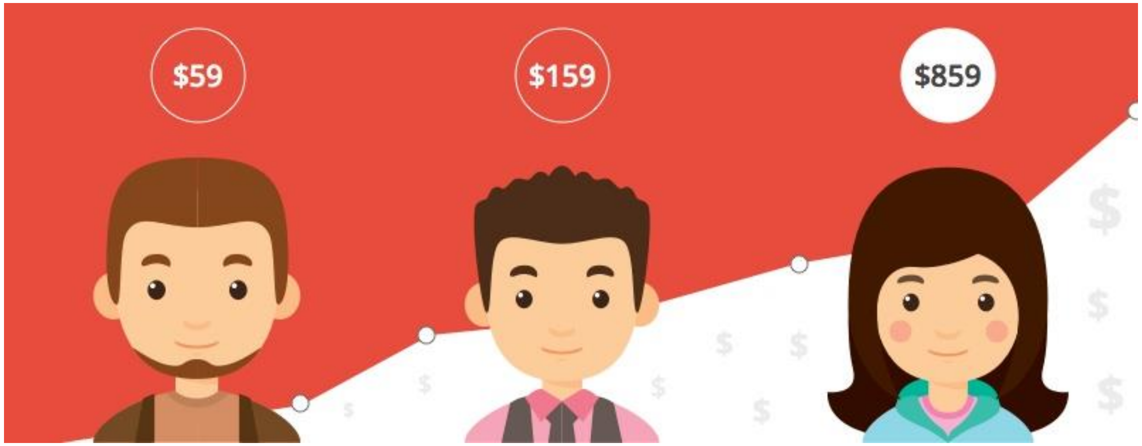


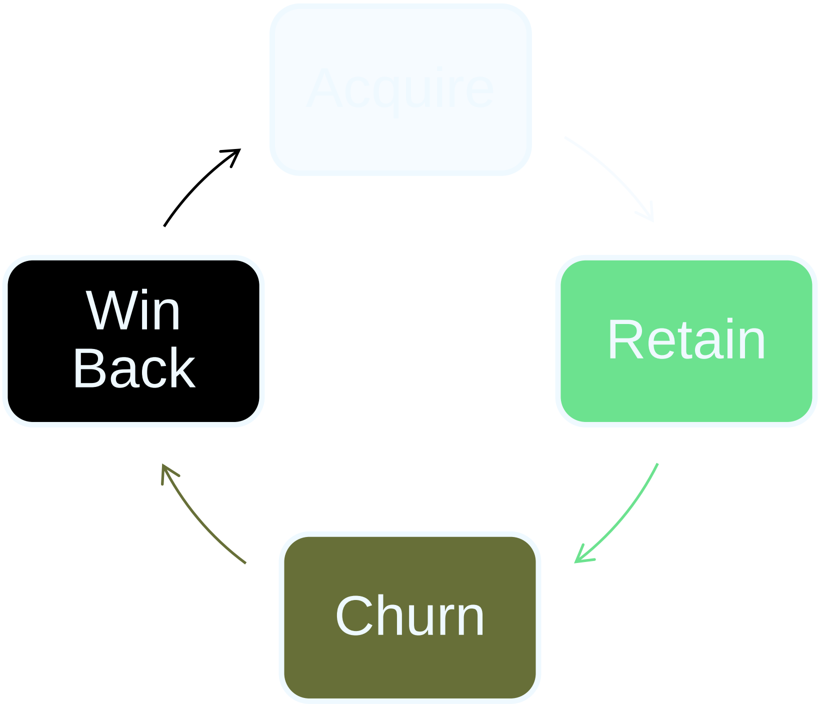
Module 2.1 Customer Lifetime Value
Customer Relationship Management

Eric Tham
Institute of Systems Science
National University of Singapore

isstyc@nus.edu.sg

Day 2 Lesson Plan

- **Day 2: Customer retention & growth**
  - » **Customer Lifetime value**
  - » **Customer Loyalty & demand**
  - » **Workshop using Excel simulation on CLV**



