin{aligned} S(t) &= S(t \mid \Theta = \theta_1) P(\Theta = \theta_1) + S(t \mid \Theta = \theta_2) P(\Theta = \theta_2) \\ &= (1 - \theta_1)^t \pi_1 + (1 - \theta_2)^t \pi_2 \end{aligned}$$

Estimating model parameters

Assuming

i) the observed data were generated according to the heterogeneous "coin flipping" story of contract renewal, and

ii) we know θ_1, θ_2 and π_1,

the probability of the observed pattern of renewals is:

$$[P(T=1 \mid \theta_1, \theta_2, \pi_1)]^{369}[P(T=2 \mid \theta_1, \theta_2, \pi_1)]^{163}[P(T=3 \mid \theta_1, \theta_2, \pi_1)]^{86} \\ \times [P(T=4 \mid \theta_1, \theta_2, \pi_1)]^{56}[S(4 \mid \theta_1, \theta_2, \pi_1)]^{326}$$

Estimating model parameters

The log-likelihood function is given by:

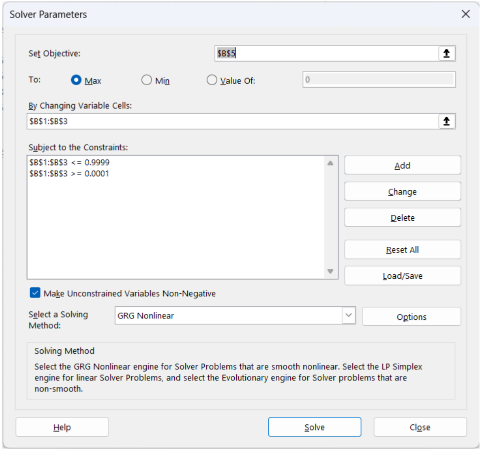
$$\begin{aligned}
LL(\theta_1, \theta_2, \pi_1 \,|\, \text{data}) = {} & 369 \times \ln[P(T = 1 \,|\, \theta_1, \theta_2, \pi_1)] + \\
& 163 \times \ln[P(T = 2 \,|\, \theta_1, \theta_2, \pi_1)] + \\
& 86 \times \ln[P(T = 3 \,|\, \theta_1, \theta_2, \pi_1)] + \\
& 56 \times \ln[P(T = 4 \,|\, \theta_1, \theta_2, \pi_1)] + \\
& 326 \times \ln[S(4 \,|\, \theta_1, \theta_2, \pi_1)]
\end{aligned}$$

The maximum value of the log-likelihood function is $LL = -1401.4$, which occurs at $\hat{\theta}_1 = 0.118$, $\hat{\theta}_2 = 0.648$ and $\hat{\pi}_1 = 0.526$.

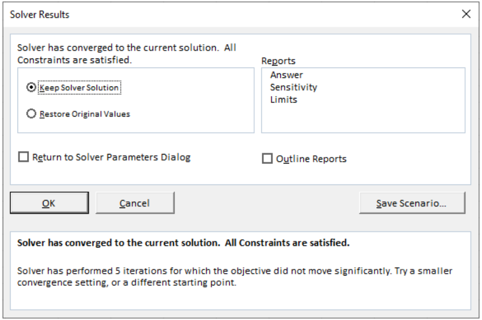
Estimating model parameters

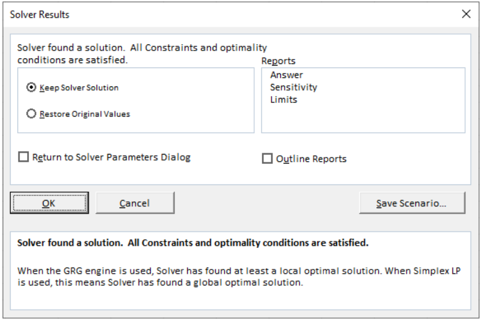
	A	B	C	D	E	F	G	H	I
1	theta_1	0.200							
2	theta_2	0.500							
3	pi_1	0.500							
4	pi_2	0.500							
5	LL	-1433.91							
6					P("die")				
7	Year	t	# Cust.	# Lost	seg 1	seg 2	total	S(t)	
8	1	0	1000					1.0000	
9	2	1	631	369	0.2000	0.5000	0.3500	0.6500	-387.38
10	3	2	468	163	0.1600	0.2500	0.2050	0.4450	-258.31
11	4	3	382	86	0.1280	0.1250	0.1265	0.3185	-177.81
12	5	4	326	56	0.1024	0.0625	0.0825	0.2361	-139.75
13									-470.65
14	B4	=1-B3							
15	G9	=E9*$B$3+F9*$B$4							

Estimating model parameters



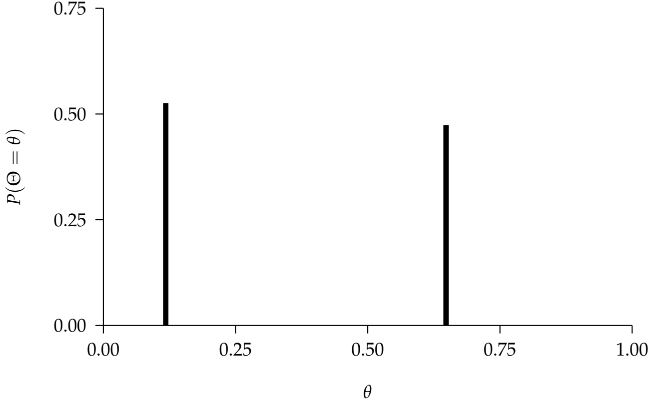
A Solver aside



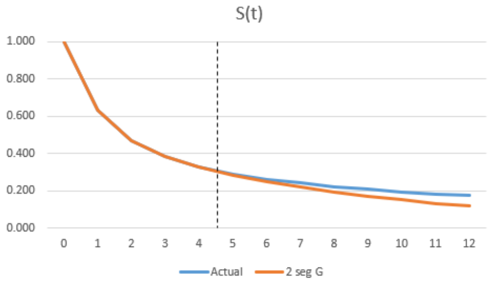


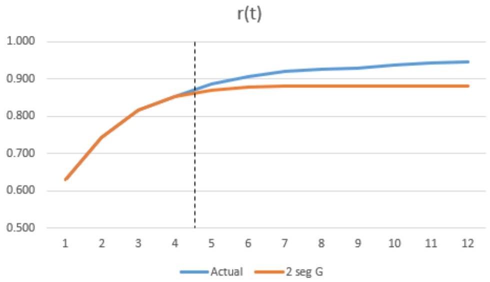
https://www.solver.com/standard-excel-solver-grg-nonlinear-solver-stopping-conditions

Implied distribution of Θ



Model performance

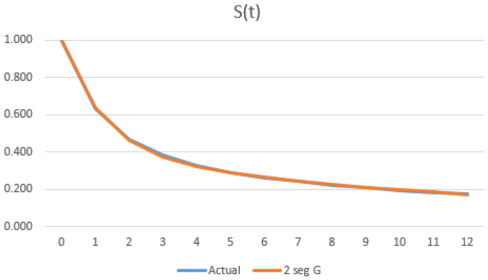


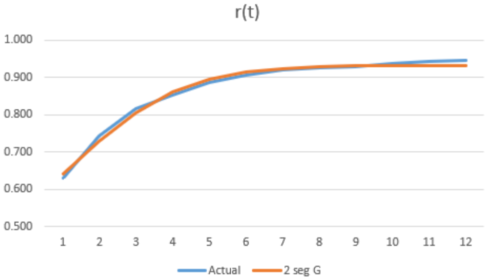


What if we use all the data?

