Customer Lifetime Value (I): Contractual Settings

Lecture 6

Customer Analytics

Tilburg University

Agenda

Conceptual definition & applications

Primitives of CLV

- Model 1: geometric model of retention
- Model 2: shifted Beta-geometric model of retention

Heterogeneity and increasing retention rates

Calculating CLV/RLV

Customer Lifetime Value

CLV is the present value of the future profits associated with a particular customer

Customer Lifetime Value

CLV is the present value of the future profits associated with a particular customer

Time value of money:
€1 today > €1 tomorrow

Customer Lifetime Value

CLV is the present value of the future profits associated with a particular customer

Uncertain from today's perspective—
it's a forecast/prediction!

Time value of money:
€1 today > €1 tomorrow

Customer Lifetime Value

CLV is the present value of the future profits associated with a particular customer

Uncertain from today's perspective—
it's a forecast/prediction!

Revenues less customer-specific variable costs: marketing, billing, servicing, tech support

Time value of money:
€1 today > €1 tomorrow

Customer Lifetime Value

CLV is the present value of the future profits associated with a particular customer

Uncertain from today's perspective—it's a forecast/prediction!

Revenues less customer-specific variable costs: marketing, billing, servicing, tech support

Time value of money:
€1 today > €1 tomorrow

Measured for (ideally) individual customers;
recognize every customer is different

Two settings

Netflix	Bol
Netflix said it added 7.05 million subscribers during the fiscal 4th quarter. For the quarter, Netflix added 1.93 million memberships in the U.S. and 5.12 million internationally.	Bol.com opened its doors on 30 March 1999. Eighteen years later, the store has more than 7 million active customers* in the Netherlands and Belgium and an assortment of almost 15 million items. * customers are considered active when they have placed an order during the preceding twelve-month period

Which firm knows how many customers they currently have?

Contractual settings	Non-contractual (a.k.a. transactional) settings
• Minority of companies, but growing	• Majority of companies
• Customers notify the company to quit (ending contract)	• Customers silently leave
• Any subscription business model: gym membership, internet/cable, bank, insurance	• Grocery store, retailers, fast moving consumer goods (FMCG), hotels, airlines, media
• Focus is on retention rate, because quitting is observed.	• Focus is on repeat purchasing
Today	Next time

Schmittlein, Morrison and Colombo (1987)

Retention rate and survivor function

$T = \{0, 1, 2, \dots\}$ is the (discrete) lifetime of a customer.

The probability that someone remains a customer longer that t periods is the survivor function

$$\mathrm{P}(T > t) = S(t)$$

where $S(0) = 1$

The probability that a customer who has already remained $t - 1$ periods, remains at least one period longer is the retention rate:

$$\mathrm{P}(T > t \mid T > t - 1) = r(t)$$

Retention rate and survivor function

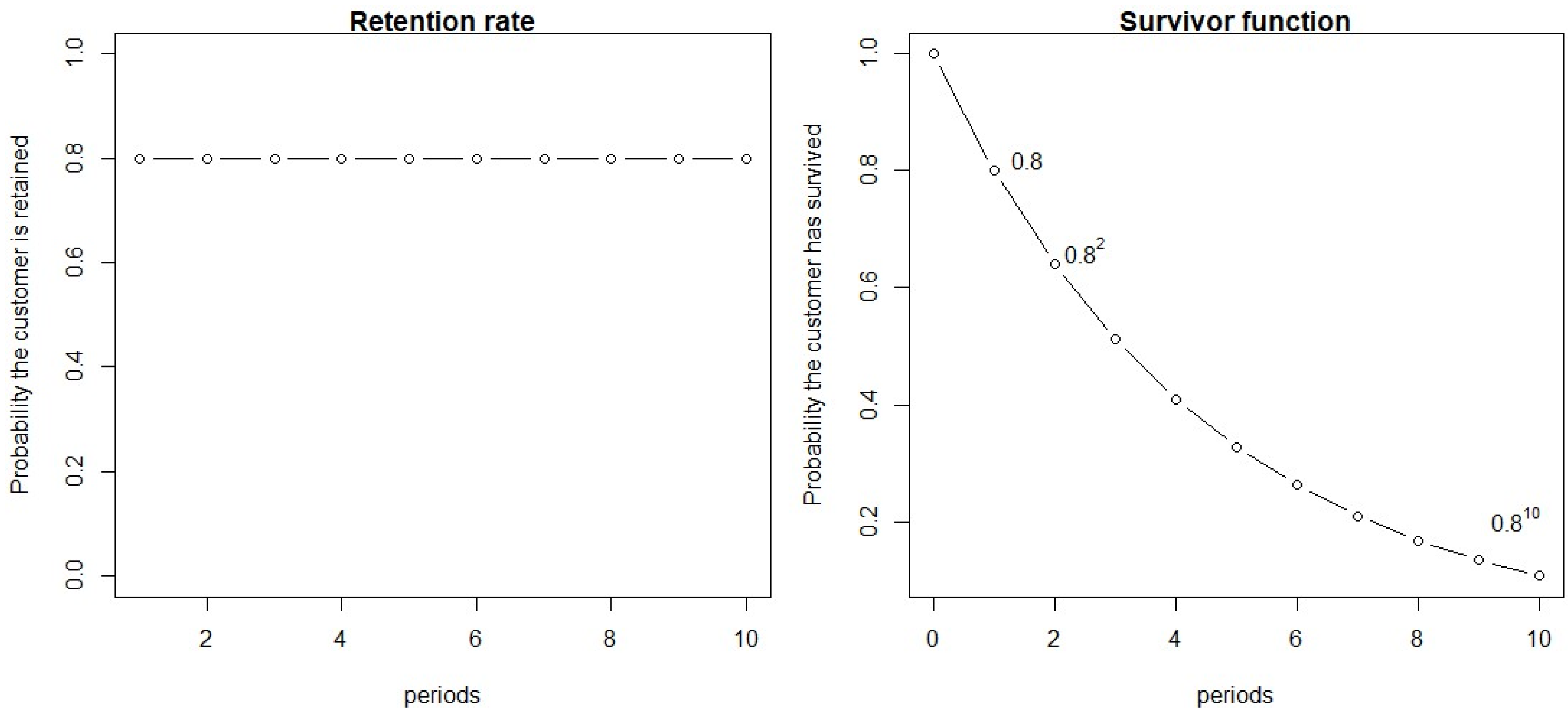
- If you know one, you can calculate the other
- The survivor function at t is the product of the retention rates until t

$$S(t) = \mathrm{r}(1)\,\mathrm{r}(2)\ldots\mathrm{r}(t) = \prod_{j=1}^{t} r(j)$$

- The retention rate at t is the ratio of survivor functions at t and $t-1$

$$r(t) = \frac{S(t)}{S(t-1)}$$

The most basic model is to say the retention rate is constant each period. Here we set $r(t) = 0.8$



CLV geometric model

(Assume all these things are constant across periods):

1. Margins (accounting)

   m = revenue less costs of marketing, selling, production, servicing

2. Discount rate per period (finance)

   d = opportunity cost of capital

3. Retention (marketing)

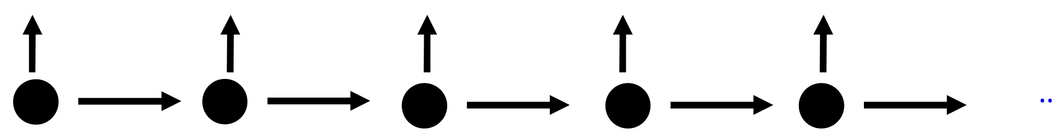
   p = probability customer survives (does not quit) for one more period

CLV geometric model: the formula

E[CLV]=

A x B x C $\quad m + \frac{mp}{1+d} + \frac{mp^2}{(1+d)^2} + \frac{mp^3}{(1+d)^3} + \frac{mp^4}{(1+d)^4} \cdots = \boxed{\frac{m(1+d)}{1+d-p}}$

A	Discounted to Present		$1/(1+d)$	$1/(1+d)^2$	$1/(1+d)^3$	$1/(1+d)^4$
B	Prob. cust. survives past t: $S(t)$		p	p^2	p^3	p^4
C	Margin	m	m	m	m	m