Content

- **Defining Customer Lifetime Value**
  - **Importance of CLV**
  - **Lifecycle of CLV**

- **Customer Lifetime Value factors**

- **Modelling Techniques**
  - **Simple Models**
  - **Complex Models: Hazard, Pareto/NBD models**

- **Practical approaches on CLV**

1. What is Customer Lifetime value (CLV)?

Defining Customer Life Time Value

Life Time Value of the Customer = Current Value + Potential Value

- The current value is whatever the customer has created in value for business as of today.
- Potential value is the future stream of ***profits*** expected from the customer as long as they continue to be a customer

- *We usually say*

  Customer Lifetime Value (CLV) = Potential Value

  as current value (computed up to now) is easy to compute

Defining CLV again … …

- The CLV of a customer *i* is the <u>discounted</u> value of the future profits yielded by this customer

$$CLV_i = \sum_{t=0}^{h} \frac{CF_{i,t}}{(1+r)^t}$$

- Where
  - $CF_{i,t}$ = net cash flow generated by the customer *i* activity at time *t*
  - h = time horizon for estimating the CLV
  - r = interest rate

- The CLV is the value added, by an individual customer, to the company over a set period in time

Why is CLV important?

- Helps to determine customer viability
  - which are the important ones to focus efforts on
  - Create and assess market strategies

- Important for company valuation especially start-ups

- Provides an input to the production and inventory management processes

CLV in Marketing

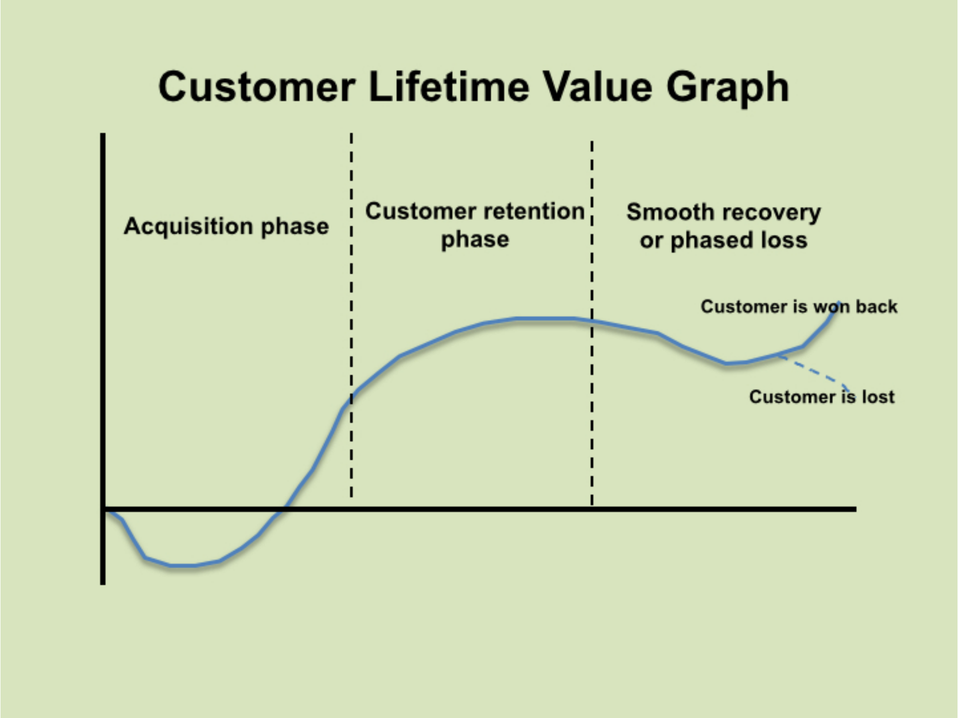
- The use of CLV in developing and assessing the performance of marketing strategies is
  - Determining the most valuable customers
  - Identifying those potential customers you want to attract
  - Determining which products will generate the most profit
  - Determining the possible profit from a marketing campaign
    - » Comparing with a budget of a marketing campaign
  - Measure the effectiveness of a past marketing campaign by evaluating the CLV change it incurred
  - Neglecting the less valuable ones, to which the company should pay less attention

- CLV should be considered whenever a significant decision is made at any level of the marketing organization

CLV Stages

- Customer goes through value stages when he is more valuable.

