Customer Segmentation

By Atufa Shireen

Why?

- Understanding customers behaviour towards business products.
- Targeted marketing.
- The future monetary value expected from customers.
- The probability of leaving the business.
- Expected number of transactions in future.

Segmentation Strategies

- Demographic,
- Psychographic,
- Geographic,
- **Purchasing Behaviour**

Information Required

- When the customer has joined.
- When or how much the customer has interacted with business.
- Amount spent on transactions.

Definitions

Recency

- How recently a customer has made a purchase.(end of obser - last purchase)

Frequency

- How often a customer makes a purchase. (num of purchases)

Monetary Value

- How much money a customer spends on purchases. (total amount spent)

Process

- Apply Kmeans to customer’s recency ,frequency,monetary.
- Rearrange labels with clusters mean (highest label for highest mean).
- Rearrange data with descending order for Recency.
- Rearrange data with ascending order for Frequency.
- Rearrange data with ascending order for monetary.
- Combine cluster label and give segment label based on RFM Score.

CLV Strategies

- aggregate analysis.
- **probabilistic models.**
- cohort analysis.
- survival analysis models.

Probabilistic Models

- Pareto/NBD (PGGG) Model.
- BG/BB (PNBD) Model.
- MBG/NBD Model.
- **BG / NBD (BGNBD) Model.**

- Gamma Gamma Model.