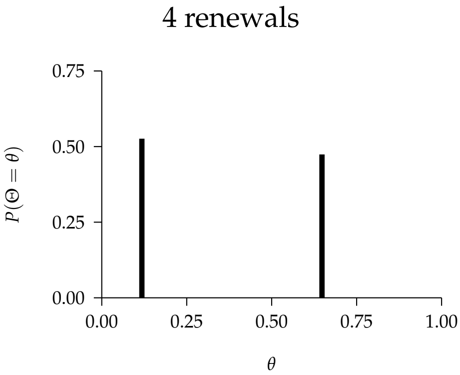
	A	B	C	D	E	F	G	H	I
1	theta_1	0.067							
2	theta_2	0.551							
3	pi_1	0.395							
4	pi_2	0.605							
5	LL	-1932.32							
6						P("die")			
7	Year	t	# Cust.	# Lost	seg 1	seg 2	total	S(t)	
8	1	0	1000					1.0000	
9	2	1	631	369	0.0672	0.5512	0.3599	0.6401	-377.07
10	3	2	468	163	0.0627	0.2474	0.1744	0.4657	-284.69
11	4	3	382	86	0.0584	0.1110	0.0902	0.3755	-206.85
12	5	4	326	56	0.0545	0.0498	0.0517	0.3238	-165.91
13	6	5	289	37	0.0509	0.0224	0.0336	0.2902	-125.53
14	7	6	262	27	0.0474	0.0100	0.0248	0.2653	-99.80
15	8	7	241	21	0.0443	0.0045	0.0202	0.2451	-81.93
16	9	8	223	18	0.0413	0.0020	0.0175	0.2276	-72.78
17	10	9	207	16	0.0385	0.0009	0.0158	0.2118	-66.40
18	11	10	194	13	0.0359	0.0004	0.0144	0.1974	-55.09
19	12	11	183	11	0.0335	0.0002	0.0134	0.1840	-47.47
20	13	12	173	10	0.0313	0.0001	0.0124	0.1716	-43.90
21									-304.90

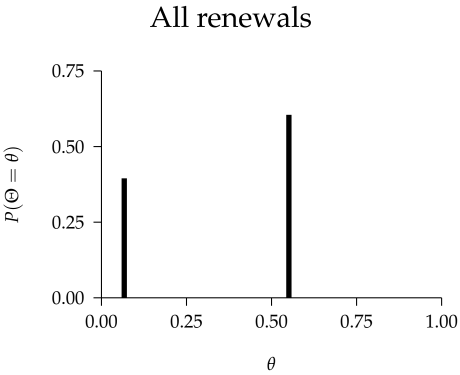
What if we use all the data?





Implied distributions of Θ



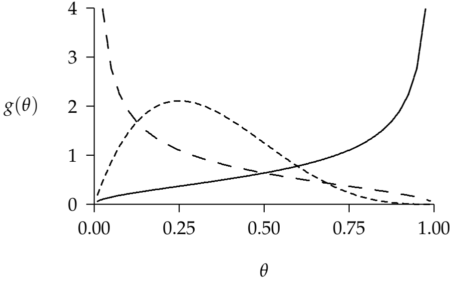


Our inferences about the distribution of Θ are sensitive to the length of the calibration period $\longrightarrow$ not a "good" model

Accommodating heterogeneity in θ

Rather than trying three (or more) segments, we move from a *discrete* distribution for Θ to a *continuous* distribution (an infinite number of segments).

The distribution of churn probabilities is captured by a continuous function $g(\theta)$ defined over $(0, 1)$:



Accommodating heterogeneity in θ

Just as the heights of the bars in our discrete distribution sum to one, so the area under the curve equals one.

For a randomly chosen customer, we go from

$$P(T = t) = \sum_{i=1}^{I} P(T = t \mid \theta_i)\pi_i$$

to

$$P(T = t) = \int_0^1 P(T = t \mid \theta)g(\theta)\, d\theta$$

The beta distribution

The beta distribution is a flexible (and mathematically convenient) two-parameter distribution bounded between 0 and 1:

$$g(\theta \mid \gamma, \delta) = \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma,\delta)},$$

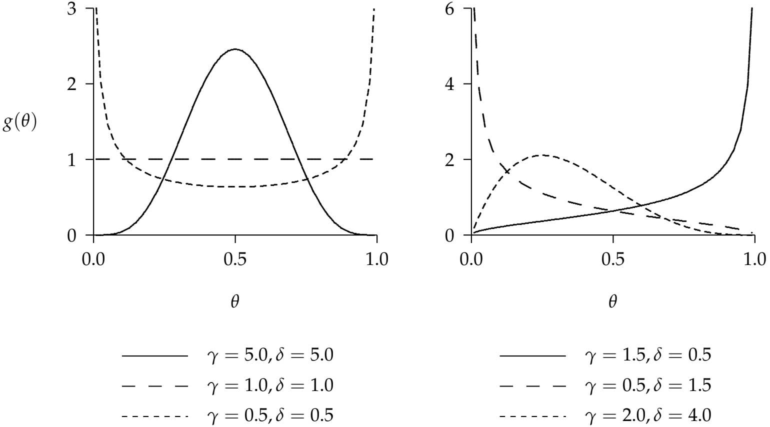
where $\gamma, \delta > 0$ and $B(\gamma, \delta)$ is the beta function.

The mean and variance of the beta distribution are

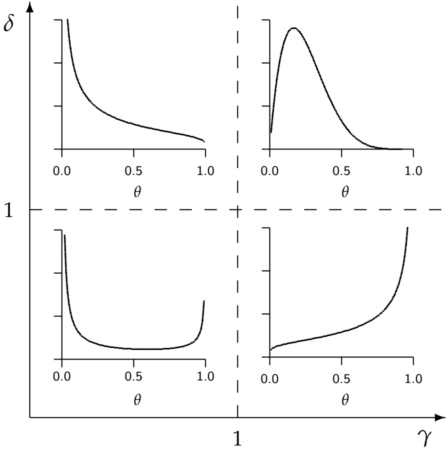
$$E(\Theta) = \frac{\gamma}{\gamma+\delta}$$

$$\text{var}(\Theta) = \frac{\gamma\delta}{(\gamma+\delta)^2(\gamma+\delta+1)}$$

Illustrative beta distributions



Five general shapes of the beta distribution



The beta function

The beta function $B(\gamma, \delta)$ is defined by the integral

$$B(\gamma, \delta) = \int_0^1 t^{\gamma-1}(1-t)^{\delta-1} dt, \;\; \gamma, \delta > 0\,,$$

and can be expressed in terms of gamma functions:

$$B(\gamma, \delta) = \frac{\Gamma(\gamma)\Gamma(\delta)}{\Gamma(\gamma+\delta)}\,.$$

The gamma function $\Gamma(\gamma)$ is a generalized factorial, which has the recursive property $\Gamma(\gamma+1) = \gamma\Gamma(\gamma)$.

Since $\Gamma(0) = 1$, $\Gamma(n) = (n-1)!$ for positive integer n.

Model development (III)

Consider the following story of customer behavior:

i) At the end of each period, an individual renews their contract with (constant and unobserved) probability $1 - \theta$.

   $\longrightarrow$ the duration of the customer's relationship with the firm is characterized by the geometric distribution.

ii) Heterogeneity in θ is characterized by the beta distribution.

Model development (III)

For a randomly chosen individual,

$$\begin{aligned}
P(T = t \mid \gamma, \delta) &= \int_0^1 P(T = t \mid \theta) g(\theta \mid \gamma, \delta) \, d\theta \\
&= \frac{B(\gamma + 1, \delta + t - 1)}{B(\gamma, \delta)} \\
S(t \mid \gamma, \delta) &= \int_0^1 S(t \mid \theta) g(\theta \mid \gamma, \delta) \, d\theta \\
&= \frac{B(\gamma, \delta + t)}{B(\gamma, \delta)}
\end{aligned}$$

We call this *continuous mixture* model the beta-geometric (BG) distribution.

Model development (III)

We can compute BG probabilities using the following forward-recursion formula from $P(T=1)$:

$$P(T=t \mid \gamma, \delta) = \begin{cases} \dfrac{\gamma}{\gamma+\delta} & t=1 \\[2ex] \dfrac{\delta+t-2}{\gamma+\delta+t-1} \times P(T=t-1) & t=2,3,\ldots \end{cases}$$

Estimating model parameters

Assuming

i) the observed data were generated according to the heterogeneous "coin flipping" story of contract renewal, and

ii) we know γ and δ,

the probability of the observed pattern of renewals is:

$$[P(T=1\,|\,\gamma,\delta)]^{369}[P(T=2\,|\,\gamma,\delta)]^{163}[P(T=3\,|\,\gamma,\delta)]^{86} \\ \times [P(T=4\,|\,\gamma,\delta)]^{56}[S(4\,|\,\gamma,\delta)]^{326}$$

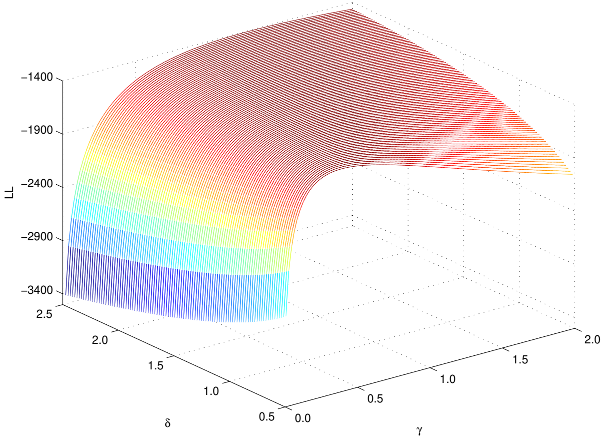
Estimating model parameters

The log-likelihood function is given by:

