Introduction to CLV

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

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

Revenues less costs of marketing, selling, production, servicing

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

Measured for individual customers; recognize every customer is different

Uses of CLV

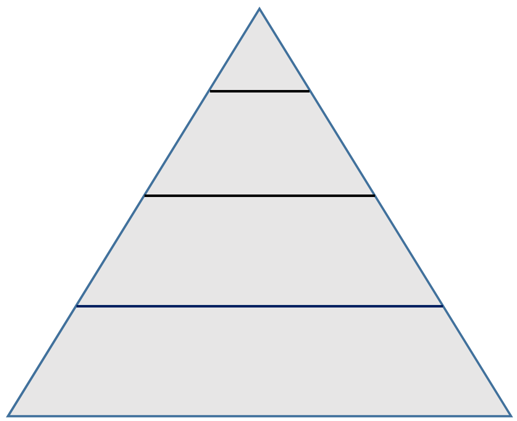
- Tell you what each of your individual customers is worth (useful beyond other measures like NPS, clicks, etc)
  - Byproduct: you can use it to predict purchasing, benchmark with marketing efforts

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

- 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 CLV

Gold (15%): 30% of CLV

Silver (30%): 20% of CLV

Lead (50%): 0% of CLV

CLV primitives

(Assume all these things are the same each period):

1. Margins (accounting)

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

2. Discount rate (finance)

   d = opportunity cost of capital

3. Retention (marketing)

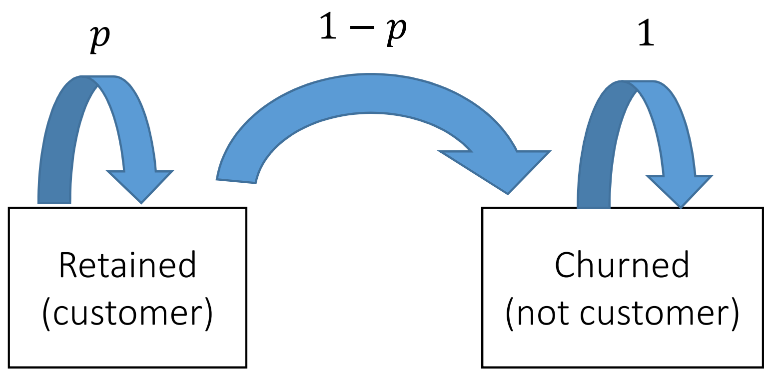
   Geometric model

   $r(t)$= P(customer **next** period | customer this period) $\boxed{= p}$

   "hazard rate"

Key source of uncertainty in CLV models: how long will customer remain?

Retention as a two-state Markov chain



Next customer state

Previous customer state		Retained	Churned
	Retained	p	$1-p$
	Churned	0	1

Random variable T is the amount of time spent as customer (in retained state)

$$r(t) = P(T > t \mid T > t-1) = p$$

Geometric distribution that starts at 1. The survival function is the probability that a customer lasts longer than t

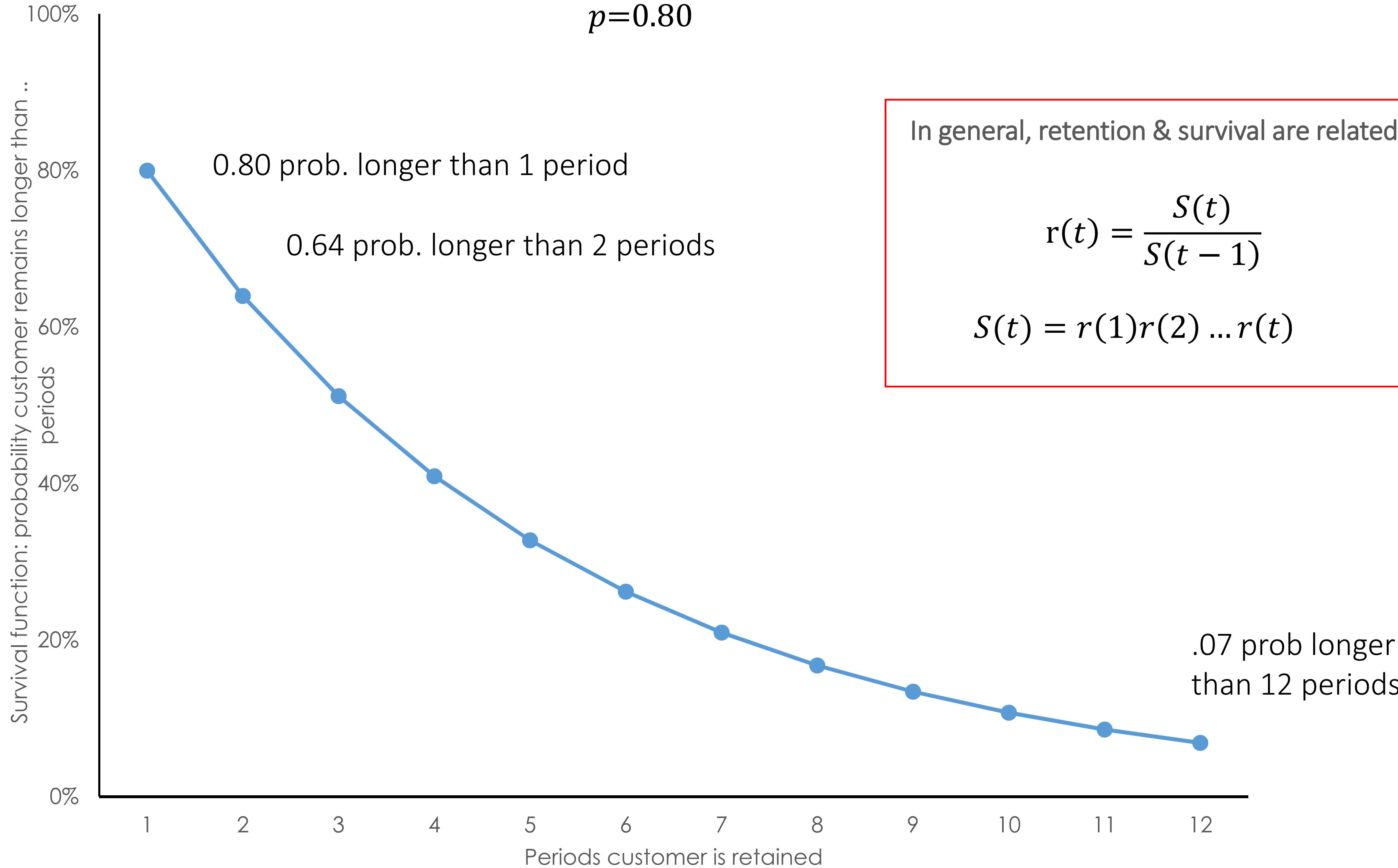
$$P(T > t) = S(t) = p^t \quad t = 1,2,\ldots$$