Time (0 is present)	0	1	2	3	4

Receive m when contract is initiated at t = 0; customer renews or stays at time period 1 with probability p, in which case we receive another m discounted by $1/(1+d)$

CLV calculations

$$E[CLV] = m + \frac{m\,p}{(1+d)} + \frac{m\,p^2}{(1+d)^2} + \frac{m\,p^3}{(1+d)^3} + \frac{m\,p^4}{(1+d)^4} + \cdots$$

$$= m\sum_{n=0}^{\infty}\left(\frac{p}{1+d}\right)^n$$

$$= \frac{m}{1-\left(\frac{p}{1+d}\right)} = \frac{m(1+d)}{1+d-p}$$

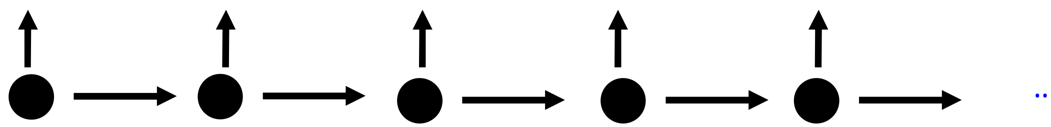
Geometric series

$$\boxed{\sum_{n=0}^{\infty} k^n = \frac{1}{1-k}, \qquad 0 < k < 1}$$

Example: m= 100, p = 0.80, d = 0.10

E[CLV]=

		100	+ 72	+ 53	+ 38	+ 28	...	= $\boxed{367}$
	A x B x C	m	+ $\frac{mp}{1+d}$	+ $\frac{mp^2}{(1+d)^2}$	+ $\frac{mp^3}{(1+d)^3}$	+ $\frac{mp^4}{(1+d)^4}$	...	= $\boxed{\frac{m(1+d)}{1+d-p}}$
A	Discounted to Present		$1/(1+d)$	$1/(1+d)^2$	$1/(1+d)^3$	$1/(1+d)^4$		
B	Prob. cust. survives past t: $S(t)$		p	p^2	p^3	p^4		
C	Margin	m	m	m	m	m		



Time (0 is present)	0	1	2	3	4	...

Expected customer lifetime value of this customer is €367

CLV vs. residual lifetime value (RLV)

CLV = for a not-yet-acquired customer, <u>from first purchase</u>

RLV = **<u>R</u>**esidual **<u>L</u>**ifetime **<u>V</u>**alue of **already existing** customer, including future purchases

- E.g., what's the value of a customer who with an "age" of 3 (3 years as a customer)?

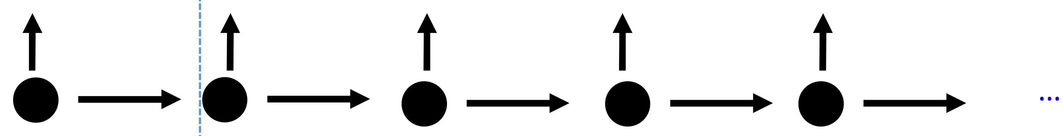
RLV: right before next renewal

A x B x C

$$mp + \frac{mp^2}{(1+d)^1} + \frac{mp^3}{(1+d)^2} + \frac{mp^4}{(1+d)^3} \cdots = \boxed{\frac{m\,p\,(1+d)}{1+d-p}}$$

present

A	Discounted to Present			$1/(1+d)^1$	$1/(1+d)^2$	$1/(1+d)^3$
B	**Prob. cust. survives past t: $S(t)$**		p	p^2	p^3	p^4
C	Margin		m	m	m	m



Time (t is present)	t-1	t*	t+1	t+2	t+3

history

future

Right before customer makes a renewal decision, receive m if contract is renewed with probability p;

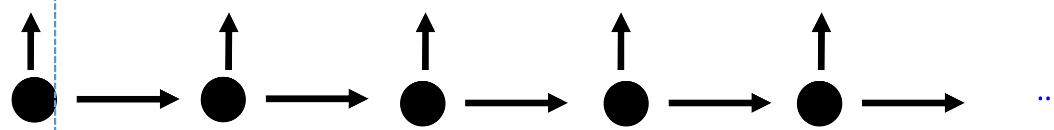
RLV: right after renewal

present

A x B x C

$$\frac{mp}{(1+d)^1} + \frac{mp^2}{(1+d)^2} + \frac{mp^3}{(1+d)^3} + \frac{mp^4}{(1+d)^4} \cdots = \boxed{\frac{m\,p}{1+d-p}}$$

A	Discounted to Present	$1/(1+d)^1$	$1/(1+d)^2$	$1/(1+d)^3$	$1/(1+d)^4$
B	**Prob. cust. survives past t: $S(t)$**	p	p^2	p^3	p^4
C	Margin	m	m	m	m



Time (t is present) t-1 t* t+1 t+2 t+3

history future

Right after customer makes a renewal decision, receive m in one period if contract is renewed with probability p, discounted to present.

Uses of CLV

- Tell you what each of your individual customers is worth (useful beyond other measures like NPS, clicks, etc)
  - Byproduct: you can use it to predict purchasing, benchmark with marketing efforts

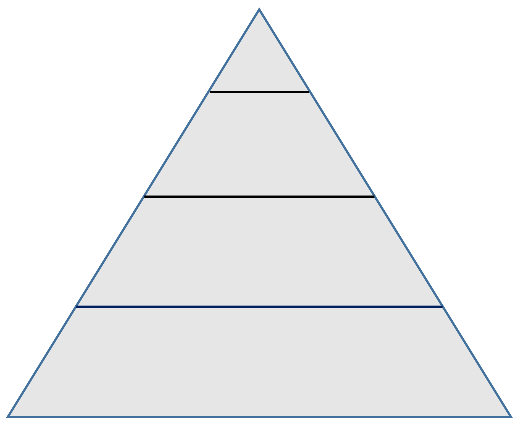
- Upper bound on spending for customer acquisition, retention, development

$$CPA < CLV$$

- Value company: add up all customer's CLV to measure customer equity (cf. brand equity)

Forming segments based on CLV

Create segments: separate most valuable customers from everyone else, focus efforts



Platinum (5% of customers): 50% of total CLV

Gold (15%): 30% of **total** CLV

Silver (30%): 20% of **total** CLV

Lead (50%): 0% (or negative!) of **total** CLV

Valid assumption?

We assume a constant retention rate over time

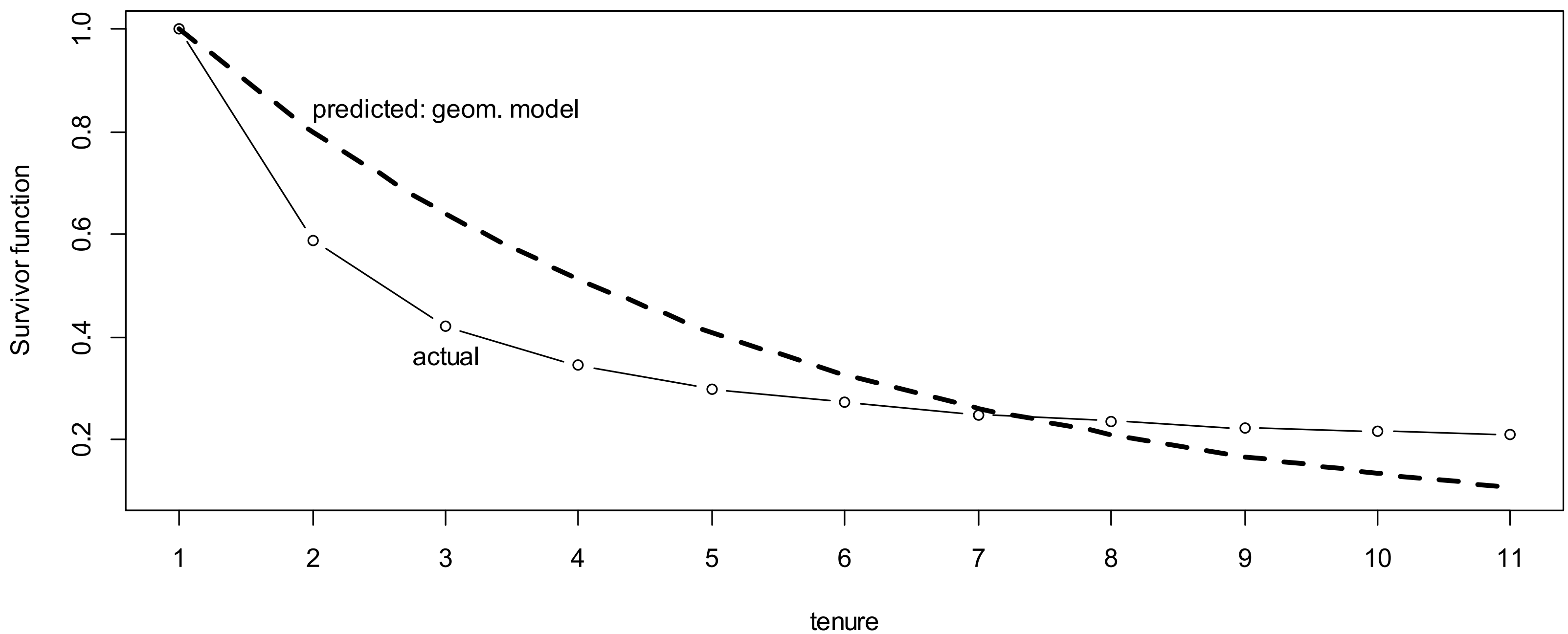
p = probability customer survives one period