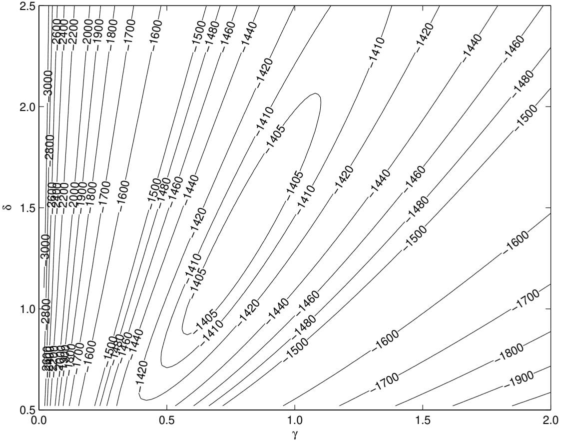
$$\begin{aligned} LL(\gamma, \delta \,|\, \text{data}) = {} & 369 \times \ln[P(T = 1 \,|\, \gamma, \delta)] + \\ & 163 \times \ln[P(T = 2 \,|\, \gamma, \delta)] + \\ & 86 \times \ln[P(T = 3 \,|\, \gamma, \delta)] + \\ & 56 \times \ln[P(T = 4 \,|\, \gamma, \delta)] + \\ & 326 \times \ln[S(4 \,|\, \gamma, \delta)] \end{aligned}$$

The maximum value of the log-likelihood function is $LL = -1401.6$, which occurs at $\hat{\gamma} = 0.764$ and $\hat{\delta} = 1.296$.

Surface plot of BG log-likelihood function



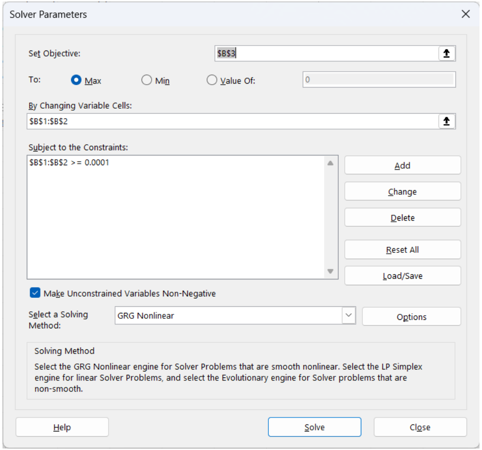
Contour plot of BG log-likelihood function



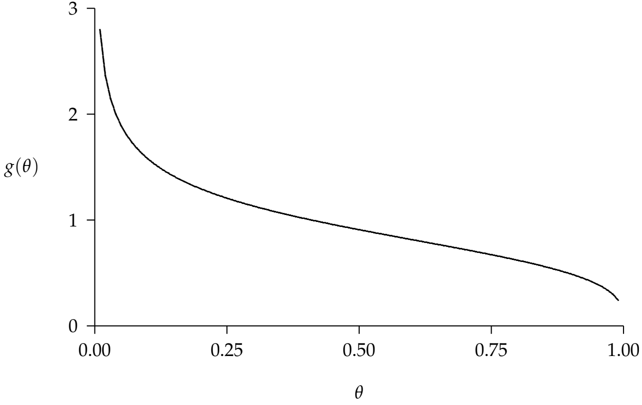
Estimating model parameters

	A	B	C	D	E	F	G
1	gamma	1.000					
2	delta	1.000					
3	LL	-1453.97					
4							
5	Year	t	# Cust.	# Lost	P("die")	S(t)	
6	1	0	1000			1.0000	
7	2	1	631	369	0.5000	0.5000	-255.77
8	3	2	468	163	0.1667	0.3333	-292.06
9	4	3	382	86	0.0833	0.2500	-213.70
10	5	4	326	56	0.0500	0.2000	-167.76
11							-524.68
12	E7	=B1/(B1+B2)					
13	E8	=($B$2+B8-2)/($B$1+$B$2+B8-1)*E7					

Estimating model parameters

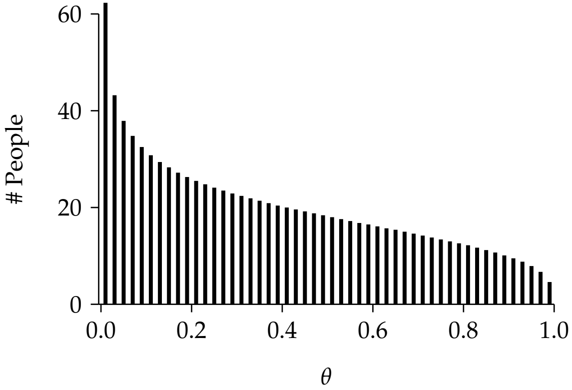


Estimated distribution of churn probabilities



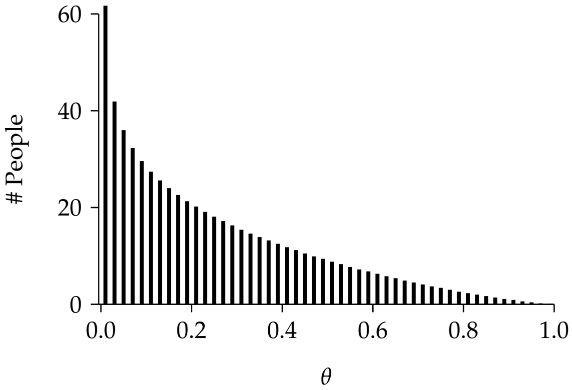
$$\hat{\gamma} = 0.764, \hat{\delta} = 1.296, \widehat{E(\Theta)} = 0.371$$

Year 1



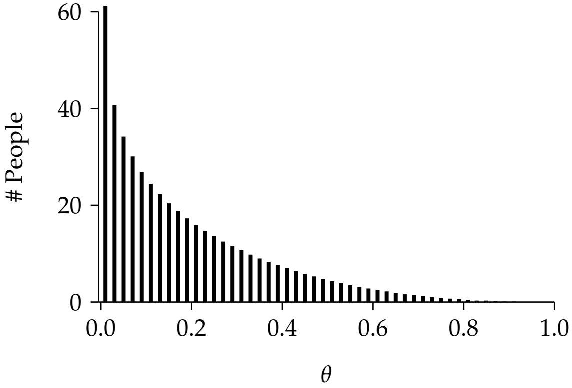
$E(\Theta) = 0.371 \longrightarrow$ expect $1000 \times (1 - 0.371)$ = 629 customers to renew at the end of Year 1.

Year 2



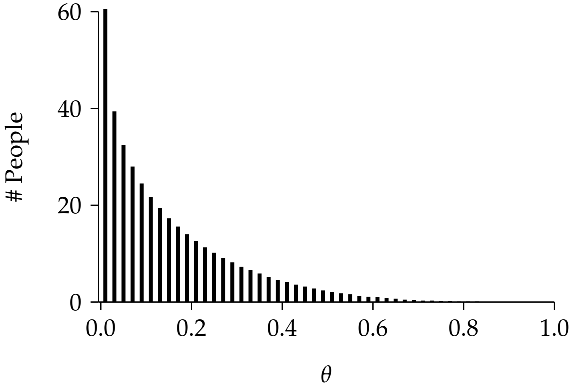
$E(\Theta) = 0.250 \longrightarrow$ expect $629 \times (1 - 0.250)$ = 472 customers to renew at the end of Year 2.

Year 3



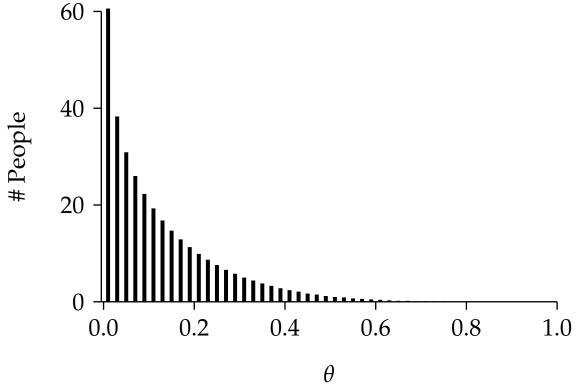
$E(\Theta) = 0.188 \longrightarrow$ expect $472 \times (1 - 0.188)$ = 383 customers to renew at the end of Year 3.

Year 4



$E(\Theta) = 0.151 \longrightarrow$ expect $383 \times (1 - 0.151)$ = 325 customers to renew at the end of Year 4.