Survival function

$S(t) = p^t$

Retention rate

p=0.80



In general, retention & survival are related

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

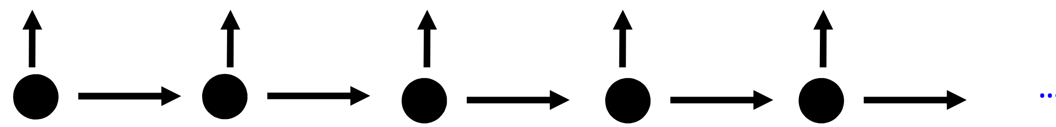
$$S(t) = r(1)r(2)\ldots r(t)$$

CLV traditional approach: the formula

A x B x C

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

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

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

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

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

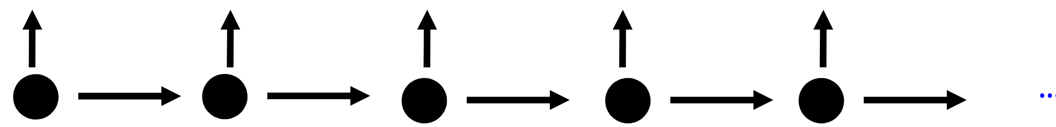
Geometric series

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

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

Sum of first 10 terms vs. infinite horizon formula

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

0	100
1	73
2	53
3	38
4	28
5	20
6	15
7	11
8	8
9	6
10	4

Each subsequent term is smaller

356 vs. 367 (from formula)

CLV vs. residual lifetime value (RLV)

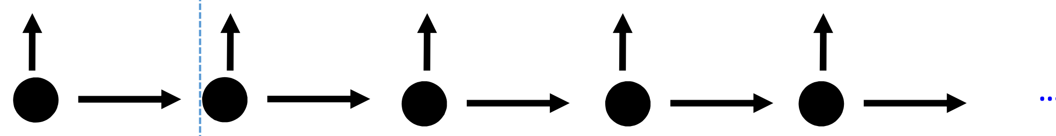
CLV = for a not-yet-acquired customer, <u>from first purchase</u>

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

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



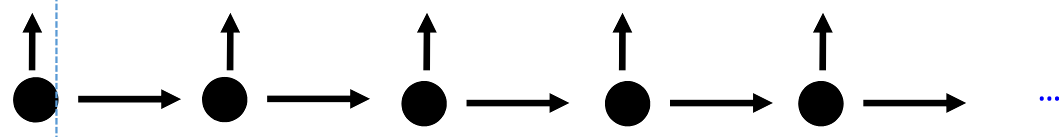
Time (t is present)	t-1	t*	t+1	t+2	t+3
	history		future		

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

RLV: right after renewal

present

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

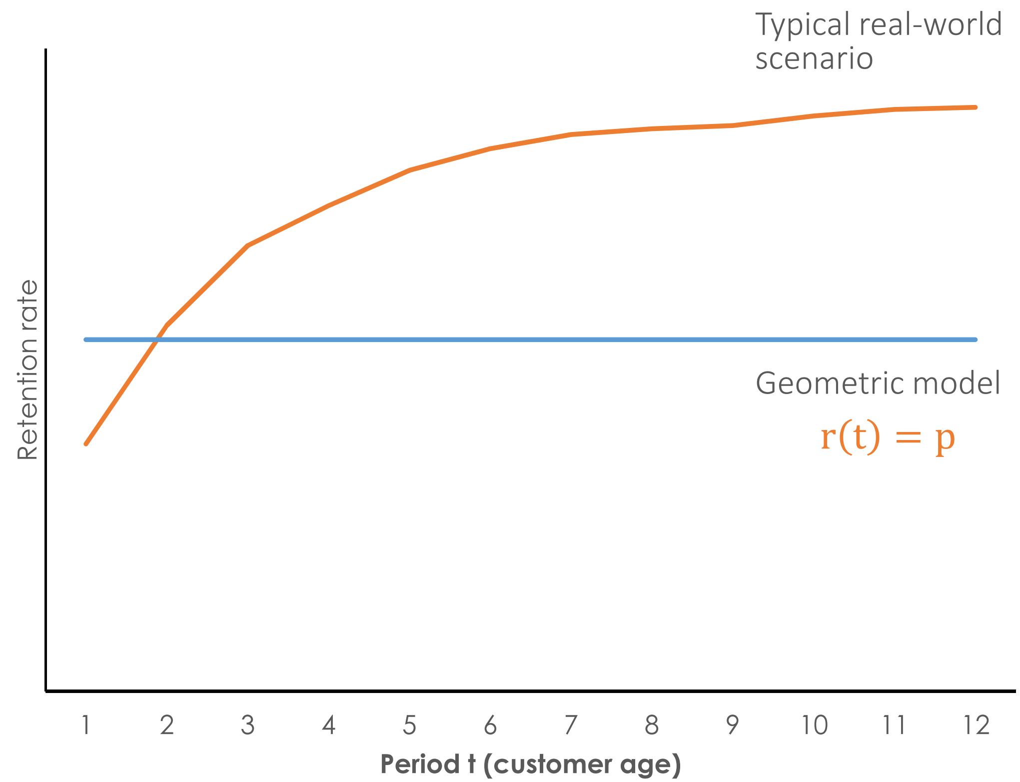
Problem: retention rates increase

The geometric model is typically **not** a good description of reality

Why?

Retention rates are **not** constant: often increase the longer the customer has remained a customer ("customer age")

Hence basing CLV & RLV calculations on constant retention will bias results



An example

RLV

Standing at the end of 2018, what is the **residual lifetime value** of the 26569 customers remaining?
Assume contracts last a year

- Profit per year = 100
- Discount rate = 10%
- How do we measure retention probability?

Retention rate = *ratio of the number of retained customers to the number at risk*

Number of customers by year of acquisition

2014	2015	2016	2017	2018
10000	6334	4367	3264	2604
	10000	6334	4367	3264
		10000	6334	4367
			10000	6334
				10000
10000	16334	20701	23965	26569

$$\bar{p} = \frac{2604 + 3264 + 4367 + 6334}{3264 + 4367 + 6334 + 10000} = 0.691$$

plug into RLV formula

$$26569 * \frac{m\, p\, (1 + d)}{1 + d - p} = 4945049$$

Issue #2: customer heterogeneity

Problem #1

Each cohort has a different retention rate, with earlier cohorts having larger retention rate:

Retention rate in 2018

Customers who started in 2014 $p_1 = \frac{2604}{3264} = 0.798$

$p_2 = \frac{3264}{4367} = 0.747$

$p_3 = \frac{4367}{6334} = 0.689$

Customers who started in 2017 $p_4 = \frac{6334}{10000} = 0.633$

$\bar{p} = 0.691$

Number of customers by year of acquisition

2014	2015	2016	2017	2018
10000	6334	4367	3264	2604
	10000	6334	4367	3264
		10000	6334	4367
			10000	6334
				10000
10000	16334	20701	23965	26569