Same retention rate applies to a:

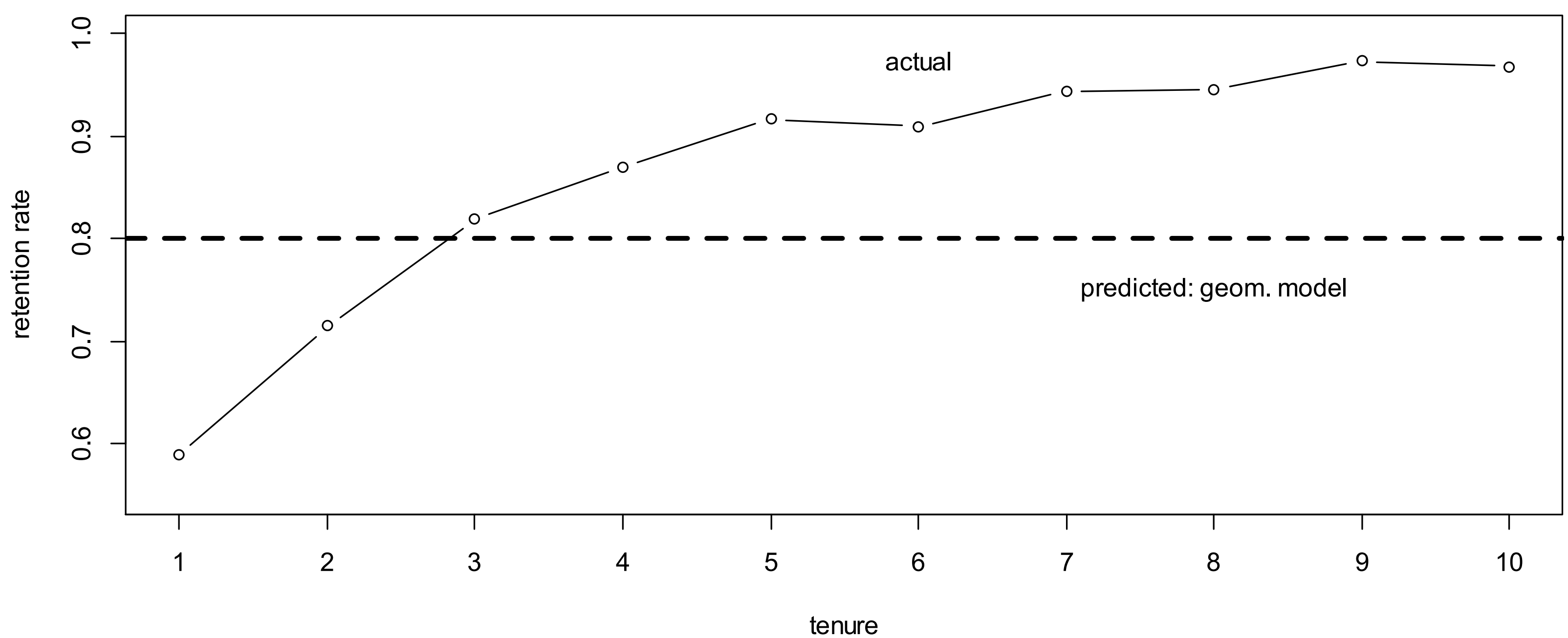
a) new customer
b) old customer

Can we test this?

- Let’s take a cohort of customers who all started at the same time and record how long they remained a customer
  - Why study customers who started at the same time?
- Cohort analysis
- We can plot the survival of these customers

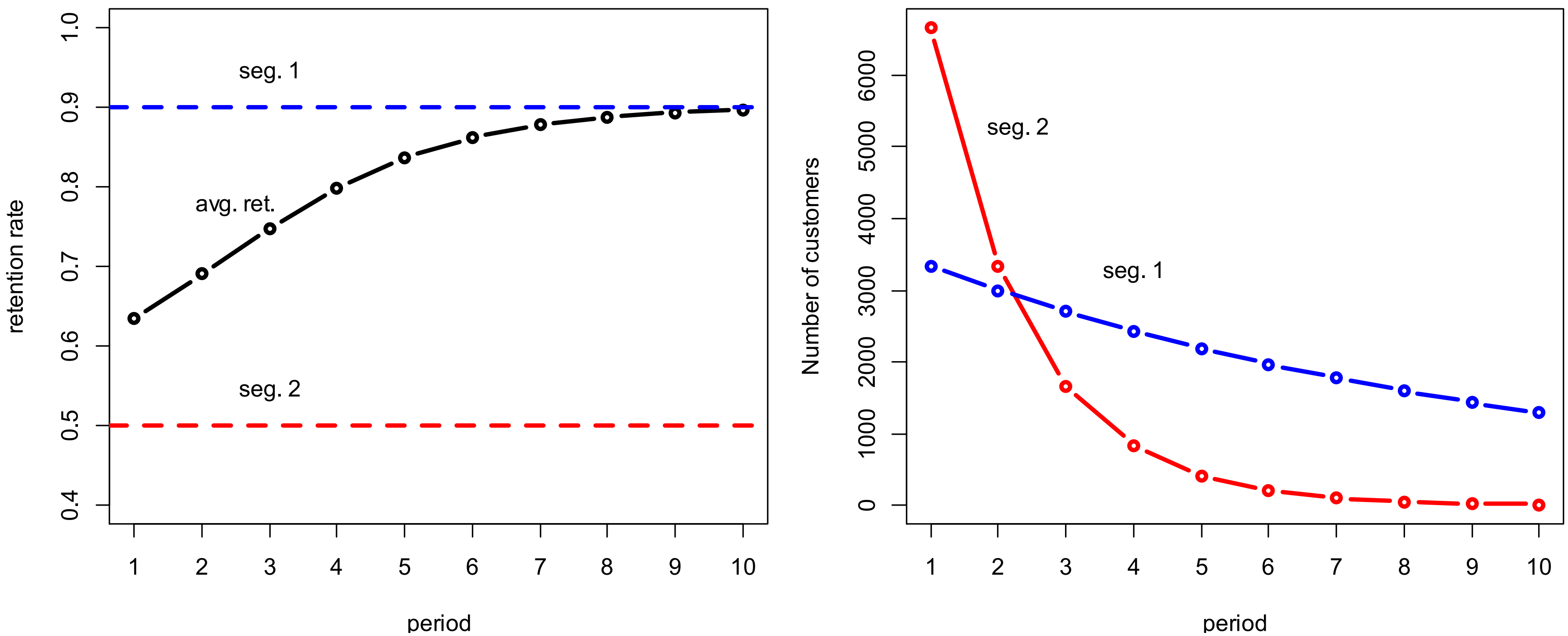


predicted: geom. model
actual
Survivor function
tenure
1.0
0.8
0.6
0.4
0.2
1
2
3
4
5
6
7
8
9
10
11



Why do cohort-level retention rates increase?