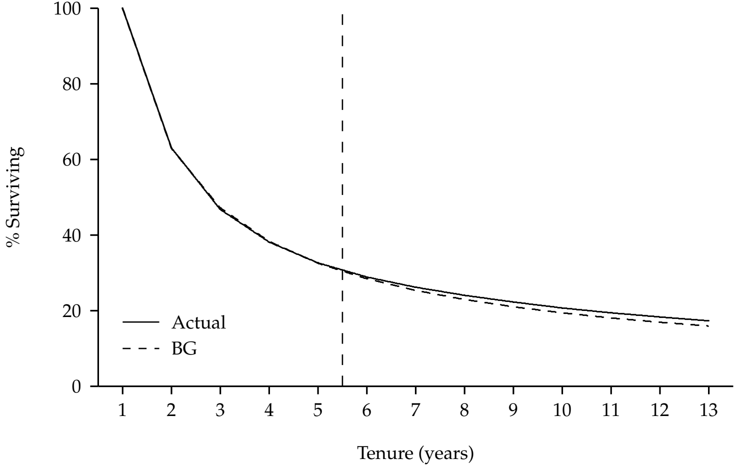
Year 5



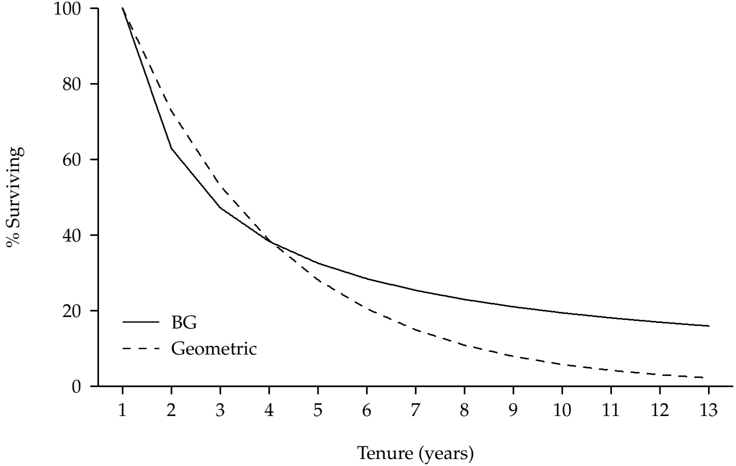
Model performance

	A	B	C	D	E	F	G
1	gamma	0.764					
2	delta	1.296					
3	LL	-1401.56					
4							
5	Year	t	# Cust.	# Lost	P("die")	S(t)	
6	1	0	1000			1.0000	
7	2	1	631	369	0.3708	0.6292	-366.08
8	3	2	468	163	0.1571	0.4721	-301.74
9	4	3	382	86	0.0888	0.3833	-208.22
10	5	4	326	56	0.0579	0.3255	-159.59
11	6	5			0.0410	0.2845	-365.93
12	7	6			0.0308	0.2537	
13	8	7			0.0240	0.2296	
14	9	8			0.0194	0.2103	
15	10	9			0.0160	0.1943	
16	11	10			0.0134	0.1809	
17	12	11			0.0115	0.1694	
18	13	12			0.0099	0.1595	

Survival curve projection



Comparing survival curves



Implied retention rates

Recall that

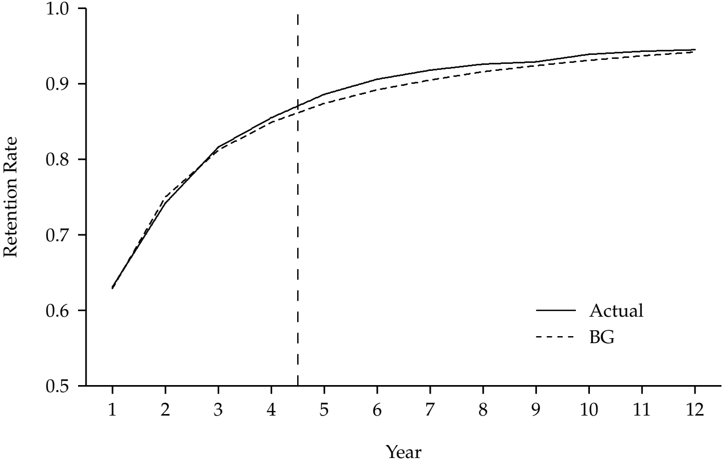
$$r(t) = \frac{S(t)}{S(t-1)}\,,\ t = 1, 2, 3, \ldots$$

Given the expression for the BG survivor function,

$$r(t \mid \gamma, \delta) = \frac{\delta + t - 1}{\gamma + \delta + t - 1}\,.$$

An increasing function of time, even though the individual-level retention probability is constant. (A sorting effect in a heterogeneous population.)

Projecting retention rates



An alternative recursion

The relationship between $r(t)$ and $S(t)$ implies that, given knowledge of $r(t)$, we can compute $S(t)$ using the *forward recursion*:

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

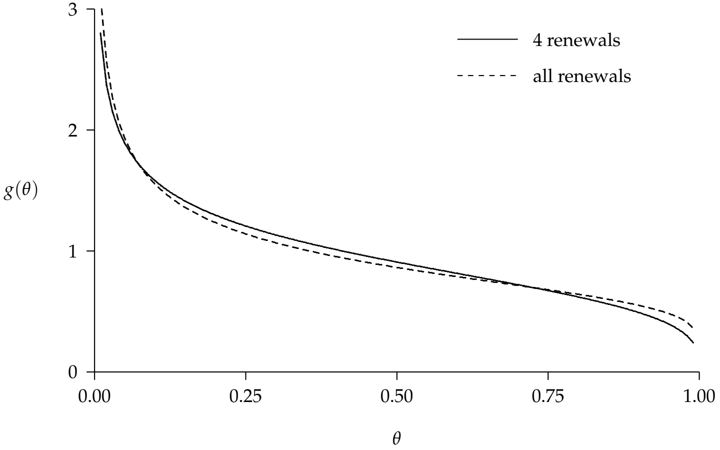
We can compute the BG survivor function using the following forward recursion formula from $S(0)$:

$$S(t \mid \gamma, \delta) = \begin{cases} 1 & \text{if } t = 0 \\ \dfrac{\delta + t - 1}{\gamma + \delta + t - 1} \times S(t-1) & \text{if } t = 1, 2, 3, \ldots \end{cases}$$

What if we use all the data?

	A	B	C	D	E	F	G
1	gamma	0.697	0.764	mean	0.374	0.371	
2	delta	1.169	1.296	var	0.082	0.076	
3	LL	-1931.49					
4							
5	Year	t	# Cust.	# Lost	P("die")	S(t)	
6	1	0	1000			1.0000	
7	2	1	631	369	0.3737	0.6263	-363.24
8	3	2	468	163	0.1524	0.4739	-306.66
9	4	3	382	86	0.0855	0.3885	-211.51
10	5	4	326	56	0.0557	0.3328	-161.75
11	6	5	289	37	0.0396	0.2932	-119.51
12	7	6	262	27	0.0298	0.2634	-94.87
13	8	7	241	21	0.0234	0.2401	-78.90
14	9	8	223	18	0.0189	0.2212	-71.45
15	10	9	207	16	0.0156	0.2056	-66.53
16	11	10	194	13	0.0132	0.1924	-56.26
17	12	11	183	11	0.0113	0.1811	-49.31
18	13	12	173	10	0.0098	0.1713	-46.24
19							-305.27

What if we use all the data?



Concepts and tools introduced

- Modelling single-event discrete-time duration data.
- Capturing unobserved heterogeneity using discrete (finite) and continuous distributions.
- The beta-geometric (BG) distribution as a robust model of contract renewal behavior.
- The method of maximum likelihood as a means of estimating model parameters.
- Using the BG model to forecast survival and retention.

Further reading

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 (2007), "How Not to Project Customer Retention." `http://brucehardie.com/notes/016/`

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

Lee, Ka Lok, Peter S. Fader, and Bruce G.S. Hardie (2007), "How to Project Patient Persistency," *FORESIGHT*, Issue 8, Fall, 31–35.

Potter, R.G. and M. P. Parker (1964), "Predicting the Time Required to Conceive," *Population Studies*, **18** (July), 99–116.

PMLR Proceedings of Machine Learning Research

Volume 146: Survival Prediction - Algorithms, Challenges and Applications, 22-24 March 2021, Stanford University, Palo Alto (CA), USA

Beta Survival Models

David Hubbard DHUBBARD@NETFLIX.COM
Benoit Rostykus BROSTYKUS@NETFLIX.COM
Yves Raimond YRAIMOND@NETFLIX.COM

Tony Jebara TONYJ@SPOTIFY.COM

"This model [...] was also studied by Fader and Hardie (2007). We find that in practice this model fits the discrete decision data quite well, and that it allows for accurate projections of future decision points."

Break for lab (part 1)

Back to the motivating problem

Year 1	Year 2	Year 3	Year 4	Year 5
1000	631	468	382	326

Q1 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)?

Q2 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?