Fader and Hardie (2010), Customer Base Valuation in a Contractual Setting

Issue #2: customer heterogeneity

Problem #2

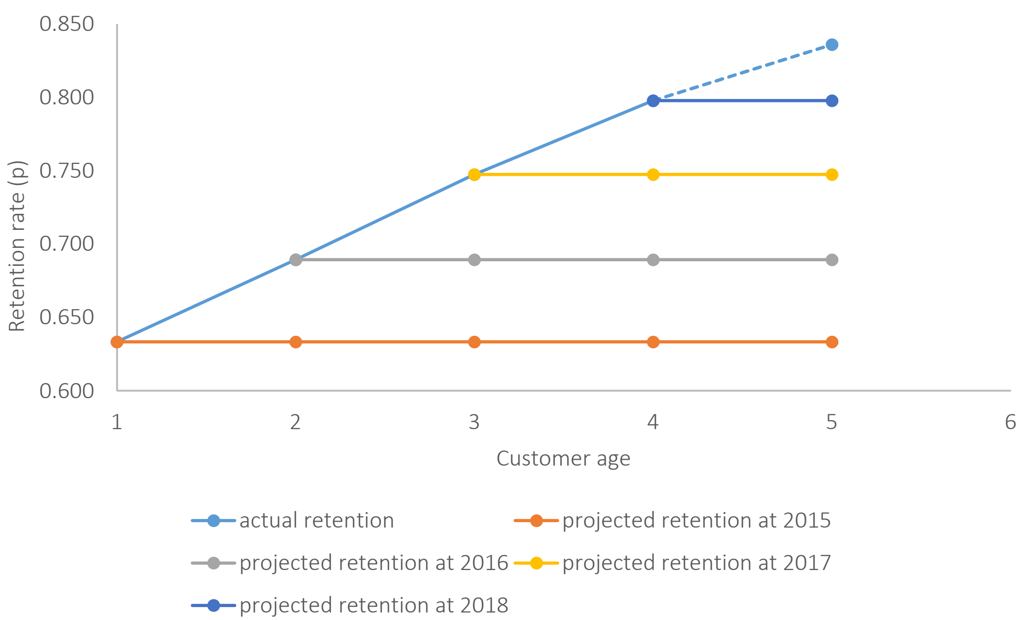
The retention rate is increasing with customer age, commonly found phenomenon

(why do you think that happens?)

In the geometric model, $\mathrm{r}(t) = p$ is the same over time.

Both these problems mean that the traditional CLV calculations are biased and actually undervalue customers.

Retention rate by year of acquisition



Fader and Hardie (2010), Customer Base Valuation in a Contractual Setting

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

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

The retention rate does not depend on t.

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, 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

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

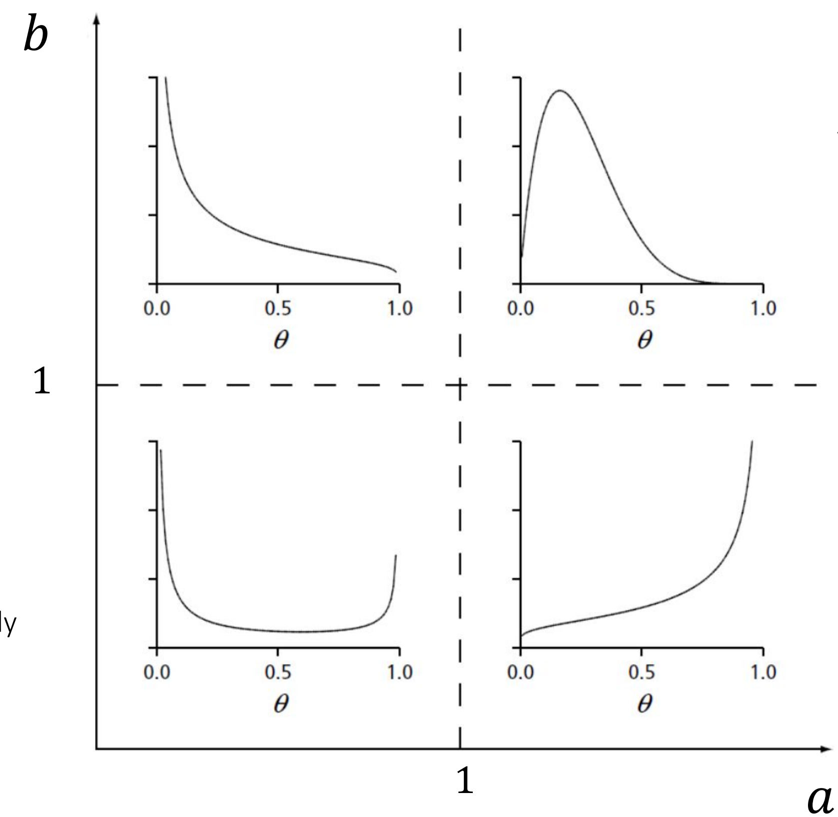
Fader and Hardie (2007), How to Project Customer Retention

General shapes of the beta distribution

Spread out across whole range, but most likely to stay

Most customers fall within a certain range

Remember:
θ is the probability of churn.
These describe possible distributions of this probability among customers



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

Shifted geo. dist. Beta dist.