Ruse of heterogeneity



Review: geometric model

At the end of each period, each customer renews his contract with (constant and unobserved) probability p.

Same as before, but we write it in terms of $\mathrm{p} = 1 - \theta$

$$S(t) = p^t = (1-\theta)^t$$

In this model, the retention rate at any time t is

$$\mathrm{r(t)} = \frac{S(t)}{S(t-1)} = (1-\theta)$$

Fader and Hardie (2007), How to Project Customer Retention

Prior distribution

Churn probabilities, θ, vary across customers according to a Beta distribution.

Instead of one θ for everyone (or 2 segment types as earlier), we have a *prior* distribution of θ that depends on two parameters, a and b.

$$f(\theta|a,b) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a,b)}$$

where $a, b > 0$. $B(a,b)$ is the beta function (details).

We will use this formula later

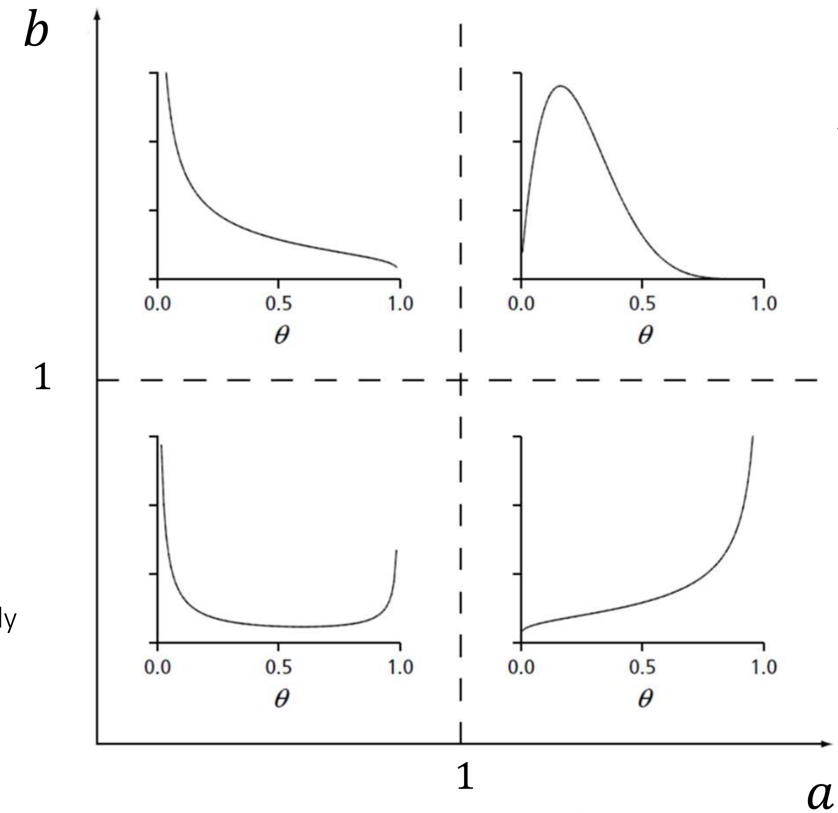
$$\frac{B(a,b+x)}{B(a,b+x-1)} = \frac{b+x-1}{a+b+x-1}$$

Fader and Hardie (2007), How to Project Customer Retention

General shapes of the beta distribution

Spread out across whole range, but most likely to stay

Most customers fall within a certain range



Two types of customers: extremely likely & unlikely

Spread out across whole range, but most likely to churn

Fader and Hardie (2007), How to Project Customer Retention

Shifted Beta-geometric model

For a randomly chosen individual, we have to integrate θ over all possibilities

$$S(t|a,b) = \int_0^1 \underbrace{S(t|\theta)}_{\text{Shifted geo. dist.}} \underbrace{f(\theta|a,b)}_{\text{Beta dist.}} \, d\theta$$

$$= \int_0^1 \frac{(1-\theta)^{t+b-1}\,\theta^{a-1}}{B(a,b)} \quad d\theta$$

$$= \frac{B(a,b+t)}{B(a,b)}$$

$$P(T=t) = \frac{B(a+1, b+t-1)}{B(a,b)}$$

Tilburg University

Retention rate in shifted beta-geometric model

$$\mathrm{r(t)} = \frac{S(t)}{S(t-1)}$$

$$= \frac{B(a, b+t)/B(a,b)}{B(a, b+t-1)/B(a,b)}$$

$$= \frac{b+t-1}{a+b+t-1}$$

Likelihood

We have data on active customers starting at the first renewal opportunity

Period	Customers Remaining
0	N
1	n_1
..	
t	n_t

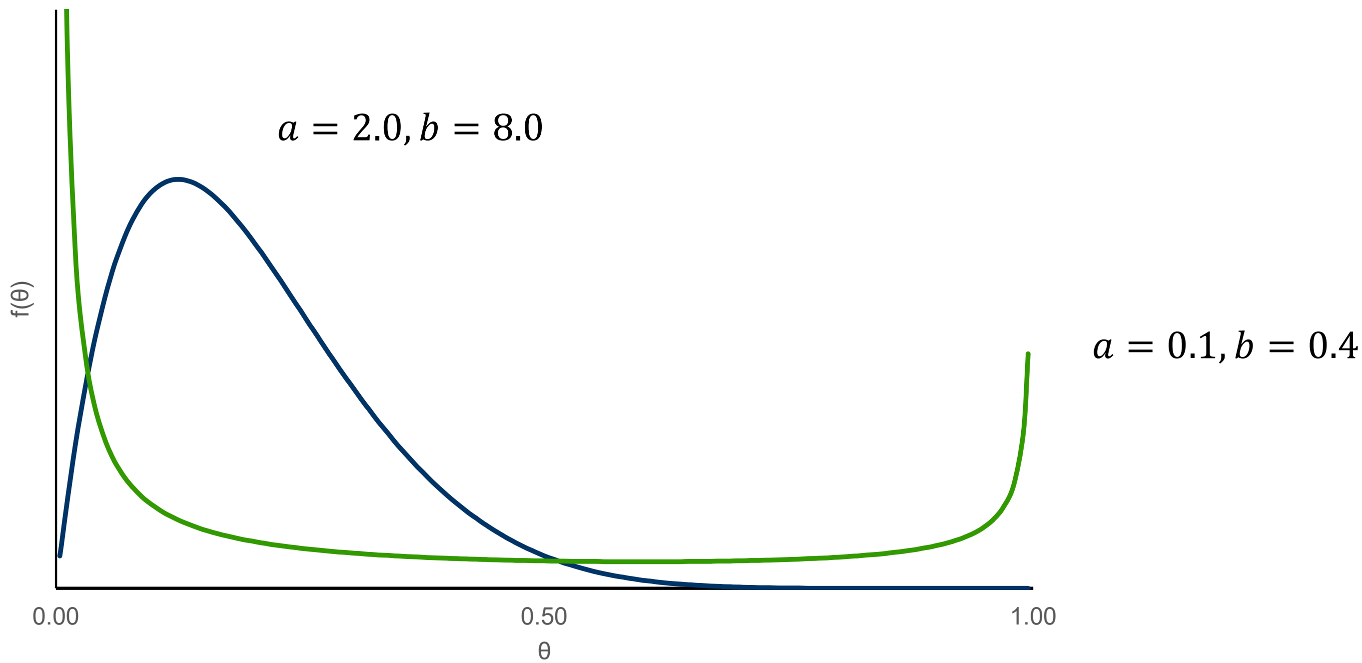
$N - n_1$ customers quit in period 1

n_t customers remaining in period t

$$L(a, b) = P(T = 1)^{N-n_1} \times P(T = 2)^{n_1-n_2} \dots \times P(T = t)^{n_{t-1}-n_t} \times S(t)^{n_t}$$

Two different groups

Two populations with the same mean, but different variance (green more than blue)