Expected value of a new customer

	Year 1	Year 2	Year 3	Year 4	Year 5
	$t = 0$ – $t = 1$	$t = 1$ – $t = 2$	$t = 2$ – $t = 3$	$t = 3$ – $t = 4$	$t = 4$ – $t = 5$
P(still a customer)	1.000	0.631	0.468	0.382	0.326
Contribution margin	\$100	\$100	\$100	\$100	\$100
Discount	1	$\frac{1}{1.1}$	$\frac{1}{(1.1)^2}$	$\frac{1}{(1.1)^3}$	$\frac{1}{(1.1)^4}$

$$\$100 + \$100 \times \frac{0.631}{1.1} + \$100 \times \frac{0.468}{(1.1)^2}$$
$$+ \$100 \times \frac{0.382}{(1.1)^3} + \$100 \times \frac{0.326}{(1.1)^4} = \$247$$

Expected value of a new customer

We can write our calculation of the expected value of a new customer (over a five-year period) as

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

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

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

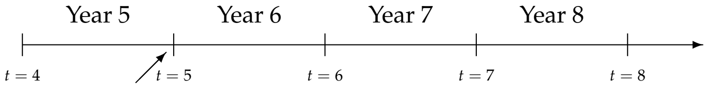
We use the BG model to project $S(t)$ out to infinity.

Expected value of a new customer

	A	B	C	D	E	F
1	gamma	0.764			E(CLV)	$362.11
2	delta	1.296				
3	d	0.1				
4	CM	$100				
5						
6	Year	t	P("die")	S(t)		disc.
7	1	0		1.0000		1.0000
8	2	1	0.3708	0.6292		0.9091
9	3	2	0.1571	0.4721		0.8264
10	4	3	0.0888	0.3833		0.7513
205	199	198	0.0001	0.0201		0.0000
206	200	199	0.0001	0.0200		0.0000
207						
208	F7	=1/(1+$B$3)^B7				
209	F1	=B4*SUMPRODUCT(D7:D206,F7:F206)				

Expected residual value of the cohort

Standing at the end of Year 5, what is the expected residual lifetime value of a customer?



$P(\text{still a customer})$	$S(5 \mid T > 4)$	$S(6 \mid T > 4)$	$S(7 \mid T > 4)$	$\cdots$
Contribution margin	\$100	\$100	\$100	$\cdots$
Discount	1	$\frac{1}{1.1}$	$\frac{1}{(1.1)^2}$	$\cdots$

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

Expected residual value of the cohort

By definition, for $t = 4, 5, 6, \ldots,$

$$\begin{aligned} S(t \mid T > 4) &= P(T > t \mid T > 4) \\ &= \frac{P(T > t \cap T > 4)}{P(T > 4)} \\ &= \frac{P(T > t)}{P(T > 4)} \\ &= \frac{S(t)}{S(4)} \end{aligned}$$

Expected residual value of the cohort

	A	B	C	D	E	F	G
1	gamma	0.764				E(RLV)	$568.38
2	delta	1.296					
3	d	0.1					
4	CM	$100					
5						Given 4 renewals	
6	Year	t	P("die")	S(t)		S(t|T>4)	disc.
7	1	0		1.0000			
8	2	1	0.3708	0.6292			
9	3	2	0.1571	0.4721			
10	4	3	0.0888	0.3833			
11	5	4	0.0579	0.3255			
12	6	5	0.0410	0.2845		0.8740	1.0000
13	7	6	0.0308	0.2537		0.7794	0.9091
205	199	198	0.0001	0.0201		0.0616	0.0000
206	200	199	0.0001	0.0200		0.0614	0.0000
207							
208	F12	=D12/$D$11					
209	G12	=1/(1+$B$3)^(B12-5)					

$\Longrightarrow$ expected residual value of the group of customers at the end of Year 5 is $326 \times \$568 = \$185{,}168$.

$E(RLV)$ as a function of customer tenure

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	gamma	0.764				E(RLV)	$288.33	E(RLV)	$394.07	E(RLV)	$467.66	E(RLV)	$523.63	E(RLV)	$568.38
2	delta	1.296													
3	d	0.1													
4	CM	$100													
5						0 renewals		1 renewal		2 renewals		3 renewals		4 renewals	
6	Year	t	P("die")	S(t)		S(t|T>0)	disc.	S(t|T>1)	disc.	S(t|T>2)	disc.	S(t|T>3)	disc.	S(t|T>4)	disc.
7	1	0		1.0000											
8	2	1	0.3708	0.6292		0.6292	1.0000								
9	3	2	0.1571	0.4721		0.4721	0.9091	0.7504	1.0000						
10	4	3	0.0888	0.3833		0.3833	0.8264	0.6092	0.9091	0.8119	1.0000				
11	5	4	0.0579	0.3255		0.3255	0.7513	0.5173	0.8264	0.6893	0.9091	0.8491	1.0000		
12	6	5	0.0410	0.2845		0.2845	0.6830	0.4521	0.7513	0.6025	0.8264	0.7421	0.9091	0.8740	1.0000
13	7	6	0.0308	0.2537		0.2537	0.6209	0.4032	0.6830	0.5373	0.7513	0.6618	0.8264	0.7794	0.9091
205	199	198	0.0001	0.0201		0.0201	0.0000	0.0319	0.0000	0.0425	0.0000	0.0523	0.0000	0.0616	0.0000
206	200	199	0.0001	0.0200		0.0200	0.0000	0.0318	0.0000	0.0423	0.0000	0.0521	0.0000	0.0614	0.0000

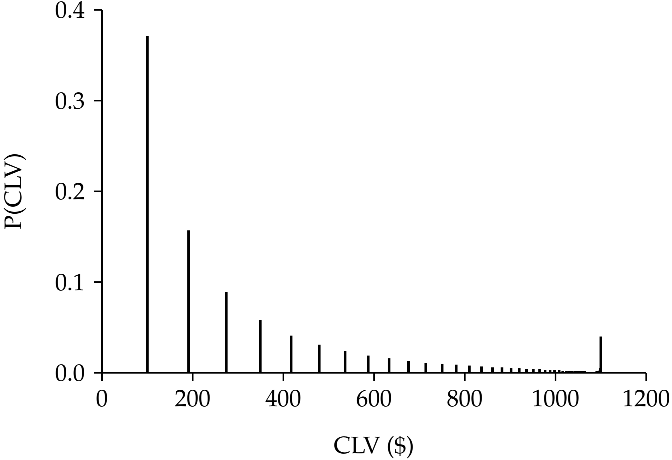
Reflecting on the value of a new customer

What is the *realized* value of a customer as a function of their tenure?

$$
\begin{aligned}
&\$100 \\
&\$100 + \frac{\$100}{1.1} = \$190.91 \\
&\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} = \$273.55 \\
&\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} + \frac{\$100}{(1.1)^3} = \$348.69 \\
&\$100 + \frac{\$100}{1.1} + \frac{\$100}{(1.1)^2} + \frac{\$100}{(1.1)^3} + \frac{\$100}{(1.1)^4} = \$416.99
\end{aligned}
$$

No one is worth $E(CLV)$! $\longrightarrow$ distribution of CLV (Fader and Hardie 2017).

Distribution of CLV



Break for lab (part 2)

Generalizing Q1

	Year 1	Year 2		Year n	
	$t=0$ to $t=1$	$t=1$ to $t=2$	$\cdots$	$t=n-1$ to $t=n$	
P(still a customer)	$S(0)$	$S(1)$		$S(n-1)$	$\cdots$
Contribution margin	$v(0)$	$v(1)$		$v(n-1)$	$\cdots$
Discount	1	$\frac{1}{(1+d)}$		$\frac{1}{(1+d)^{n-1}}$	$\cdots$

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

$$\text{If we assume } v(t) = \bar{v}\ \forall t, E(CLV) = \bar{v} \underbrace{\sum_{t=0}^{\infty} \frac{S(t)}{(1+d)^t}}_{\text{expected discounted lifetime, } E(DL)}$$

Generalizing Q2

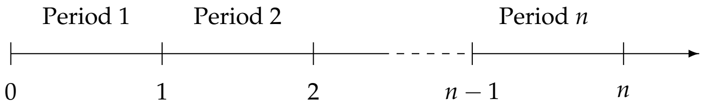
$t = n$ $\quad t = n+1$ $\quad t = n+2$

Year n	Year $n+1$	Year $n+2$	
P(still a customer)	$S(n \mid T > n-1)$	$S(n+1 \mid T > n-1)$	$\cdots$
Contribution margin	$v(n)$	$v(n+1)$	$\cdots$
Discount	1	$\dfrac{1}{(1+d)}$	$\cdots$

$$E(RLV \mid \text{active in } n) = \sum_{t=n}^{\infty} \frac{v(t)\, S(t \mid T > n-1)}{(1+d)^{t-n}}$$

$$\text{If } v(t) = \bar{v}\ \forall t,\ E(RLV \mid \text{active in } n) = \bar{v} \underbrace{\sum_{t=n}^{\infty} \frac{S(t \mid T > n-1)}{(1+d)^{t-n}}}_{\text{expected discounted residual lifetime, } E(DRL)}$$

Computing $E(DL)$



Standing at time 0 (i.e., before the customer is acquired),

$$E[DL(d) \mid \theta] = \sum_{t=0}^{\infty} \frac{S(t \mid \theta)}{(1+d)^t}$$
$$= \frac{1+d}{d+\theta}\,.$$

But θ is unobserved ...