$$S(t|a,b) = \int_0^1 S(t|\theta)\, f(\theta|a,b)\, d\theta$$

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

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

Retention rate in shifted beta-geometric model

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

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

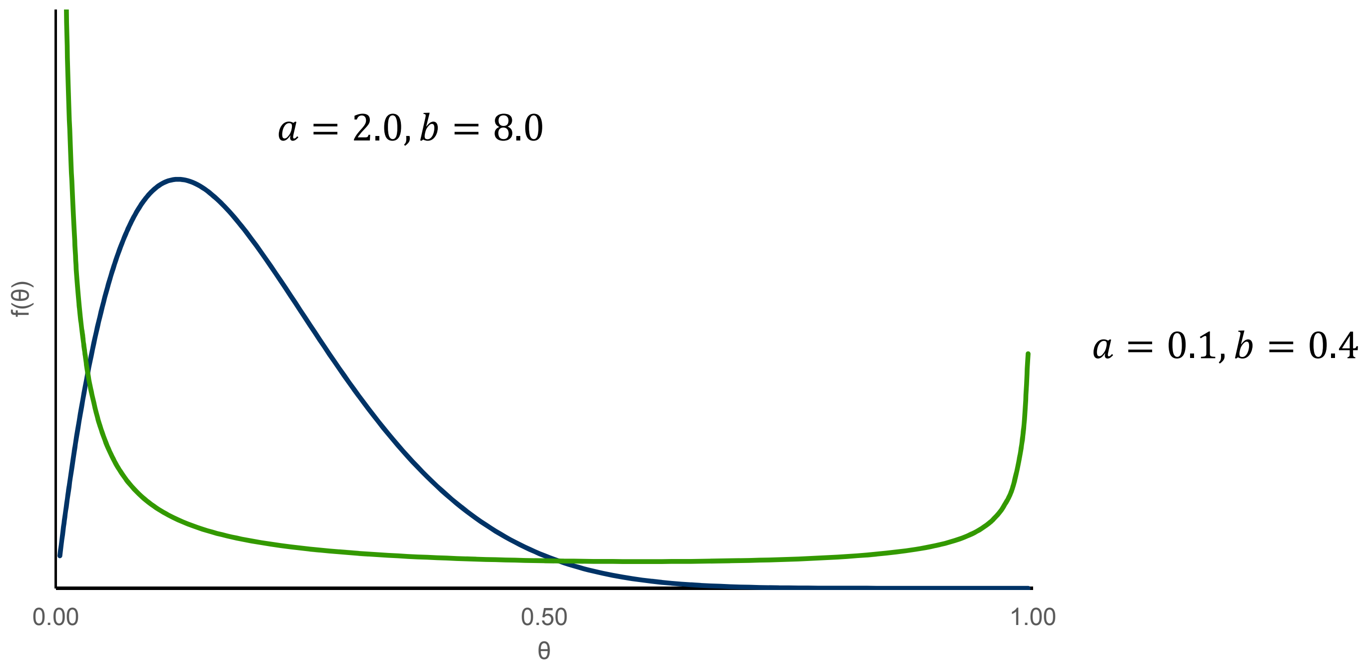
$$= \frac{b+t-1}{a+b+t-1}$$

Does retention increase with t?

How fast?

Two different groups

Two populations with the same mean, but different variance



Beta distribution

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

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

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

Implications for retention rate

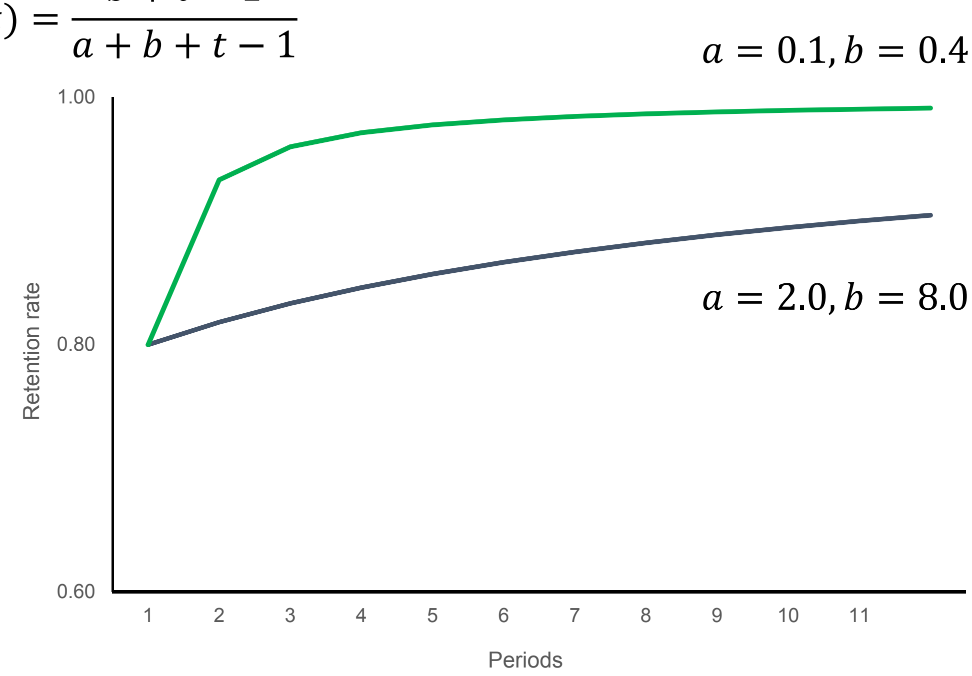
Customer heterogeneity

- 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
  $a = 2.0, b = 8.0$

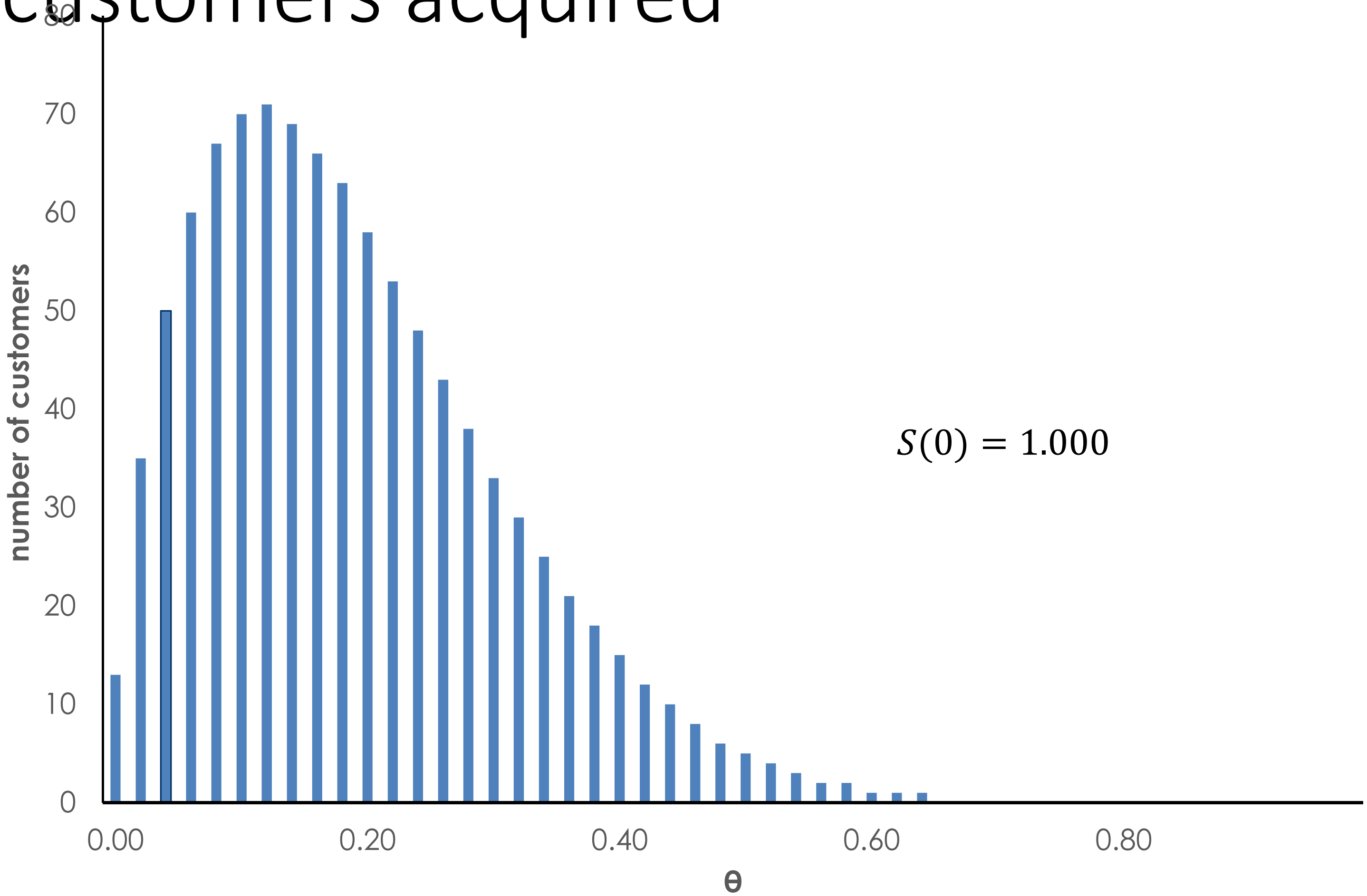
What's the intuition? (next slides)