- Different for different industries.



2. CLV Factors

Factors in CLV modelling

- Time Horizon
  - *Where do we want to predict to?*
- Profit and costs estimation
  - *Cost of maintaining and retaining customers*
  - *Cost of attracting new customers*
  - *Determining the risky discount rate*
- Other issues
  - Different customer characteristics
  - Competition

Time Horizon

- We usually set the time horizon to be one that we can relate to strategic business objectives
  - Long-term relationship is important
    - » Take a long horizon, e.g. 10 years
  - Short-term relationship is important
    - » Take a small horizon, e.g. 1 year

- It is a function of your product, your industry, and your company strategy (and your perception of the general marketplace)
  - What is the lifespan of a computer game?
  - What is the lifespan of expendable goods?

Estimating Costs and Revenue

- The Gross Profit is simply the total revenue minus total costs
  - Costs should include production, distribution, marketing etc.
- Costs as a result of trying to optimize CLV will have to be factored into the equation
- Revenue is projected revenue.

Multiply gross profit by a discount rate to get the Net Present Value of the expected profits.

Discount Rate

- This is essentially unknown as it depends on economic factors which may be difficult to predict
- The mathematical formula for the discount rate by year t is: (to tie with slide 6)

$$\text{discount rate, d} \quad = \frac{1}{(1+r)^t}$$

  - where r = interest rate, and
  - t = number of years that you have to wait.

- Example: With an interest rate of 5%, the discount rate in the third year is
  - $d = 1/(1 + 0.05)^3 = 0.863$;
- r : defines the uncertainty of revenue.
  - For higher certainty, use a larger r; lower uncertainty use smaller r

Other Factors on CLV Calculation

- Customer Profiles
  - Changes in Demographics
  - Customer Lifestyle changes
  - Customer ageing
  - Customer Mobility
  - The demographics of new Customers

- Changes in Marketing Strategies
  - Promotions/End-of-range
  - Response to Rivals/New Rivals
  - Disruptive technology

3. CLV Modelling

Two main types of CLV:

- Contractual relationships
  – Examples are telco, magazine etc.
  – Easier to compute retention rates.
  – Also more customer information

- Non-Contractual relationships
  – No telling on customer transactions. Doesn't know 'time of loss' (or observed)
  – Can't tell distinct customer sometimes
  – Less customer information/ depends on transactional information. More 'complex' model using joint distributions.

More correct 'CLV'

$$E(CLV) = \int_0^\infty E[CF_t] S(t) d_t dt$$

The CLV needs to consider how long the customer stays on or survives S(t) – probability of survival.

Following, we model this survival function.

dt - discount factor

CLV Modelling Techniques

CLV is about ***survival analysis*** – how long a customer stays with you?

- 'Simpler' Models
  - Simulation models
  - Kaplan-Meier model
  - Hazard & Cox Models
- Complex Models (more for non-contractual)
  - Stochastic probability models
    - » Pareto\NDB
  - Markov Models*
  - Persistence Models*

It should be noted that this is a developing field, and there is no approach that is universally accepted

* Not covered in this course

CLV Simple Illustration

Refer to excel worksheet ‘CLV Simple’.

This illustrates segmentation by frequent and occasional buyers. It illustrates the concept how different customers can have different CLV values, and the underlying assumptions on them.

CLV Simulation model

An example case study will be covered in the workshop on how the CLV can be simulated from a transactions database using assumptions on the factors discussed earlier on.