Beta distribution

$$\mathrm{E}[\theta|a,b] = \frac{a}{a+b}$$

$$\mathrm{var}[\theta|a,b] = \frac{ab}{(a+b)^2(a+b+1)}$$

https://en.wikipedia.org/wiki/Beta_distribution

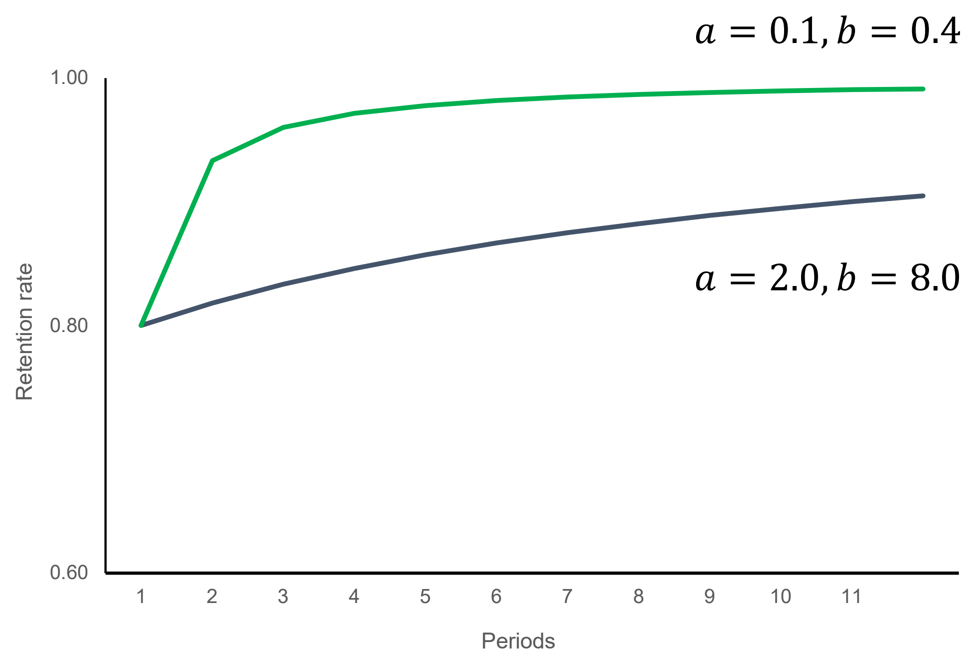
Implications for retention rate

Customer heterogeneity

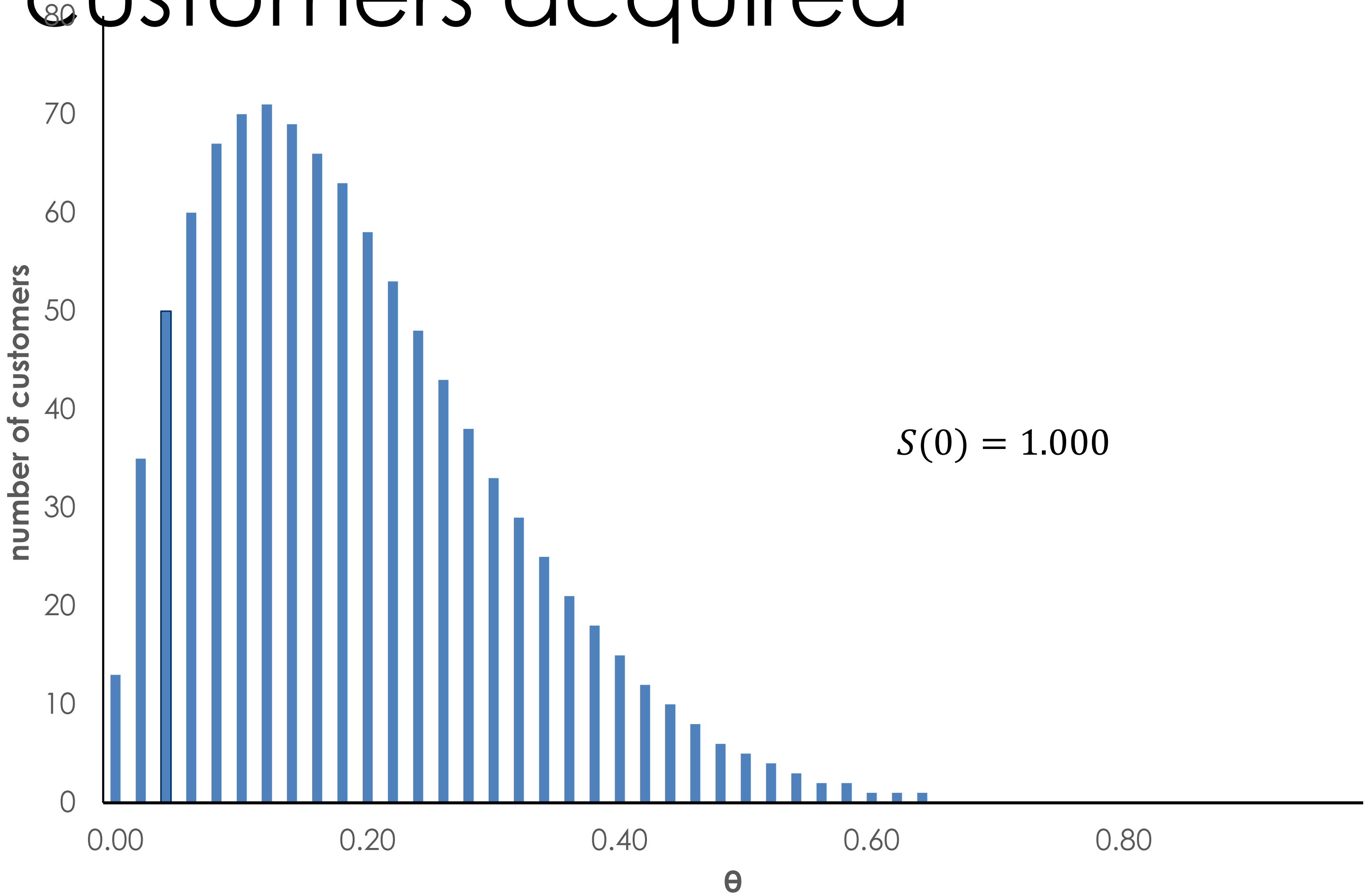
- Turns out the distribution of customer types creates that increasing retention rates over time
- If customer types differ a lot, i.e., the customer base is heterogeneous, retention rates rise quickly
  $$a = 0.1, b = 0.4$$
- If customer types differ a little, i.e., the customer base is homogeneous, retention rates rise slowly
  $$a = 2.0, b = 8.0$$

What's the intuition? (next slides)