Retention rate dynamics

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

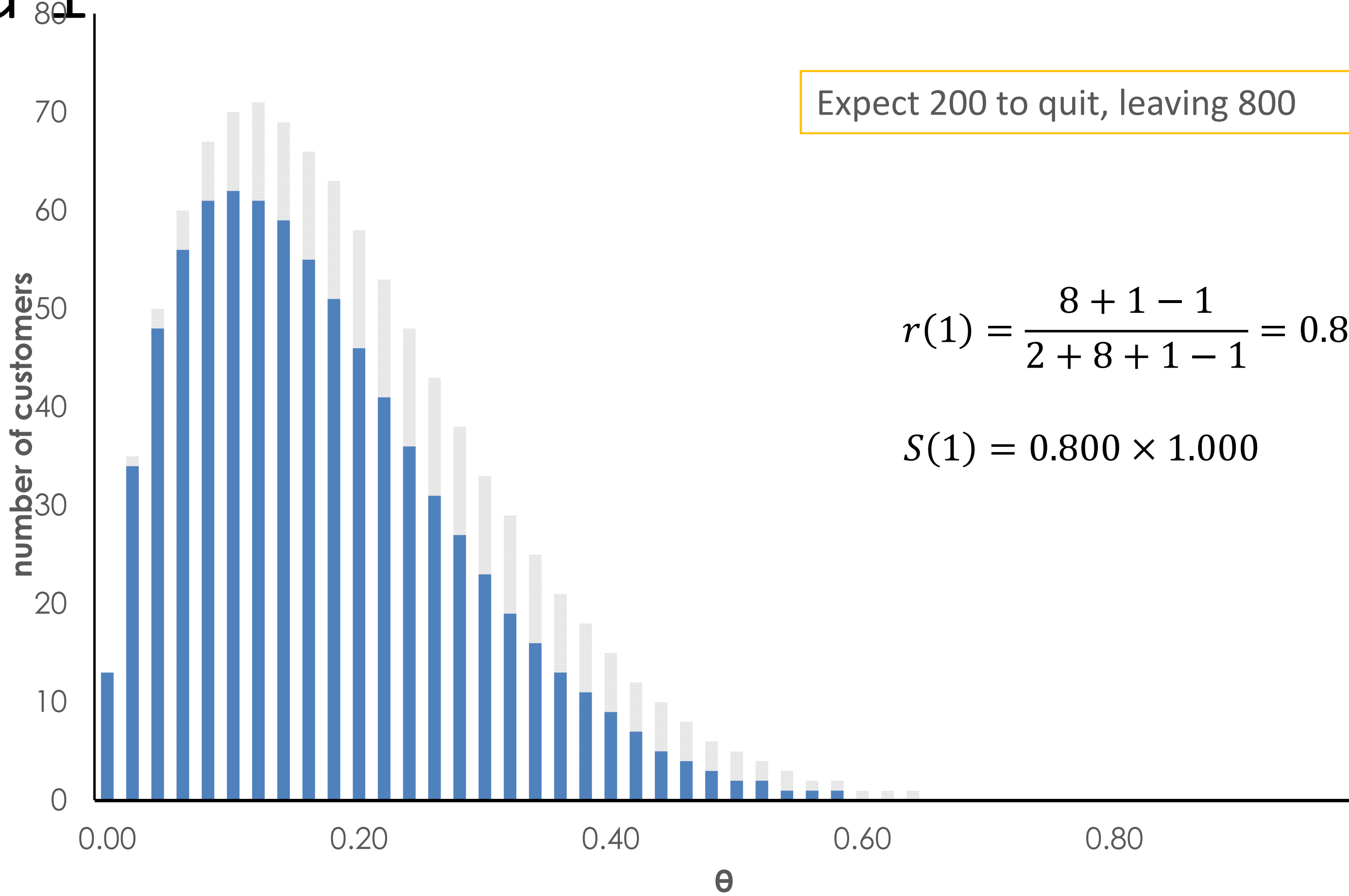


1000 customers acquired



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

Period 1



Expect 200 to quit, leaving 800

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

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

Where does this distribution come from? (next slide)

Bayes Rule: updating the distribution of $\boldsymbol{\theta}$

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

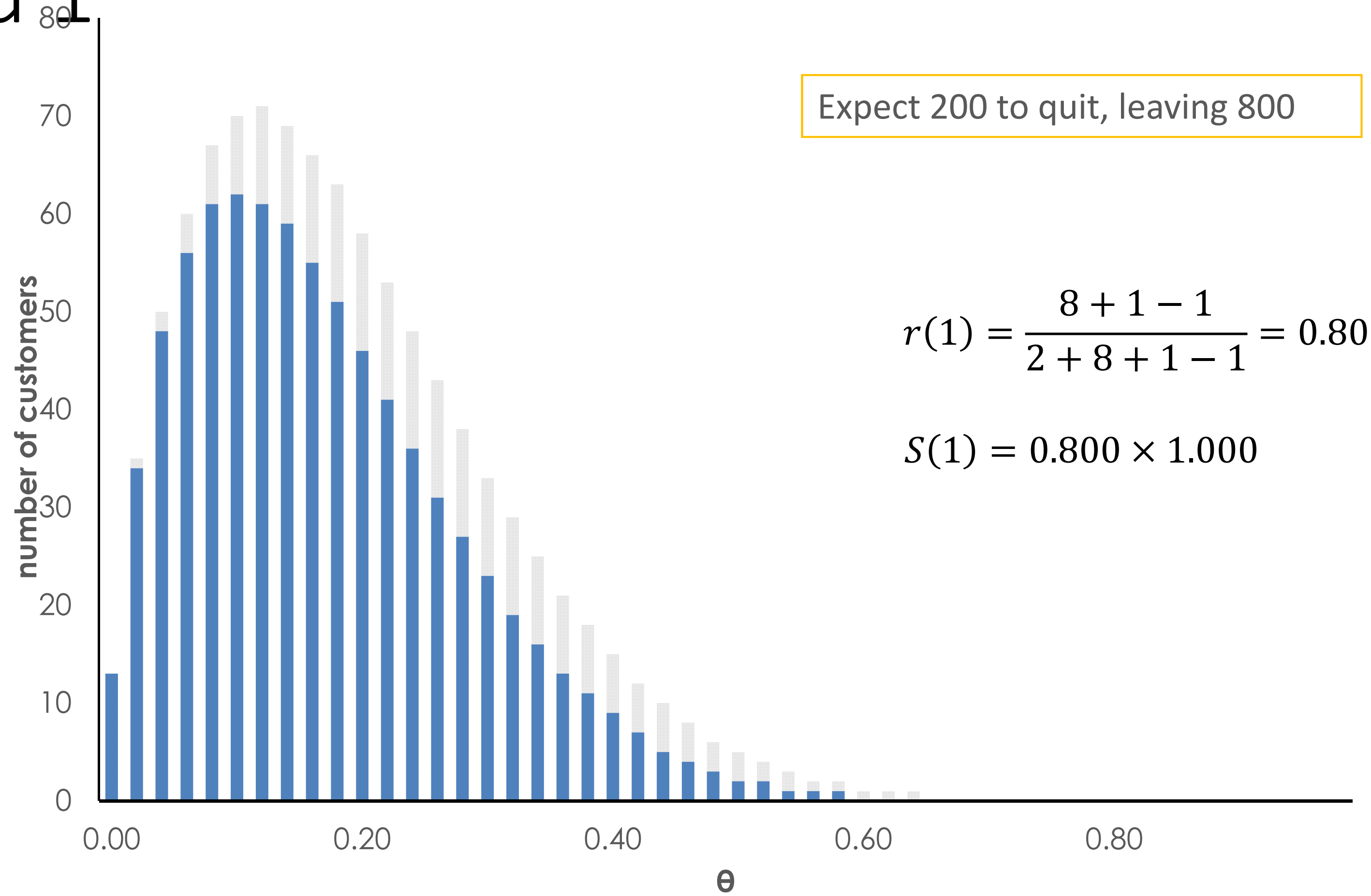
$$f(\theta|T > t) = \frac{S(t|\theta)f(\theta|a,b)}{\int S(t|\theta)f(\theta|a,b)\,d\theta}$$