CLV Modelling – Survival analysis

- A key assumption in CLV calculation is how long the customer will *stay*…or rather ***churn***
  - *Survival analysis is a set of methods to analyze data where the outcome variable is the time until the occurrence of an event of interest (customer churn)*

S(t) = probability that a customer will survive until time t (assuming measurement began at time t=0)

- Two main difficulties with doing survival analysis
  - censoring
  - non-normality

  Regression methods cannot be used.

CLV Modelling – Survival analysis

- Censoring refers to information being incomplete for survival analysis.
  - Survival analysis data for a customer that has not 'churned' till time t – only know it survives up to time t but can't be sure thereafter how long it survives.

- Normality is a key assumption for most regressions.
  - Normality implies possibility of negative data whilst in survival analysis, it is always *positive* time for event to happen.

CLV Modelling – Survival analysis

- Two key concepts
  - Survival rate
  - Hazard rate

  - The survival rate refers to the proportion of customers staying on as such up to *time t*. The Kaplan-Meier method is used to estimate survival rates St.

  - The hazard rate refers to the probability of observing customer(s) churning from amongst the population.

Kaplan Meier Estimator

- This is used to estimate survival rates from the observed data.
- Define:

$$S_t = \frac{\text{number of individuals surviving longer than t}}{\text{total number of individuals studied}}$$

- Time t can be 1, 2, 3 years etc.
- Then the estimated survival rate can be:

$$\hat{S}_t = \prod_{t_i \le t} \left[1 - \frac{d_i}{n_i} \right]$$

- t_i is duration of study at point i, d_i is number of churns up to point i and ni is number of customers at risk just prior to t. S is the conditional probability that a customer survives at the end of a time interval t_i, provided customer is still there at start of the time interval.

Kaplan Meier Estimator

Suppose after a successful marketing campaign, your company has 1000 customers. Your director estimates that after third year without any promotional efforts, the number of customers is expected to drop to 750, 600, 450 at the end of the 1st , 2nd, 3rd years. Based on this data estimate the survival rate of customers.

Time	No of Customers
0	1000
1	750
2	600
3	450

Kaplan Meier Estimator

The Kaplan-Meier estimator is calculated below where the customers that left at each year are the *censored* customers.

Time	No of Customers	d_i	1-di/ni
0	1000	250	0.750
1	750	150	0.800
2	600	150	0.750
3	450		
	Survival Rate,	$\hat{S}_t$	0.450

The calculation assumes that censored and uncensored customers have the same probability of leaving. This means that even as more customers leave, those that remain are still as ‘loyal’ as those before.

Kaplan Meier Estimator

The *average (survival) time* a customer stays is now estimated from the time that number of customer left is halved from the beginning.

	A	B	C	D
1				
2	**Time**	**No of Customers**	**d_i**	**1-di/ni**
3	0	1000	250	0.750
4	1	750	150	0.800
5	2	600	150	0.750
6	3	450		
7		Survival Rate,	$\hat{S}_t$	0.450
8		**Mean survival time of a customer**		
9				2.667
10				
11				

=(500-B6)/(B5-B6)*A5+(B5-500)/(B5-B6)*A6

Kaplan Meier Estimator

The *cumulative hazard rate H(t)* is estimated as:

$$H(t) = -\ln(\hat{S}_t)$$

In this case, it is just 0.798. This is the cumulative hazard rate up to the *3rd year*. It refers to the accumulated proportion of customers leaving up till that year.

As an exercise, compute the cumulative hazard rate up to the 2nd year.

This brings us to another hazard rate – ***instantaneous*** hazard rate and the Poisson distribution.

Poisson Distribution

- The Poisson distribution has a parameter – (instantaneous) hazard rate λ and its probability density function is given by:

$$Pr(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

  - Here, X is the number of customers (=k) that churn in a unit time (say 1 year). The hazard rate λ indicates the likeliest number (mode) of customers that will churn in the given time.

Hazard Rate(s)

- The instantaneous and cumulative hazard rate, and the survival function are related.