Retention rate dynamics



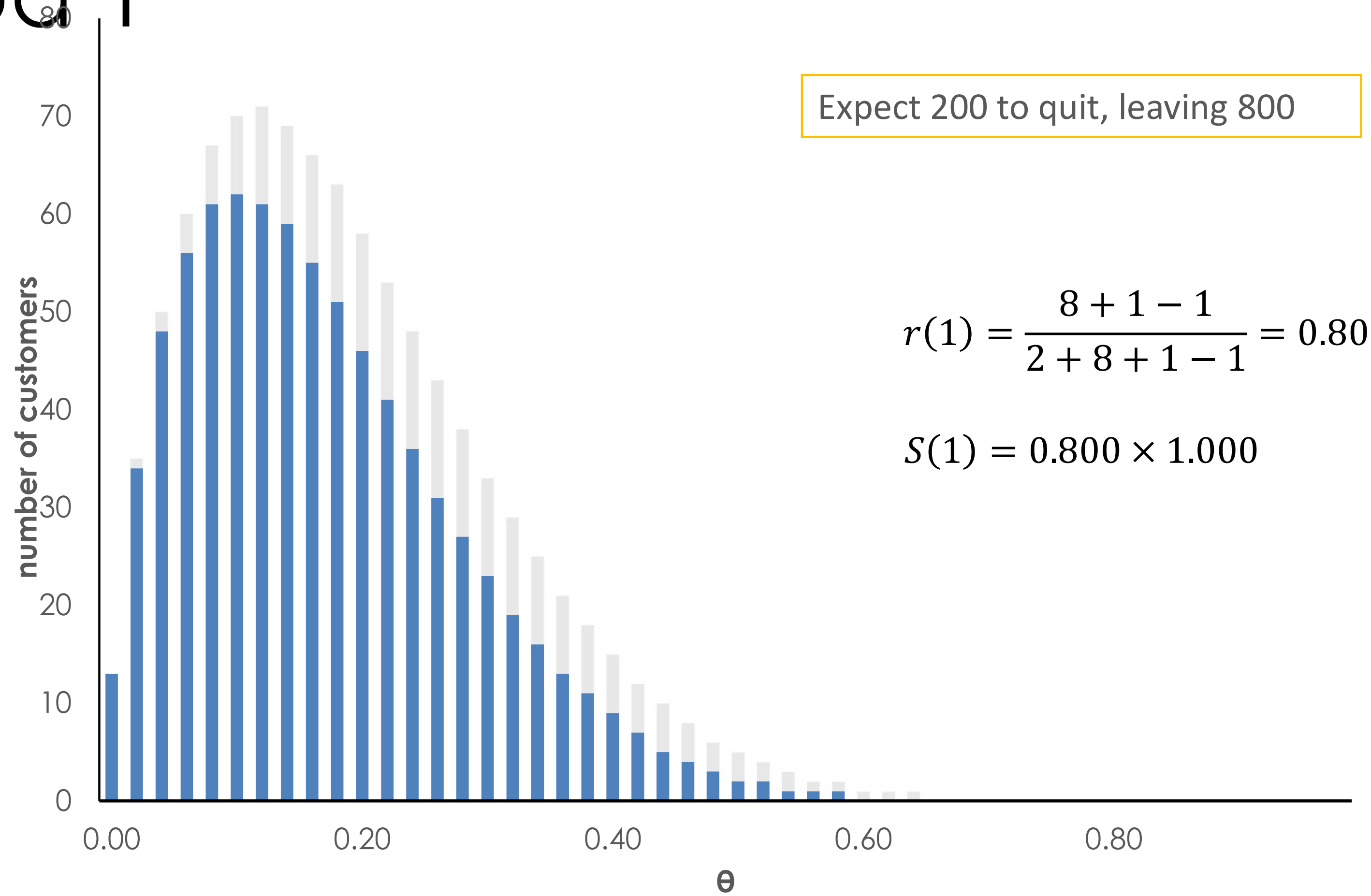
1000 customers acquired



1000 draws from a Beta distribution where $a = 2, b = 8$

Where does this distribution come from? (appendix)