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

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

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

Period 1



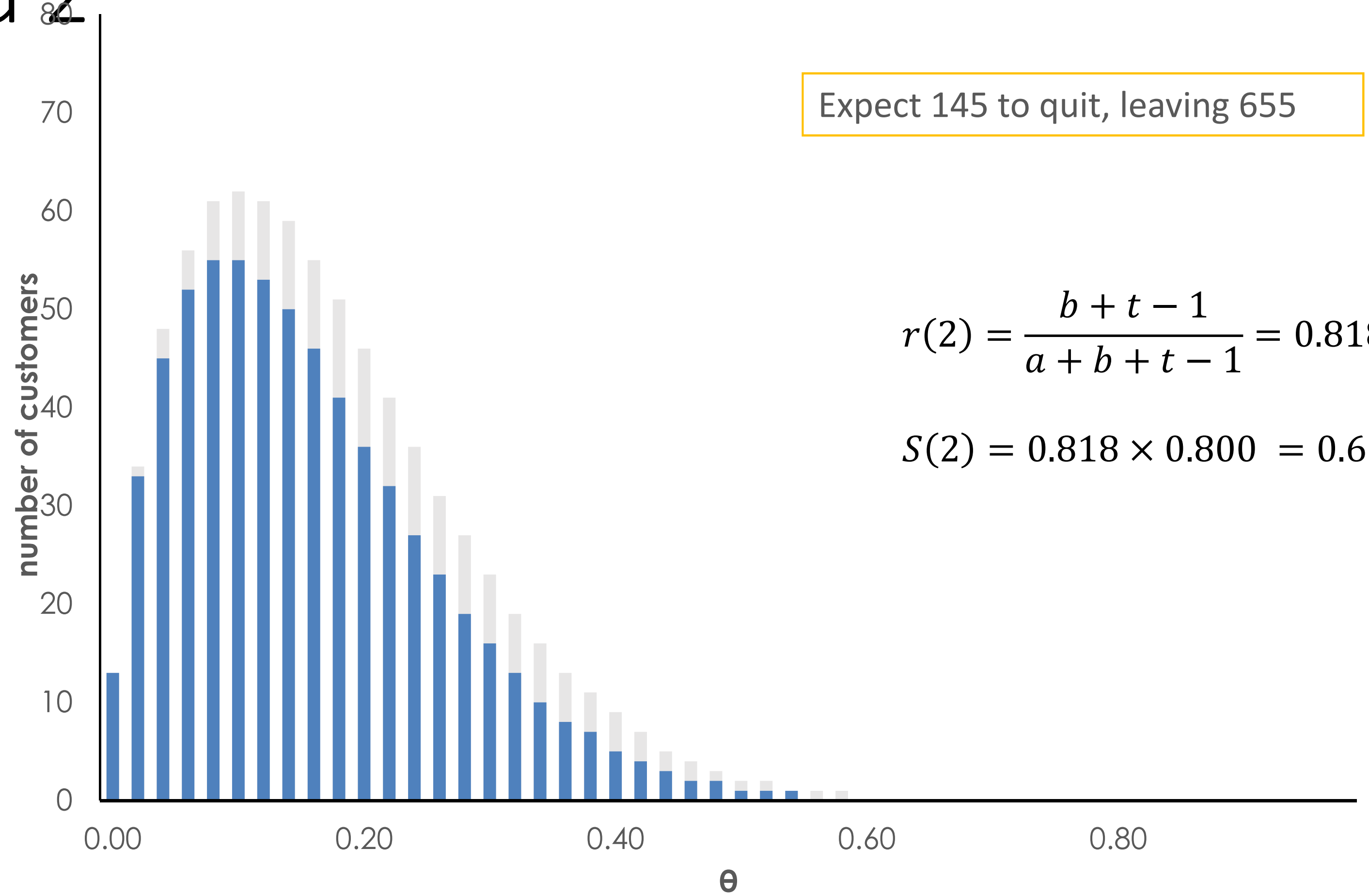
Expect 200 to quit, leaving 800

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

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

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

Period 2



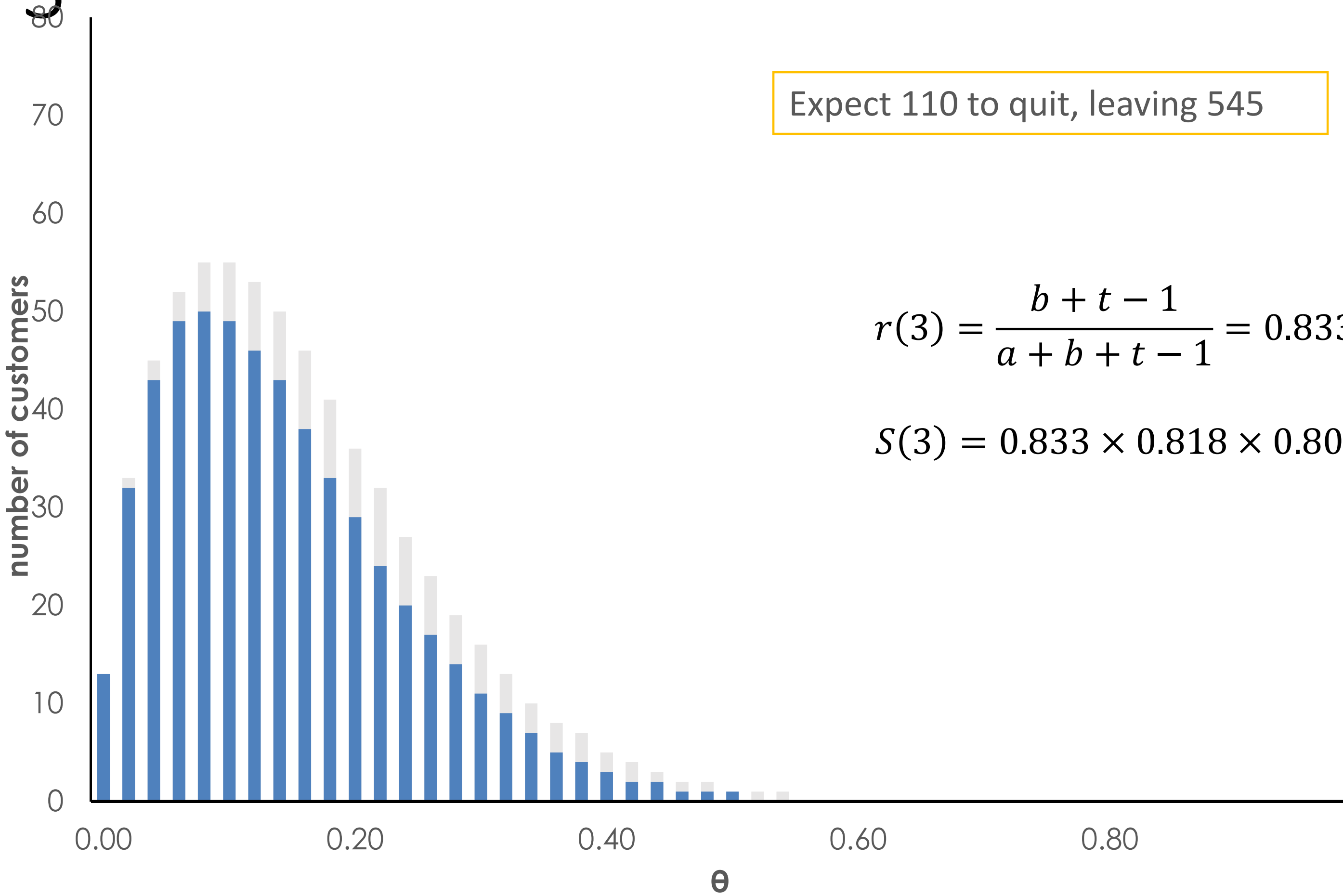
Expect 145 to quit, leaving 655

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

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

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

Period 3



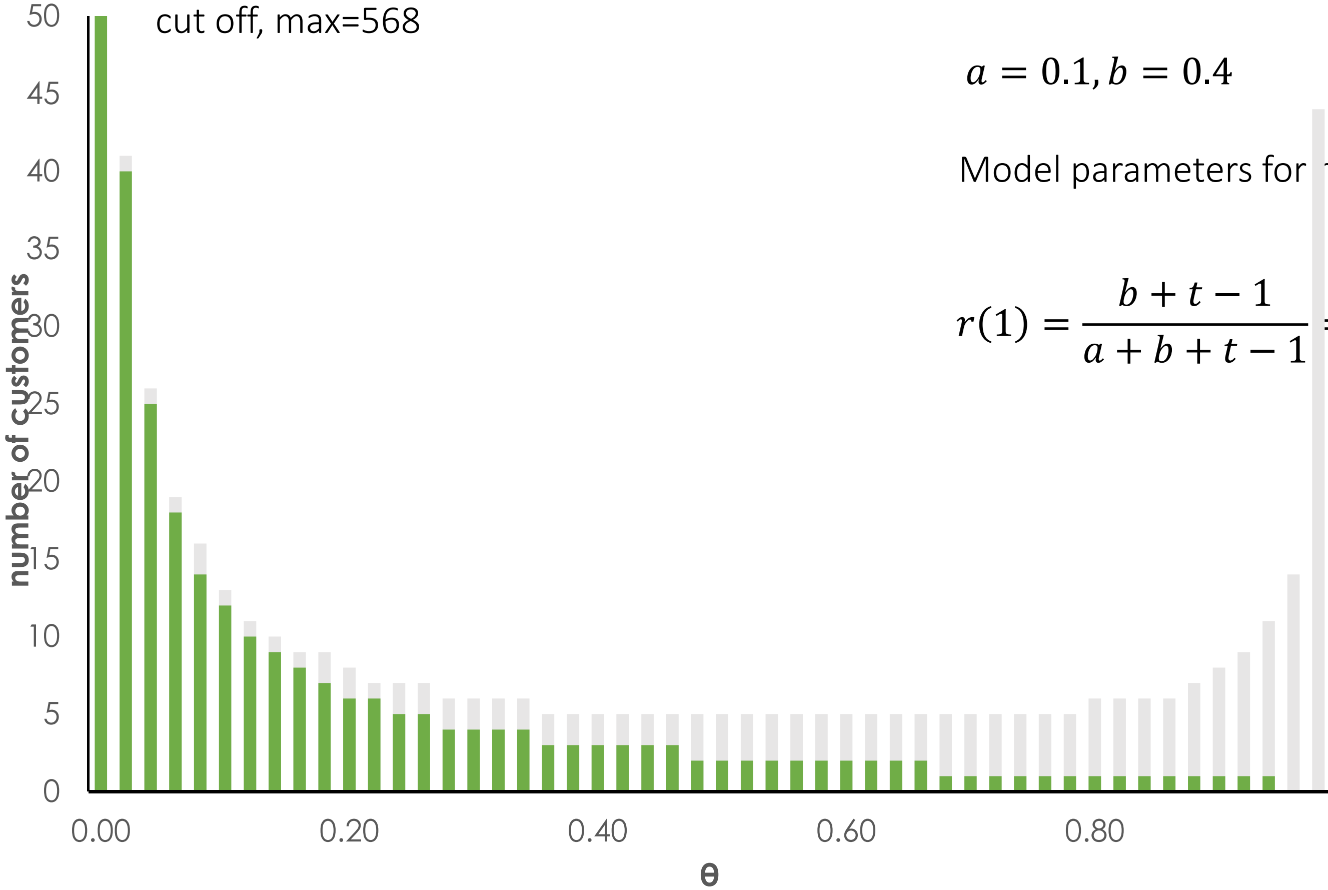
Expect 110 to quit, leaving 545

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

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

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