Computing $E(DL)$

Integrating over the distribution of θ:

$$\begin{aligned}
E[DL(d) \mid \gamma, \delta] &= \int_0^1 E[DL(d) \mid \theta] g(\theta \mid \gamma, \delta)\, d\theta \\
&= \int_0^1 \left(\frac{1+d}{d+\theta} \right) \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma, \delta)}\, d\theta \\
&= \frac{1+d}{B(\gamma, \delta)} \int_0^1 \theta^{\gamma-1}(1-\theta)^{\delta-1}(d+\theta)^{-1}\, d\theta
\end{aligned}$$

Mathematical Diversion

The Gaussian hypergeometric function is the series

$$ {}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)} \frac{z^j}{j!} \, . $$

The series converges for $|z| < 1$ and is divergent for $|z| > 1$. If $|z| = 1$, the series converges for $c - a - b > 0$.

Euler's integral representation of the function is

$$ {}_2F_1(a,b;c;z) = \frac{1}{B(b, c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-zt)^{-a} dt \, , $$

where $c > b > 0$.

Computing $E(DL)$

Letting $s = 1 - \theta$,

$$\begin{aligned}\int_0^1 \theta^{\gamma-1}(1-\theta)^{\delta-1}(d+\theta)^{-1}\,d\theta &= \int_0^1 s^{\delta-1}(1-s)^{\gamma-1}(1+d-s)^{-1}ds \\ &= \frac{1}{1+d}\int_0^1 s^{\delta-1}(1-s)^{\gamma-1}(1-\tfrac{1}{1+d}s)^{-1}ds\end{aligned}$$

which, letting $a = 1$, $b = \delta$ and $c = \gamma + \delta$,

$$= \frac{B(\gamma,\delta)}{1+d}\,{}_2F_1\big(1,\delta;\gamma+\delta;\tfrac{1}{1+d}\big)$$

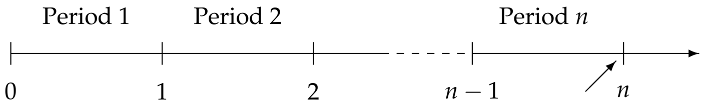
Therefore, $E[DL(d) \mid \gamma,\delta] = {}_2F_1\big(1,\delta;\gamma+\delta;\tfrac{1}{1+d}\big)$.

An alternative derivation

Assuming lifetimes are distributed BG,

$$
\begin{aligned}
E[DL(d) \mid \gamma, \delta] &= \sum_{t=0}^{\infty} \frac{S(t \mid \gamma, \delta)}{(1+d)^t} \\
&= \sum_{t=0}^{\infty} \frac{B(\gamma, \delta + t)}{B(\gamma, \delta)} \left(\frac{1}{1+d} \right)^t \\
&= \frac{\Gamma(\gamma + \delta)}{\Gamma(\delta)} \sum_{t=0}^{\infty} \frac{\Gamma(\delta + t)}{\Gamma(\gamma + \delta + t)} \left(\frac{1}{1+d} \right)^t \\
&= \frac{\Gamma(\gamma + \delta)}{\Gamma(\delta)} \sum_{t=0}^{\infty} \frac{\Gamma(t+1)\Gamma(\delta + t)}{\Gamma(\gamma + \delta + t)} \frac{1}{t!} \left(\frac{1}{1+d} \right)^t \\
&= {}_2F_1\left(1, \delta; \gamma + \delta; \tfrac{1}{1+d}\right).
\end{aligned}
$$

Computing $E(DRL)$



Standing at the end of period n, just prior to the point in time at which the customer makes their contract renewal decision,

$$E[DRL(d) \mid \theta, \text{ active in } n] = \sum_{t=n}^{\infty} \frac{S(t \mid T > n-1; \theta)}{(1+d)^{t-n}}$$
$$= \frac{(1-\theta)(1+d)}{d+\theta}.$$

But θ is unobserved ....

Bayes' theorem

The *prior distribution* $g(\theta)$ captures the possible values θ can take on, prior to collecting any information about the specific individual.

The *posterior distribution* $g(\theta \mid \text{data})$ is the conditional distribution of θ, given the observed data. It represents our updated opinion about the possible values θ can take on, now that we have some information about the specific individual.

According to Bayes' theorem:

$$g(\theta \mid \text{data}) = \frac{f(\text{data} \mid \theta) g(\theta)}{\int f(\text{data} \mid \theta) g(\theta)\, d\theta}$$

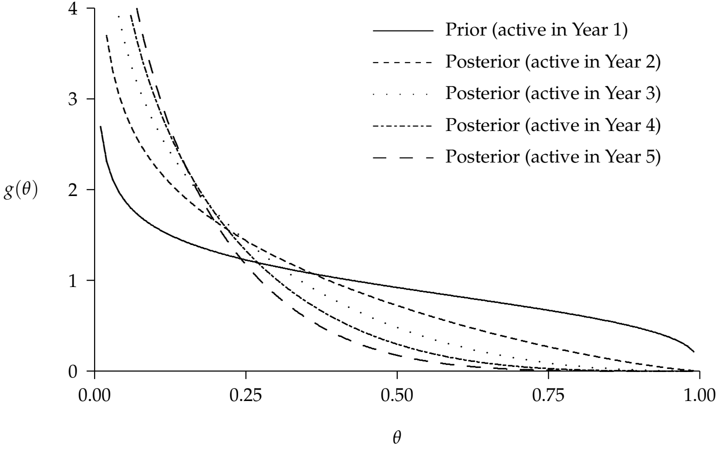
Computing $E(DRL)$

By Bayes' Theorem, the posterior distribution of θ is

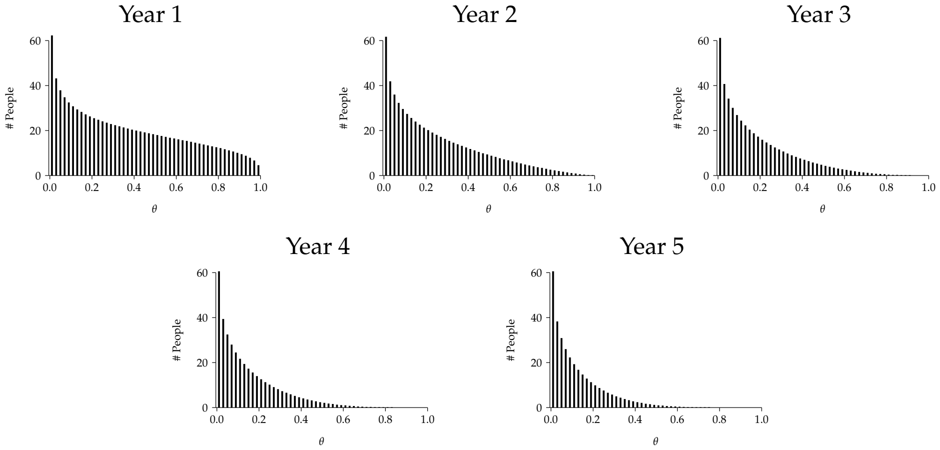
$$g(\theta \mid \gamma, \delta, \text{ active in } n) = \frac{S(n-1 \mid \theta) g(\theta \mid \gamma, \delta)}{S(n-1 \mid \gamma, \delta)}$$
$$= \frac{\theta^{\gamma-1}(1-\theta)^{\delta+n-2}}{B(\gamma, \delta+n-1)},$$

which is a beta distribution with parameters γ and $\delta + n - 1$.