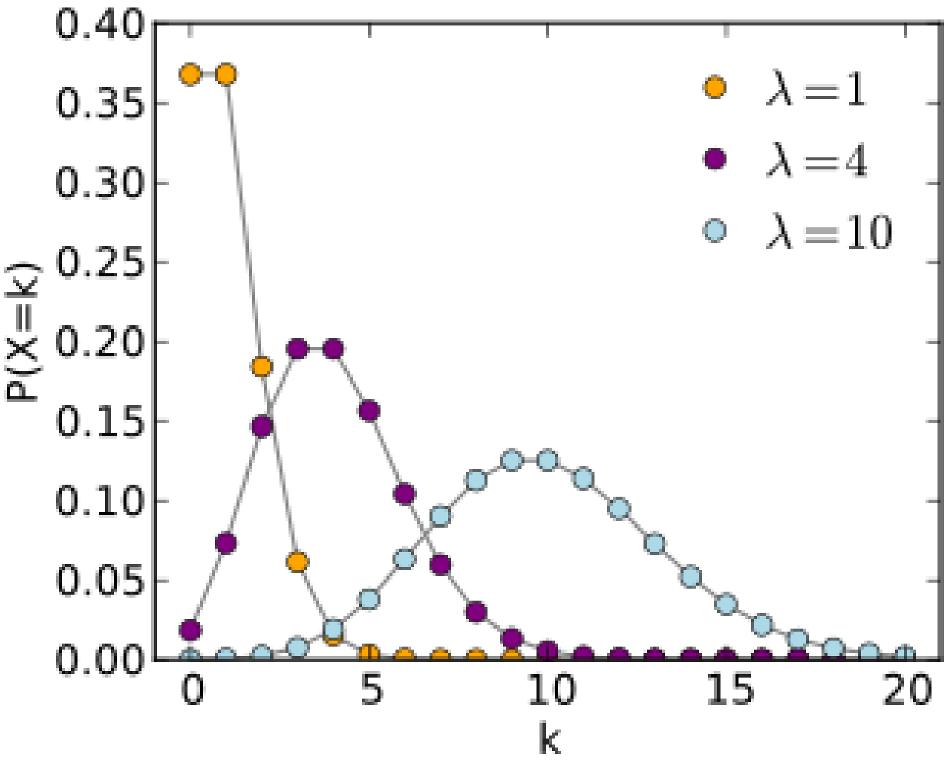
$$H(t) = -\log(S(t))$$

$$S(t) = \exp(-H(t))$$

$$\lambda(t) = -\frac{\delta log(S(t)}{\delta t}$$

Go back to the spreadsheet on slide 25, and compute the instantaneous hazard rate.

Assumptions of the Poisson Distribution



As hazard rate λ increases, it becomes like a normal distribution.