Compare to more heterogeneous customer base, period 1



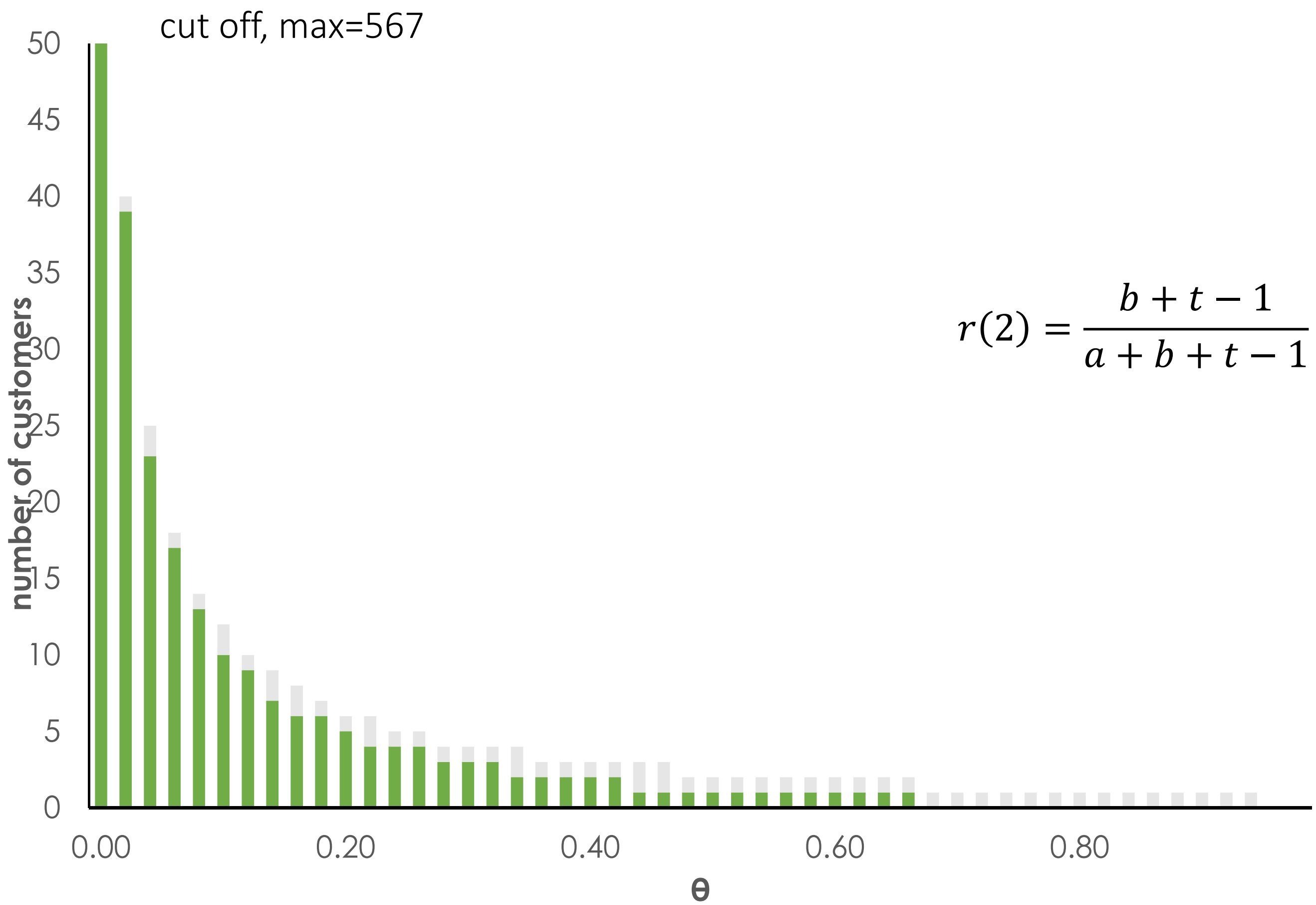
$a = 0.1, b = 0.4$

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

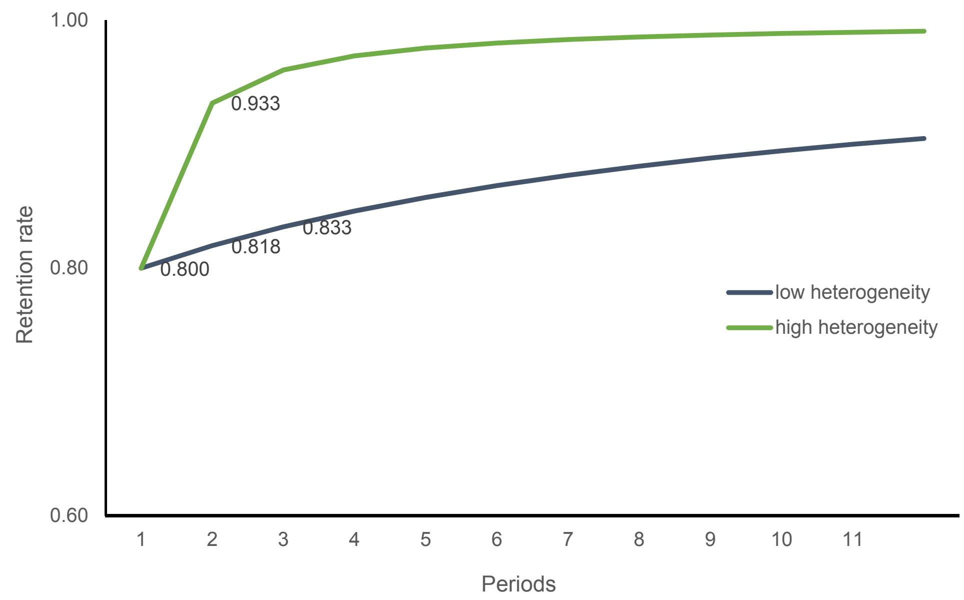
Customer heterogeneity

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



Estimation

- This method is called "(parametric) Empirical Bayes"
  - We use a (parametric) prior distribution and Bayes theorem to update it
  - Empirical because we estimate the parameters of the prior distribution

- How do you estimate the a & b parameters of the (shifted) Beta-geometric model?
  - Data on retention or survival over time
  - Estimating the parameters of the BG model on go.uvt.nl/customer-analytics
  - You can use maximum likelihood based on the survival or (in the video) minimize least squares

- Once you have them a & b estimated, you can calculate CLV and RLV

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. (see how the distribution changes here)

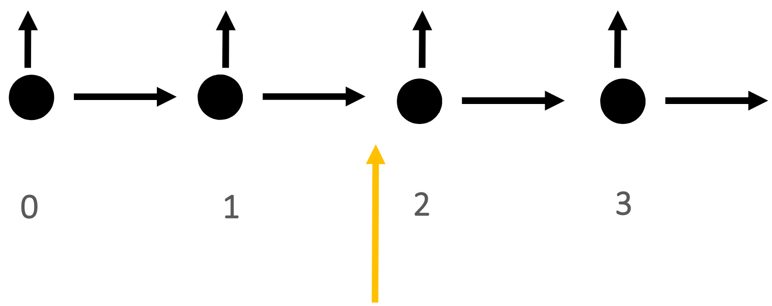
- A customer of t^* years is likely to have a higher retention rate than a customer of $t < t^*$ years

$$r(t^*) \geq r(t^* - 1) \geq \cdots r(2) \geq r(1)$$

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

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

Residual lifetime value



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

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

How do you create a spreadsheet to calculate CLV/RLV?

CLV & RLV with BG video

Customer Lifetime Value

- Future value of your customers
- Individual level
- Incorporate customer heterogeneity in the model
- Retention rate increases

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.

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