Distributions of churn probabilities



Distributions of churn probabilities



Computing $E(DRL)$

Integrating over the posterior distribution of θ:

$$
\begin{aligned}
&E[DRL(d) \mid \gamma, \delta, \text{ active in } n] \\
&\quad = \int_0^1 \Big\{ E[DRL(d) \mid \theta, \text{ active in } n] \\
&\qquad\qquad \times g(\theta \mid \gamma, \delta, \text{ active in } n) \Big\} d\theta \\
&\quad = \int_0^1 \frac{(1-\theta)(1+d)}{d+\theta} \frac{\theta^{\gamma-1}(1-\theta)^{\delta+n-2}}{B(\gamma, \delta+n-1)} d\theta \\
&\quad = \frac{1+d}{B(\gamma, \delta+n-1)} \int_0^1 \theta^{\gamma-1}(1-\theta)^{\delta+n-1}(d+\theta)^{-1} d\theta \\
&\quad = \left(\frac{\delta+n-1}{\gamma+\delta+n-1} \right) {}_2F_1\big(1, \delta+n; \gamma+\delta+n; \tfrac{1}{1+d}\big)
\end{aligned}
$$

An alternative derivation

Assuming lifetimes are distributed BG,

$$
\begin{aligned}
&E[DRL(d) \mid \gamma, \delta, \text{active in } n] \\
&\qquad = \sum_{t=n}^{\infty} \frac{S(t \mid T > n-1; \gamma, \delta)}{(1+d)^{t-n}} \\
&\qquad = \sum_{t=n}^{\infty} \frac{S(t \mid \gamma, \delta)}{S(n-1 \mid \gamma, \delta)} \left(\frac{1}{1+d}\right)^{t-n} \\
&\qquad = \sum_{t=n}^{\infty} \frac{B(\gamma, \delta + t)}{B(\gamma, \delta + n - 1)} \left(\frac{1}{1+d}\right)^{t-n} \\
&\qquad = \left(\frac{\delta + n - 1}{\gamma + \delta + n - 1}\right) {}_2F_1\big(1, \delta + n; \gamma + \delta + n; \tfrac{1}{1+d}\big)
\end{aligned}
$$

Further reading

Fader, Peter S. and Bruce G. S. Hardie (2010), "Customer-Base Valuation in a Contractual Setting: The Perils of Ignoring Heterogeneity," *Marketing Science*, **29** (January–February), 85–93. `http://brucehardie.com/papers/022/`

Fader, Peter S. and Bruce G.S. Hardie (2017), "Exploring the Distribution of Customer Lifetime Value (in Contractual Settings)." `http://brucehardie.com/notes/035/`

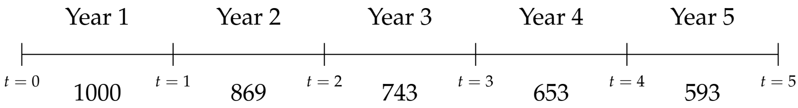
Fader, Peter S. and Bruce G.S. Hardie (2018), "The Mean and Variance of Customer Lifetime Value in Contractual Settings." `http://brucehardie.com/notes/036/`

Summary

- How to compute CLV in contractual settings
  - Recognising the need to project customer survival beyond the observed data
  - Understanding the distinction between the value of a new versus existing customer
- How to use a probability model to forecast customer survival
- Understanding the phenomenon of retention-rate dynamics

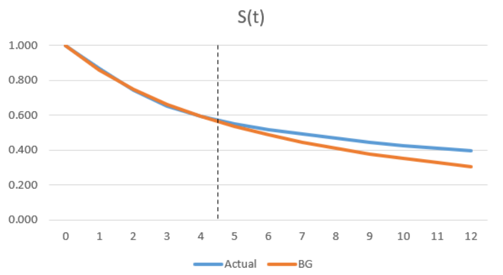
Beyond the basic BG model

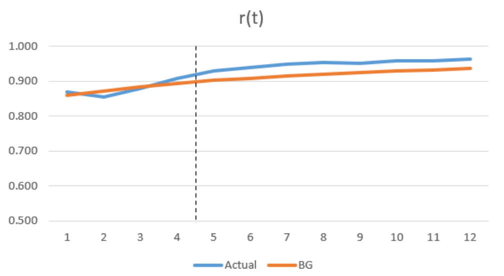
A second dataset



	A	B	C	D	E	F	G
1	gamma	1.281					
2	delta	7.790					
3	LL	-1225.13					
4							
5	Year	t	# Cust.	# Lost	P("die")	S(t)	
6	1	0	1000			1.0000	
7	2	1	869	131	0.1412	0.8588	-256.43
8	3	2	743	126	0.1092	0.7496	-279.00
9	4	3	653	90	0.0867	0.6628	-220.05
10	5	4	593	60	0.0703	0.5925	-159.27
11							-310.38

Model performance





What assumptions have we made?

- ✓ Heterogeneity
- ✗ Duration dependence
- ✗ Covariates
- ✗ Contagion

Adding duration dependence

We use Nakagawa and Osaki's (1975) discrete Weibull (dW) distribution:

$$S(t \mid \theta, c) = (1-\theta)^{t^c},\ c > 0\,.$$

Simple, mathematically convenient, and flexible; the best analogue of the (continuous) Weibull distribution.

- When $c > 1$, we have *positive duration dependence* $\Longrightarrow$ the churn probability increases over time (retention probability decreases over time).
- When $c < 1$, we have *negative duration dependence* $\Longrightarrow$ the churn probability decreases over time.

The beta-discrete-Weibull (BdW) model

Assuming heterogeneity in θ is distributed beta,

$$\begin{aligned}
S(t \mid \gamma, \delta, c) &= \int_0^1 S(t \mid \theta, c) g(\theta \mid \gamma, \delta)\, d\theta \\
&= \int_0^1 (1-\theta)^{t^c} \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma, \delta)}\, d\theta \\
&= \frac{1}{B(\gamma, \delta)} \int_0^1 \theta^{\gamma-1}(1-\theta)^{\delta+t^c-1}\, d\theta \\
&= \frac{B(\gamma, \delta+t^c)}{B(\gamma, \delta)}\,.
\end{aligned}$$

Numerical evaluation of the beta function

Not all computing environments have a `beta` function (or even a `gamma` function). Recall that the beta function $B(\gamma, \delta)$ can be expressed in terms of gamma functions:

$$B(\gamma, \delta) = \frac{\Gamma(\gamma)\Gamma(\delta)}{\Gamma(\gamma + \delta)} .$$

Most computing environments have a function that evaluates $\ln\big(\Gamma(\cdot)\big)$. In Excel, we have `gammaln`.

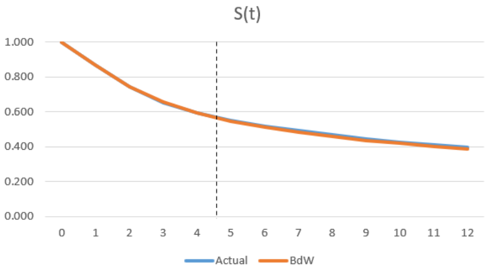
Therefore,

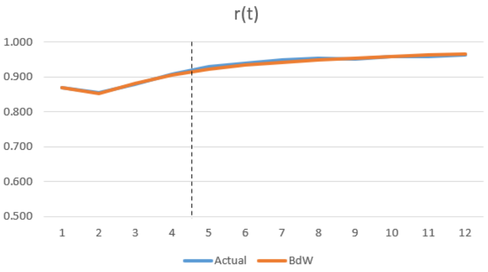
$$\begin{aligned} \Gamma(\gamma) &= \texttt{exp(gammaln(}\gamma\texttt{))} \\ B(\gamma, \delta) &= \texttt{exp(gammaln(}\gamma\texttt{)+gammaln(}\delta\texttt{)-gammaln(}\gamma\texttt{+}\delta\texttt{))} \end{aligned}$$

Estimating model parameters

	A	B	C	D	E	F	G
1	gamma	0.259		B(gamma, delta)	3.210		
2	delta	1.722					
3	c	1.585					
4	LL	-1222.75					
5							
6	Year	t	# Cust.	# Lost	P("die")	BdW	
7	1	0	1000			1.0000	
8	2	1	869	131	0.1309	0.8691	-266.41
9	3	2	743	126	0.1272	0.7419	-259.79
10	4	3	653	90	0.0871	0.6548	-219.62
11	5	4	593	60	0.0618	0.5929	-166.98
12							-309.94
13							
14	E1	=EXP(GAMMALN(B1)+GAMMALN(B2)-					
15		GAMMALN(B1+B2))					
16	F7	=EXP(GAMMALN($B$1)+GAMMALN($B$2+B7^$B$3)-					
17		GAMMALN($B$1+$B$2+B7^$B$3))/$E$1					
18	E8	=F7-F8					

Model performance





Further reading

Fader, Peter S., Bruce G. S. Hardie, Yuzhou Liu, Joseph Davin, and Thomas Steenburgh (2018), “"How to Project Customer Retention" Revisited: The Role of Duration Dependence,” *Journal of Interactive Marketing*, **43** (August), 1–16.

Jaganathan, Srihari and Ka Lok Lee (2019), “Simple Probability Models for Predicting Aggregate or Sparse Data: An Empirical Analysis of Projecting Patient Persistency,” *Journal of the Pharmaceutical Management Science Association*, Spring, Article 1. `http://www.pmsa.net/jpmsa-vol07-article01`

[R Package] “foretell: Projecting Customer Retention Based on Fader and Hardie Probability Models.” `https://CRAN.R-project.org/package=foretell`

From discrete to continuous time

We have considered settings where the discrete contract period is annual.