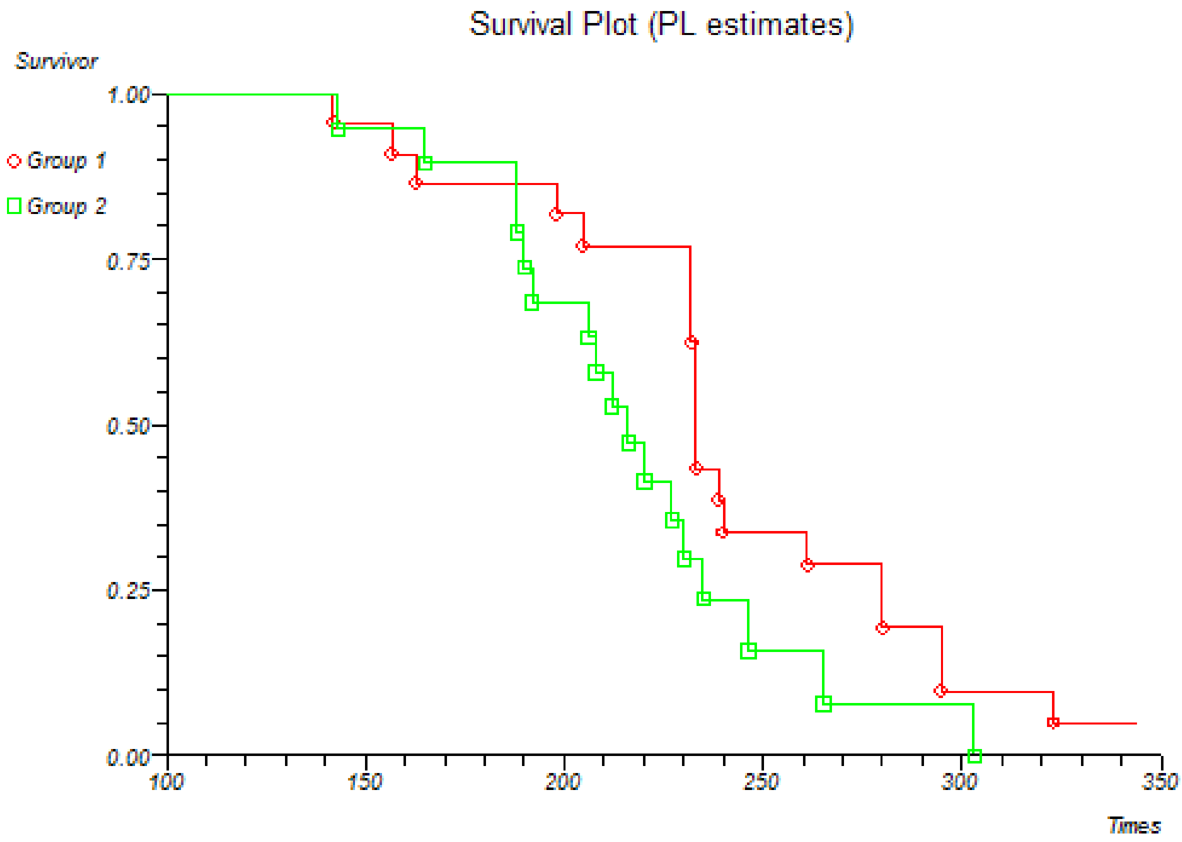
Assumptions of the Poisson Distribution

Distribution is appropriately used only if:

- k is discrete
- There are no ‘herds’. If some customers churn, it doesn’t lead to manic herds. That is customer churning is independent of one another.
- Churning rate is constant over time
- The probability of a customer churning is higher the longer the time.

Different Customer Groups

In general, it is possible to segment customers and then define their different churn (survival) rates accordingly. (illustrate in workshops



Computation of the survival rates and hazard rates

Please refer to excel worksheet for Day 2 ‘CLV Complex’

Do the calculations for the 2nd group of customers 2. Compute the hazard rate, survival rates for these group and plot the survival plot for these groups.

What can you say about these groups of customers?

Go menti.com 760317 and test your understanding.

Proportional Hazard or Cox Models

- The Cox Model relates the hazard rate λ to predictor variables X_1, X_2, … by:

$$\lambda(t|X) = \lambda_0(t) exp(\beta_1 X_1 + \ldots + \beta_p X_p)$$

- Here $\lambda_0(t)$ is the baseline hazard rate. The model parameters can be estimated by maximum likelihood method, which is essentially an optimising technique to fit the model to data reducing the errors as far as possible.

- The 'proportional' comes about as if X_1 increases, the impact on the hazard rate λ increases proportionally.

Predicting λ with predictor variables

$$\lambda = c_0 + \beta_1 * \text{gender} + \beta_2 * \text{age} + \ldots\ldots..$$

Some constant

Some constant coefficient

- The hazard rate can be regressed against predictor variables to determine the coefficients. In the above case:
  - Gender would be a binomial variable

    Gender = 0 if customer is female

    = 1 if customer is male
  - Customer age (in years) would be a real positive variable

That is supposedly age and gender can impact how long a customer stays on.