Period 1



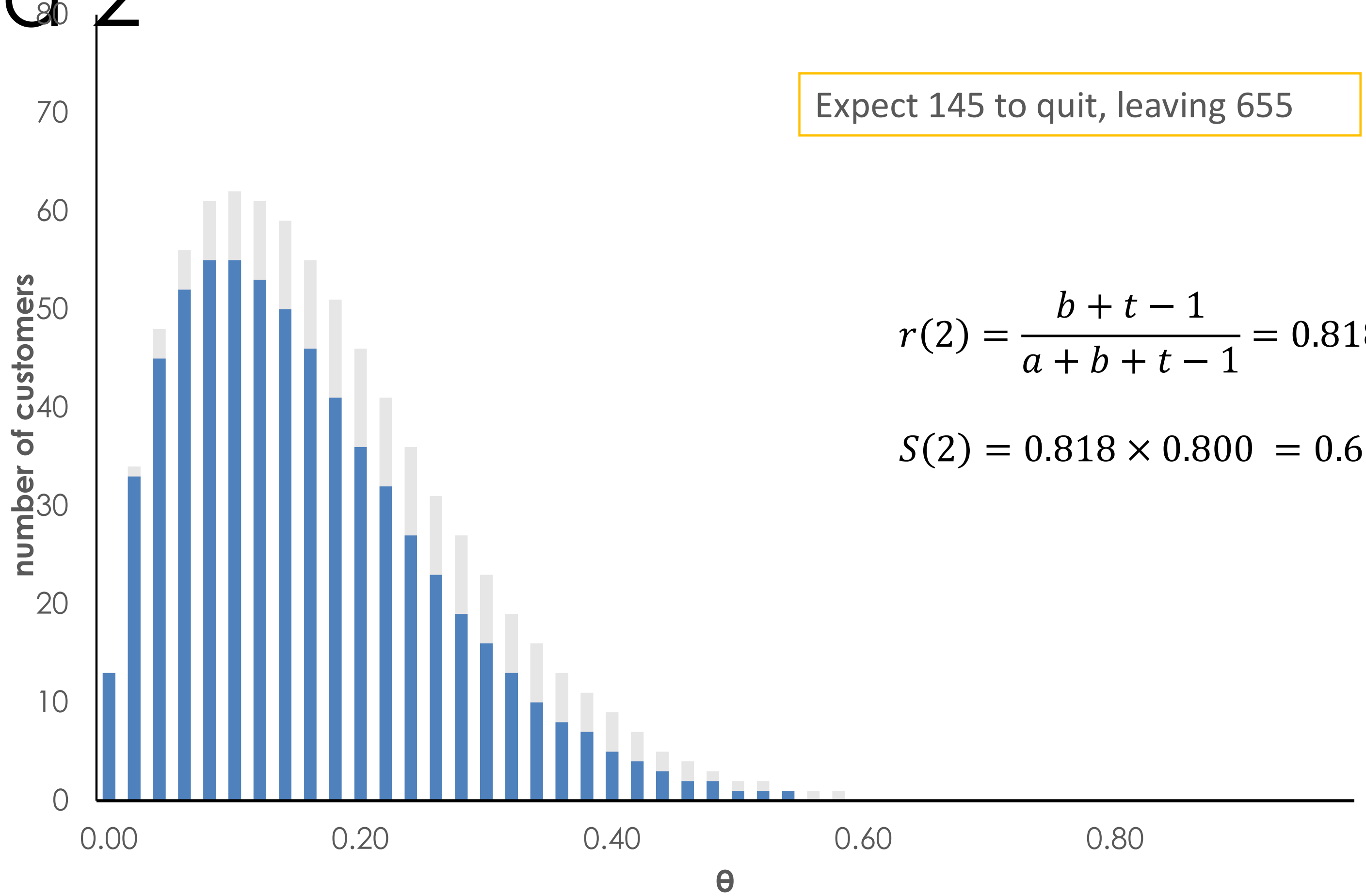
Expect 200 to quit, leaving 800

$$r(1) = \frac{8+1-1}{2+8+1-1} = 0.800$$

$$S(1) = 0.800 \times 1.000$$

800 draws from a Beta distribution where $a^* = 2, b^* = b + 1 = 9$

Period 2



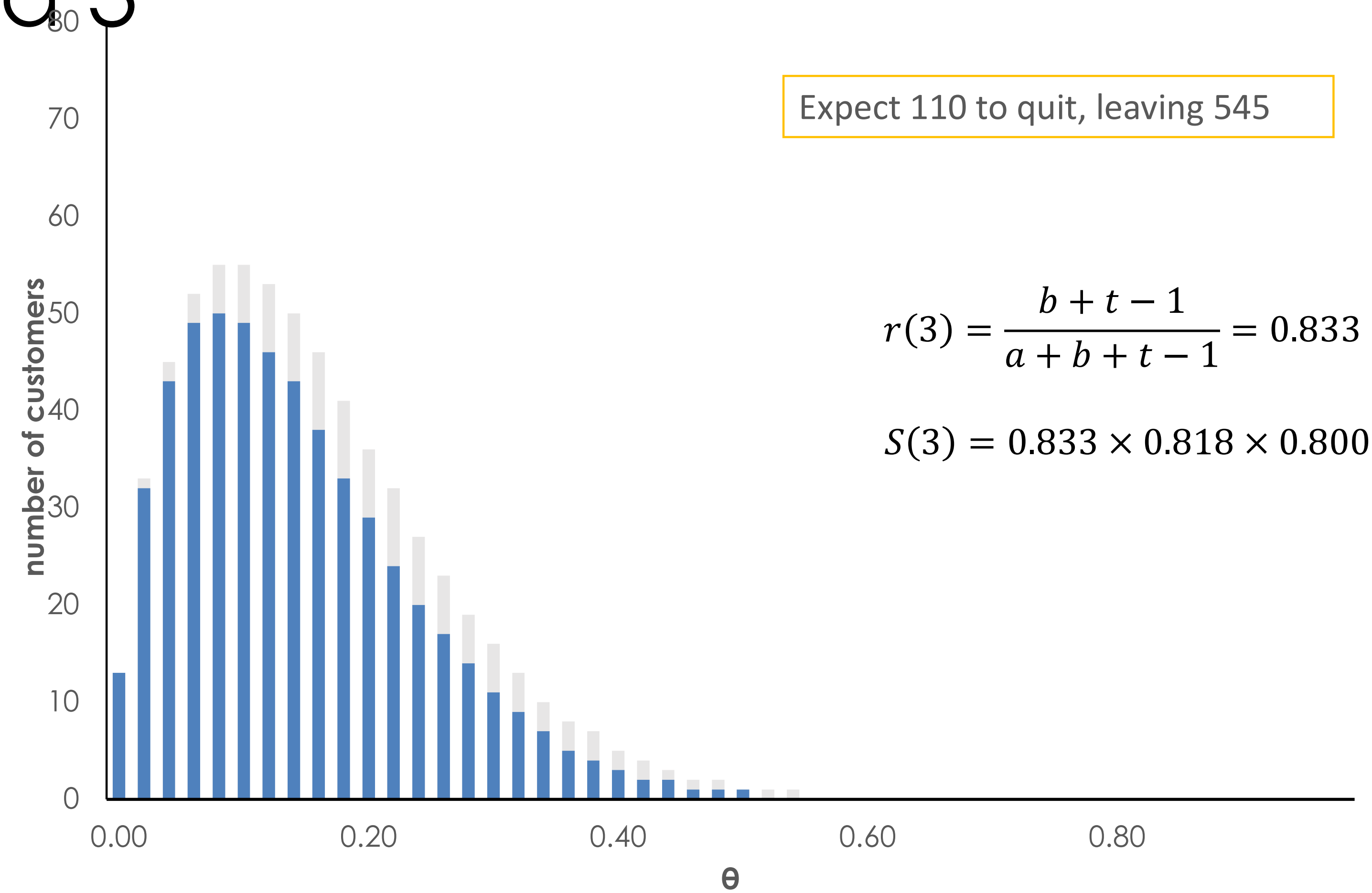
Expect 145 to quit, leaving 655

$$r(2) = \frac{b+t-1}{a+b+t-1} = 0.818$$

$$S(2) = 0.818 \times 0.800 \ = 0.655$$

655 draws from a Beta distribution where $a^* = 2, b^* = b + 2 = 10$

Period 3



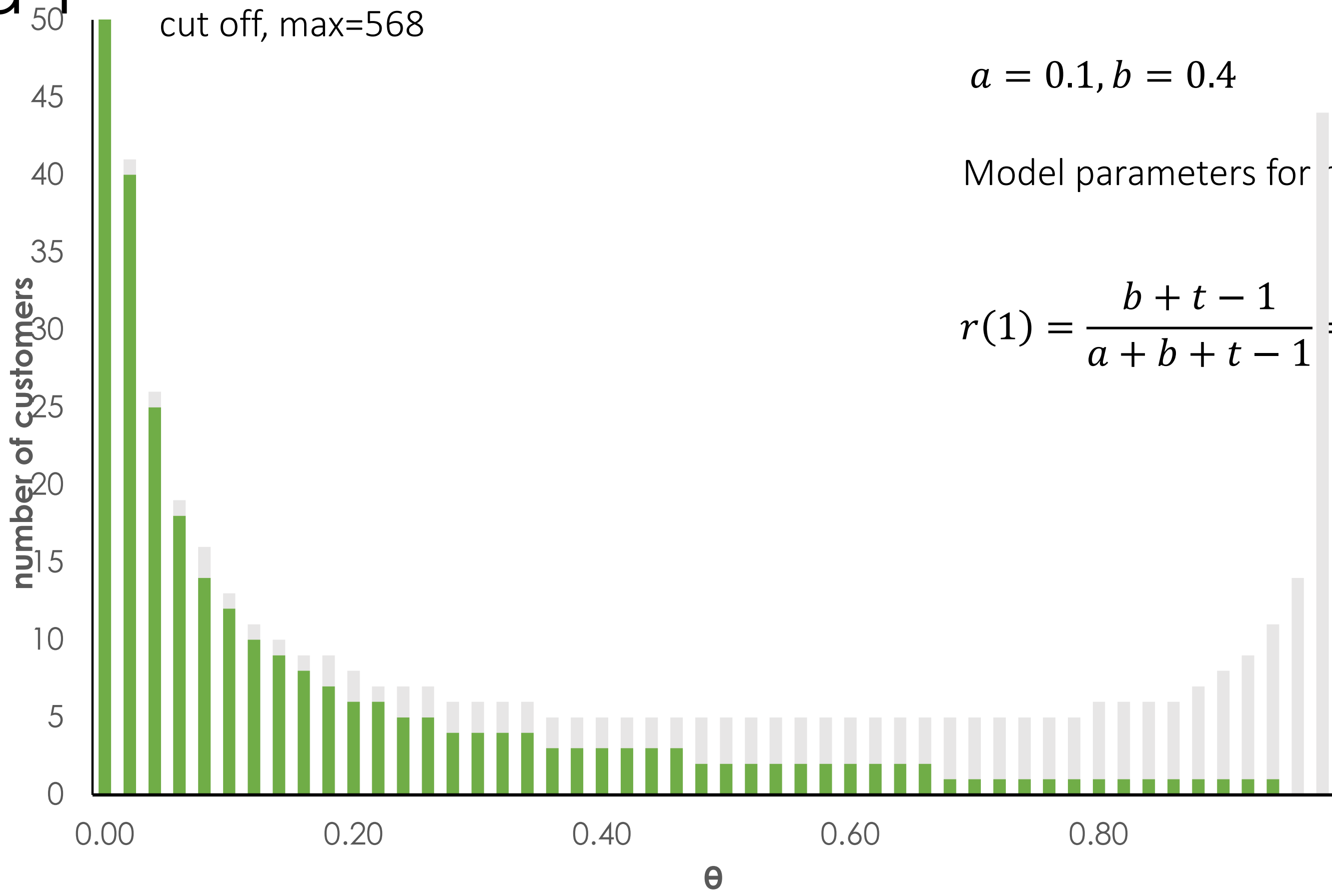
545 draws from a Beta distribution where $a^* = 2, b^* = b + 3 = 10$