In other settings, we have a quarterly contract period. Or a monthly contract. Or a weekly contract. Or the customer can cancel at any point in time with "immediate" effect.

At some level, we stop treating time as discrete and view it as continuous.

From discrete to continuous time

As we go from flipping the coin every year, to every month, to every week, to every day, to every hour, …

BG $\longrightarrow$ gamma mixture of exponentials
Pareto Type II

$$S(t \mid r, \alpha) = \left(\frac{\alpha}{\alpha + t}\right)^r$$

BdW $\longrightarrow$ gamma mixture of Weibulls
(Generalized) Burr Type XII

$$S(t \mid r, \alpha, c) = \left(\frac{\alpha}{\alpha + t^c}\right)^r$$

From discrete to continuous time

The fit and associated forecasts of the Pareto Type II are exactly the same as those of the BG. However, we tend to favor a discrete-time model given ease of story telling.

THE AMERICAN STATISTICIAN
2019, VOL. 73, NO. 3, 288–295: Teacher's Corner
https://doi.org/10.1080/00031305.2019.1543134

Taylor & Francis
Taylor & Francis Group

Check for updates

Exploring the Equivalence of Two Common Mixture Models for Duration Data

Peter S. Fader[a], Bruce G. S. Hardie[b], Daniel McCarthy[c], and Ramnath Vaidyanathan[d]

[a]The Wharton School, University of Pennsylvania, Philadelphia, PA; [b]London Business School, London, UK; [c]Goizueta Business School, Emory University, Atlanta, GA; [d]DataCamp Inc., New York, NY

ABSTRACT
The beta-geometric (BG) distribution and the Pareto distribution of the second kind (P(II)) are two basic models for duration-time data that share some underlying characteristics (i.e., continuous mixtures of memoryless distributions), but differ in two important respects: first, the BG is the natural model to use when the event of interest occurs in discrete time, while the P(II) is the right choice for a continuous-time setting. Second, the underlying mixing distributions (the beta and gamma for the BG and P(II), respectively), are very different—and often believed to be noncomparable with each other. Despite these and other key differences, the two models are strikingly similar in terms of their fit and predictive performance as well as their parameter estimates. We explore this equivalence, both empirically and analytically, and discuss the implications from both a substantive and methodological standpoint.

ARTICLE HISTORY
Received September 2017
Accepted October 2018

KEYWORDS
Beta-geometric; Grassia(II); Pareto of the second kind

Adding time-varying covariate effects

Discrete time:

Fader, Peter S. and Bruce G.S. Hardie (2020), "Incorporating Time-Varying Covariates in a Simple Mixture Model for Discrete-Time Duration Data."
`https://brucehardie.com/notes/037/`

Continuous time:

Fader, Peter S., Bruce G.S. Hardie, and Robert Zeithammer (2003), "Forecasting New Product Trial in a Controlled Test Market Environment," *Journal of Forecasting*, **22** (August), 391–410.

Schweidel, David A., Peter S. Fader, and Eric T. Bradlow (2008), "Understanding Service Retention Within and Across Cohorts Using Limited Information," *Journal of Marketing*, **72** (January), 82–94.

But when do we really want to include the effects of covariates?

Reflections on data structures

Raw data:

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
$\vdots$		$\vdots$		$\vdots$	
0999	1	1	1	0	0
1000	1	0	0	0	0
	1000	631	468	382	326

Standard summary (single cohort):

Calendar Time $\rightarrow$				
n_1	n_2	n_3	$\ldots$	n_I

Reflections on data structures

Standard summary (multiple cohorts):

Cohort	Calendar Time $\rightarrow$				
1	n_{11}	n_{12}	n_{13}	$\dots$	n_{1I}
2		n_{22}	n_{23}	$\dots$	n_{2I}
3			n_{33}	$\dots$	n_{3I}
$\vdots$				$\ddots$	$\vdots$
I					n_{II}
	$n_{.1}$	$n_{.2}$	$n_{.3}$	$\dots$	$n_{.I}$

Reflections on data structures

Cohort	Calendar Time →				
1	n_{11}				n_{1I}
2		n_{22}			n_{2I}
⋮					⋮
I−1				$n_{I-1,I-1}$	$n_{I-1,I}$
I					n_{II}

Cohort	Calendar Time →				
1	n_{11}				
2		n_{22}			
⋮					
I−1				$n_{I-1,I-1}$	
I					n_{II}
	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.I-1}$	$n_{.I}$

Cohort	Calendar Time →				
1					n_{1I}
2					n_{2I}
⋮					⋮
I−1					$n_{I-1,I}$
I					n_{II}
	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.I-1}$	$n_{.I}$

Cohort	Calendar Time →	
1	n_{1I-1}	n_{1I}
2	n_{2I-1}	n_{2I}
⋮	⋮	⋮
I−1	$n_{I-1,I-1}$	$n_{I-1,I}$
I		n_{II}

Fader, Peter S. and Bruce G.S. Hardie (2007), “Fitting the sBG Model to Multi-Cohort Data.” `http://brucehardie.com/notes/017/`

Reflections on data structures

Netflix (10Q/10K Data)

Year	Qtr	Begin	Add	Subtract	End
1999	4	NA	NA	NA	NA
2000	1	NA	NA	NA	NA
2000	2	NA	NA	NA	NA
2000	3	NA	NA	NA	NA
2000	4	NA	NA	NA	NA
2001	1	NA	NA	NA	NA
2001	2	NA	88	NA	308
2001	3	308	107	NA	334
2001	4	334	NA	NA	456
2002	1	456	312	165	603
2002	2	603	236	169	670
⋮	⋮	⋮	⋮	⋮	⋮
2009	1	9390	2413	1493	10310
2009	2	10310	1936	1647	10599
2009	3	10599	2180	1670	11109
2009	4	11109	2803	1644	12268
2010	1	12268	3492	1793	13967
2010	2	13967	3059	2025	15001
2010	3	15001	4101	2169	16933
2010	4	16933	5649	2572	20010
2011	1	20010	6299	3512	22797
2011	2	22797	5315	3518	24594
2011	3	24594	4714	5519	23789
2011	4	23789	NA	NA	24395
2012	1	24395	NA	NA	26074
2012	2	26074	NA	NA	26494
2012	3	26494	NA	NA	27507
2012	4	27507	NA	NA	29368
2013	1	29368	NA	NA	31396
2013	2	31396	NA	NA	32029
2013	3	32029	NA	NA	33314
2013	4	33314	NA	NA	35642
2014	1	35642	NA	NA	37896

Reflections on data structures

Daniel M. McCarthy, Peter S. Fader, & Bruce G.S. Hardie

Valuing Subscription-Based Businesses Using Publicly Disclosed Customer Data

The growth of subscription-based commerce has changed the types of data that firms report to external shareholders. More than ever, companies are discussing and disclosing information on the number of customers acquired and lost, customer lifetime value, and other data. This has fueled an increasing interest in linking the value of a firm's customers to the overall value of the firm, with the term "customer-based corporate valuation" being used to describe such efforts. Although several researchers in the fields of marketing and accounting have explored this idea, their underlying models of customer acquisition and retention do not adequately reflect the empirical realities associated with these behaviors, and the associated valuation models do not meet the standards of finance professionals. The authors develop a framework for valuing subscription-based firms that addresses both issues, and they apply it to data from DISH Network and Sirius XM Holdings.

Keywords: customer lifetime value, customer equity, shareholder value, valuation

ISSN: 0022-2429 (print)
1547-7185 (electronic)

Journal of Marketing
Vol. 81 (January 2017), 17–35
DOI: 10.1509/jm.15.0519

Looking ahead to tomorrow

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
0006	1	0	0	0	0
⋮		⋮		⋮	
0998	1	0	0	0	0
0999	1	1	1	0	0
1000	1	0	0	0	0
	1000	631	468	382	326

ID	1995	1996	1997	1998	1999	2000	2001
100001	1	0	0	0	0	0	0
100002	1	0	0	0	0	0	0
100003	1	0	0	0	0	0	0
100004	1	0	1	0	1	1	1
100005	1	0	1	1	1	0	1
100006	1	1	1	1	0	1	0
100007	1	1	0	1	0	1	0
100008	1	1	1	1	1	1	1
100009	1	1	1	1	1	1	0
100010	1	0	0	0	0	0	0
⋮		⋮		⋮		⋮	
111102	1	1	1	1	1	1	1
111103	1	0	1	1	0	1	1
111104	1	0	0	0	0	0	0
	11104	5652	4674	4019	3552	3555	3163