Stochastic CLV models

- Thus far, the hazard rate model considers the proportion of customers that churn and doesn't consider the transactional amount.
- Stochastic CLV models consider it differently. They consider:
  - It uses a "coin" to determine whether a customer churns → Pareto distribution
  - Then it uses "dice" to determine how many items a customer will order → Negative binomial distribution

Stochastic Models:The Pareto/NBD model

- From the stochastic CLV model, the CLV is calculated as

$$CLV_i = \sum_{t=0}^{h} \frac{x_{i,t} \times m_{i,t}}{(1+d)^t}$$

- Where
  - $x_{i,t}$ = no of customer *i's* surviving till the period *t*
  - $m_{i,t}$ = profit per transaction performed by customer *i* in the period *t*
  - d = discount rate
  - h = time horizon of the prediction

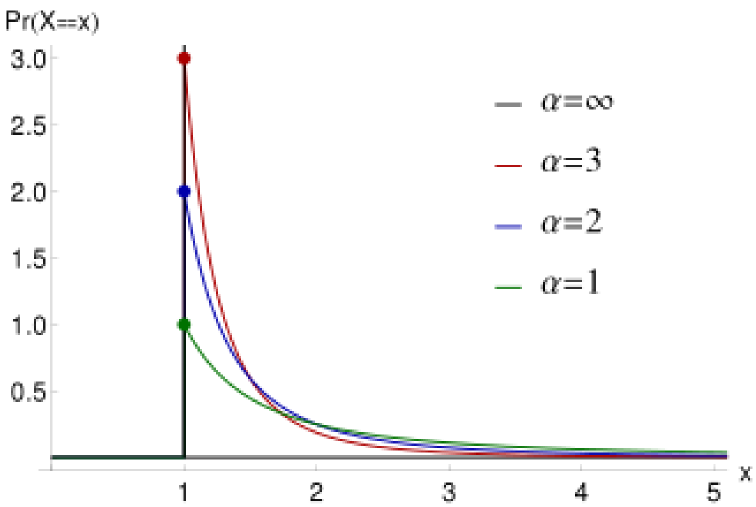
- Advantage of Pareto/ NBD is the heterogeneity of the individual customers' propensity to churn modelled by $x_{i,t}$

Pareto/NBD Model

- Pareto distribution: recall Pareto distribution of wealth.
  - 20% of the population contains 80% of the wealth. In this context, most of the churns occur in the initial years. Assume a customer is less likely to churn with increasing year.
  - The Pareto distribution has a shape parameter α that determines this initial proportion. The location parameter μ determines which year customer is most likely to churn.

Reference: http://www.brucehardie.com/talks/ho_cba_tut_art_09.pdf

Pareto Distribution



As shape α ↑, greater proportion of customers likely to churn in first year.
As location μ ↑, the curve shifts to the right.
Its CDF is given by for x > α: $1 - (\frac{\mu}{x})^{\alpha}$

Pareto/NBD Model

- The negative binomial distribution is a discrete probability distribution for the number of 'Head' in coin tosses, before a tail appears.
- Models lesser propensity to churn in later years. It has 2 parameters to determine the pdf.

$$X \sim \mathrm{NB}(r; p) \Rightarrow \begin{bmatrix} k+r-1 \\ k \end{bmatrix} (1-p)^r p^k$$

- k– the no of churn (=1 in our context)
- p – probability of churning at each stage
- r – year no

Go to menti.com code 760317 to calculate CLV based on NBD model and Excel sheet 'Day2-NegBino'.

Practical issues to compute the CLV Model

- Be simple : Start from a baseline (simple) probability model and make modifications from it.
- Mind the major features is probably enough. For eg. certain customer types will be more likely to churn.
- There will be a lack of data. Use expert opinion to determine parameters if necessary. May be better than 'calibrating' model.
- Know the limitations of your model. It may work under certain scenarios (or not) and the future may not repeat these scenarios. Know what these scenarios are.

Conclusion

This session gives:

- Gave an overview of the concept of CLV (CLV)
- Explained how CLV can be modelled by survival analysis
  - Using a simple approach
  - Using proportional hazards
  - Using a Pareto/NBD model based approach
- A *practical* approach to computing CLV