Expect 110 to quit, leaving 545

$$r(3) = \frac{b + t - 1}{a + b + t - 1} = 0.833$$

$$S(3) = 0.833 \times 0.818 \times 0.800 \ = 0.545$$

Compare to more heterogeneous customer base, period 1



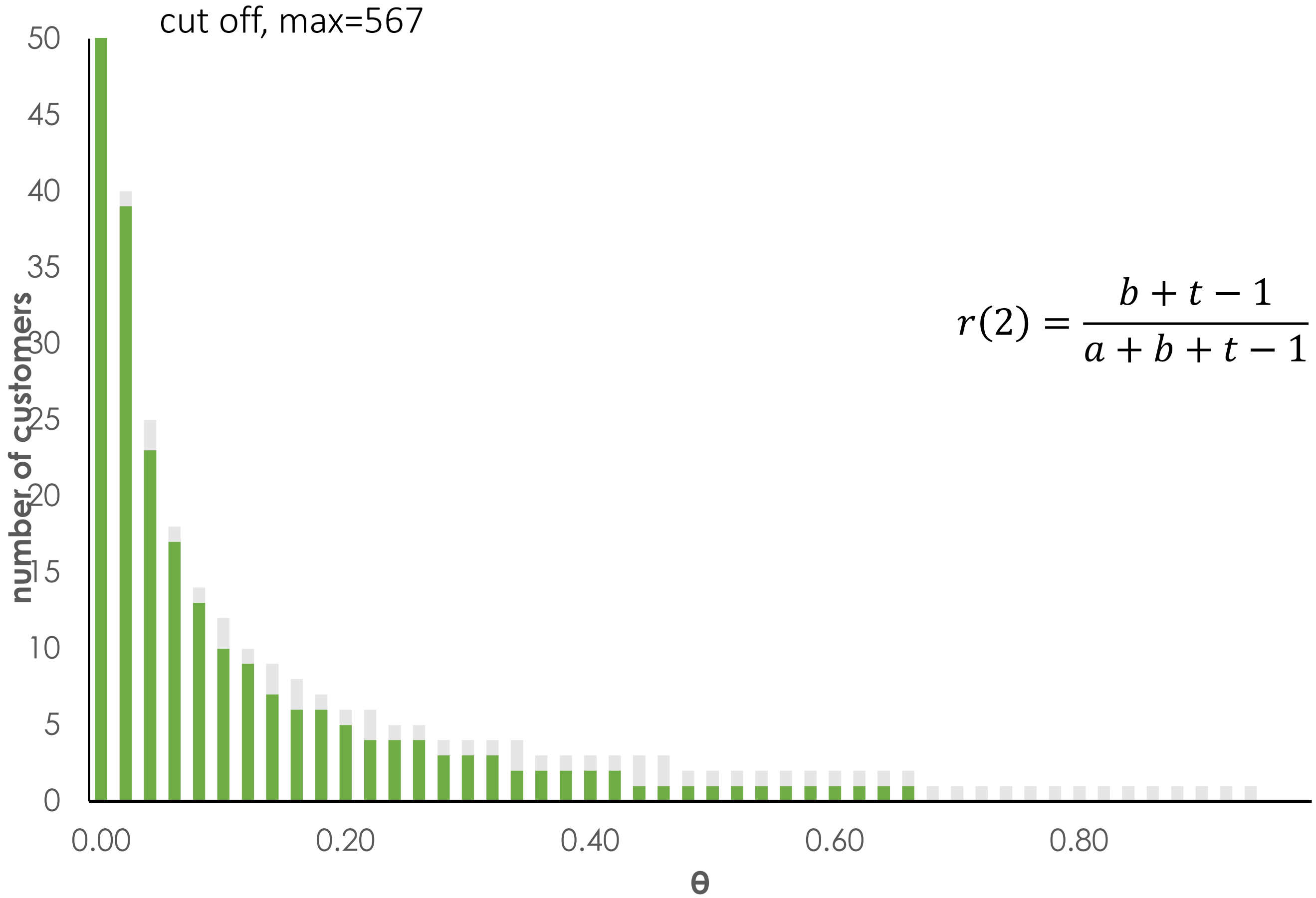
$a = 0.1, b = 0.4$

Model parameters for high heterogeneity

$$r(1) = \frac{b+t-1}{a+b+t-1} = 0.800$$

800 draws from a Beta distribution where $a^* = 0.1, b^* = 0.1 + 1 = 1.1$

period 2



$$r(2) = \frac{b+t-1}{a+b+t-1} = 0.933$$

747 draws from a Beta distribution where $a^* = 0.1, b^* = 0.1 + 2 = 2.1$

Concepts

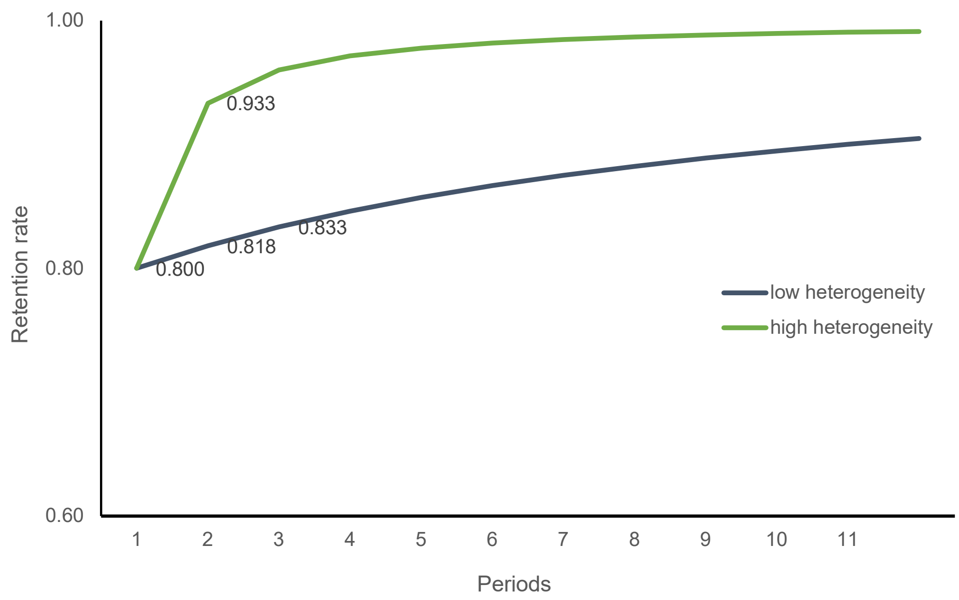
Customer heterogeneity

Retention rates slope upward even if customers propensity to churn stays the same over time.

Why?

- High-churn customers drop out early
  - Sorting effect in a heterogeneous population
- Remaining customers have lower churn probabilities
  - Happens more quickly the more heterogeneity there is
- Ignoring this will bias your CLV estimates downwards

Retention rate dynamics



CLV with the BG model

from before:

$$E[CLV] = m + \frac{m\,S(1)}{1+d} + \frac{m\,S(2)}{(1+d)^2} + \frac{m\,S(3)}{(1+d)^3} + \cdots$$

in the geometric model, $S(t) = p^t$.

- Formula for infinite series available: simple expression (here)

in the shifted Beta-geometric model, $S(t) = B(a, b+t)/B(a, b)$

- Formula for infinite series unavailable: sum up the first N terms, until contribution is small.

RLV under the BG model

Sorting population also means that there is a difference in a customer's lifetime value depending on how long they have been a customer.

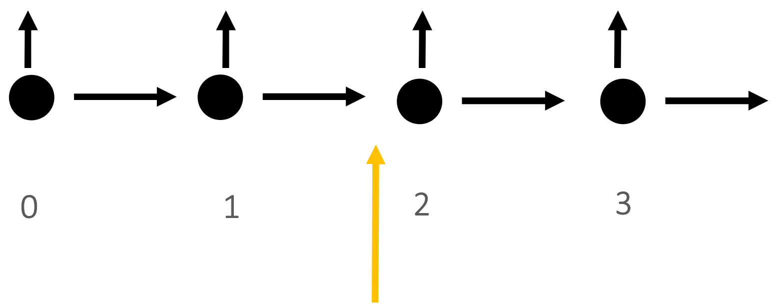
- How likely will a customer who has been with us for two periods (renewed once) renew at least a second time?

$$P(T > 2 | \, T > 1) = \frac{P(T > 2)}{P(T > 1)} = \frac{S(2)}{S(1)}$$

A third time?

$$P(T > 3 | \, T > 1) = \frac{S(3)}{S(1)}$$

RLV



Residual lifetime value: an existing customer for 2 periods, what is expected future discounted profits right before the renewal decision?

$$E[RLV] = m\frac{S(2)}{S(1)} + \frac{m}{1+d}\frac{S(3)}{S(1)} + \frac{m}{(1+d)^2}\frac{S(4)}{S(1)} + \cdots$$

Customer Lifetime Value

- CLV is a forward looking metric: different models give different results
  - Geometric model: constant retention over time
  - Shifted Beta-geometric: increasing
- Dynamic patterns can arise from simple heterogeneity
- Not accounting for retention rate increases will bias CLV and RLV calculations downward.

Appendix

Aside: beta function

Formally, defined by this integral:

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} \, dx$$

where a, b > 0. we can write the beta function in terms of gamma functions:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

The gamma function $\Gamma(a)$ is a generalized factorial, which has the recursive property:

$$\Gamma(x+1) = x\,\Gamma(x)$$

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

$$\Gamma(4) = 3! = 6$$

Bayes Rule: updating the distribution of θ

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

$$f(\theta|T > t) = \frac{S(t|\theta)f(\theta|a,b)}{\int S(t|\theta)f(\theta|a,b)\,d\theta}$$

$$= \frac{(1-\theta)^{t+b-1}\theta^{a-1}/B(a,b)}{B(a,b+t)/B(a,b)}$$

$$= \frac{\theta^{a-1}(1-\theta)^{t+b-1}}{B(a,b+t)}$$

This is just a beta distribution with $a^* = a$ and $b^* = b + t$