Assumptions

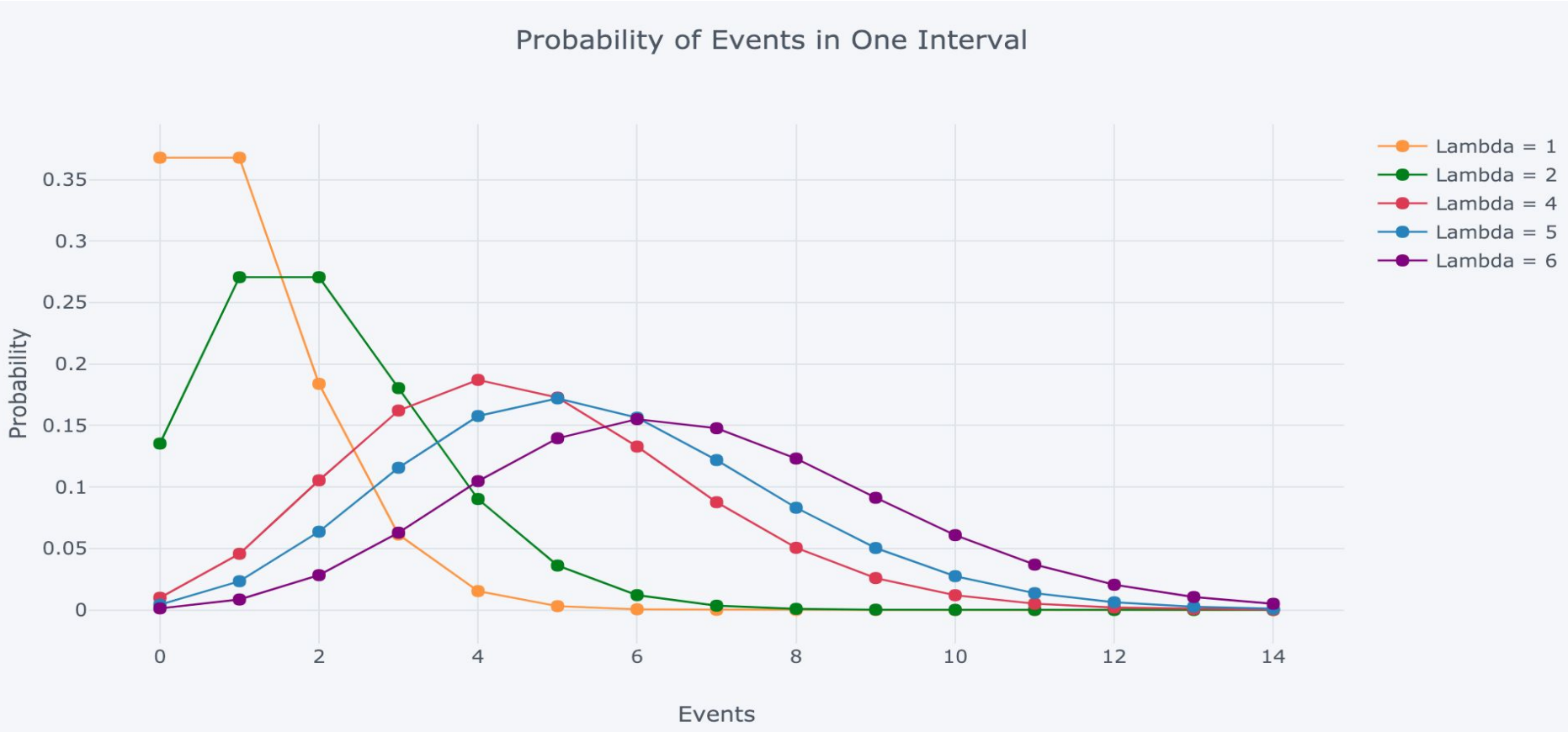
For BGNBD Model

1. When a user is active, number of transactions in a time t is described by ***Poisson distribution***
   (Poisson distribution helps to predict certain events happening using the data about how often event occurred in past)
2. Heterogeneity in transaction across users (difference in purchasing behavior across users) has ***Gamma distribution***.
   (Gamma distribution arises naturally in processes with waiting time between Poisson distributed events (as in our case for transaction rate λ)
3. Users may become inactive after any transaction with probability p and their dropout point is distributed between purchases with ***Geometric distribution***.
   (Geometric distribution is similar to Bernoulli trials and is used for modeling the number of trials up to and including the first success.)
4. Heterogeneity in dropout probability has ***Beta distribution*** .
   (Beta distribution is the best for representing a probabilistic distribution of probabilities — the case where we don't know what a probability is in advance, but we have some reasonable priors, described by α and β (mean of a Beta distribution α / (α+β)))
5. Transaction rate and dropout probability vary independently across users.

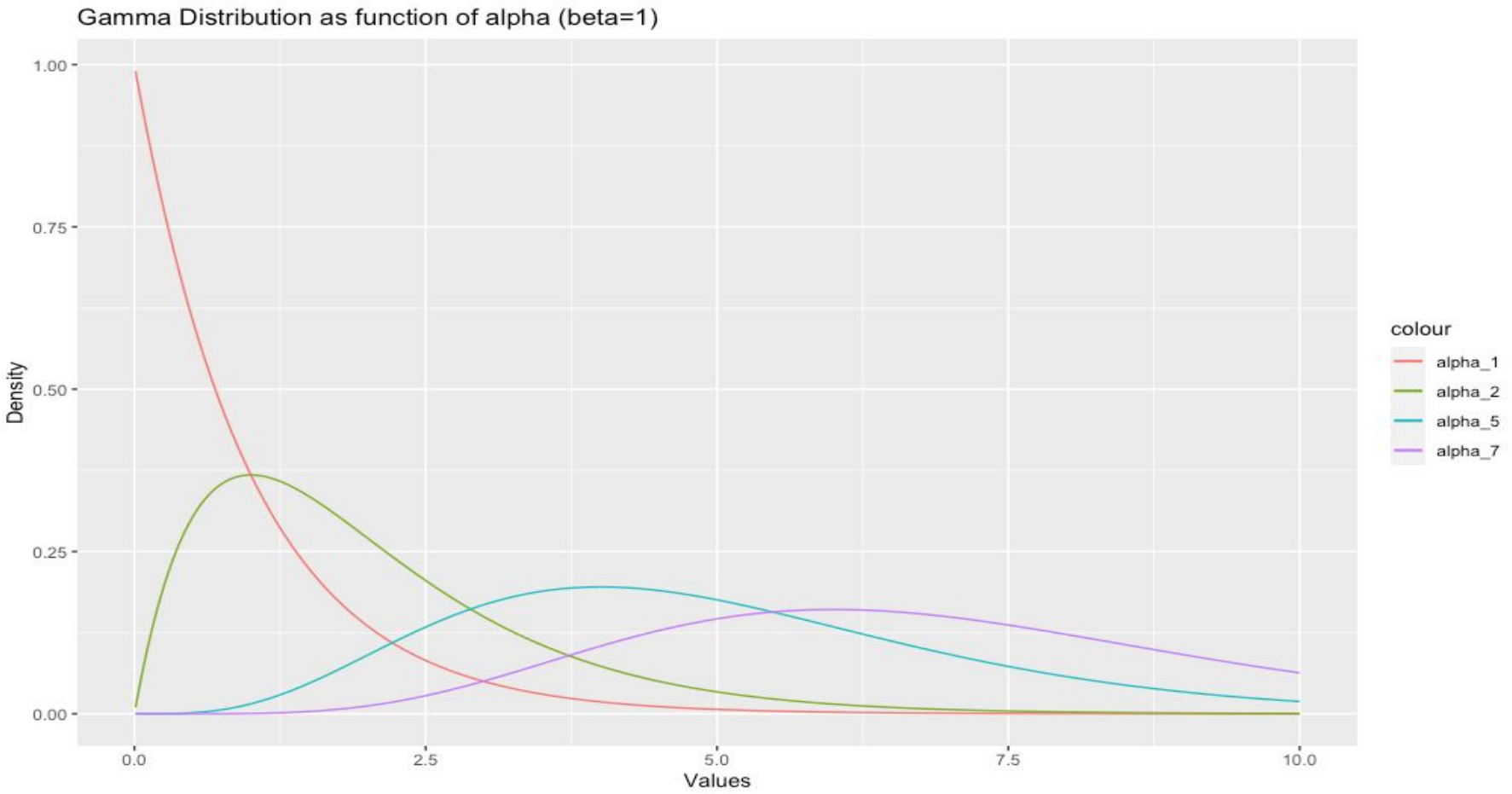
For GammaGamma Model

6. Monetary value of users' transactions is random around their mean transaction value.
7. Mean transaction value varies across users but doesn't vary for an individual user over time.
8. Mean transaction values is Gamma distributed across customers.

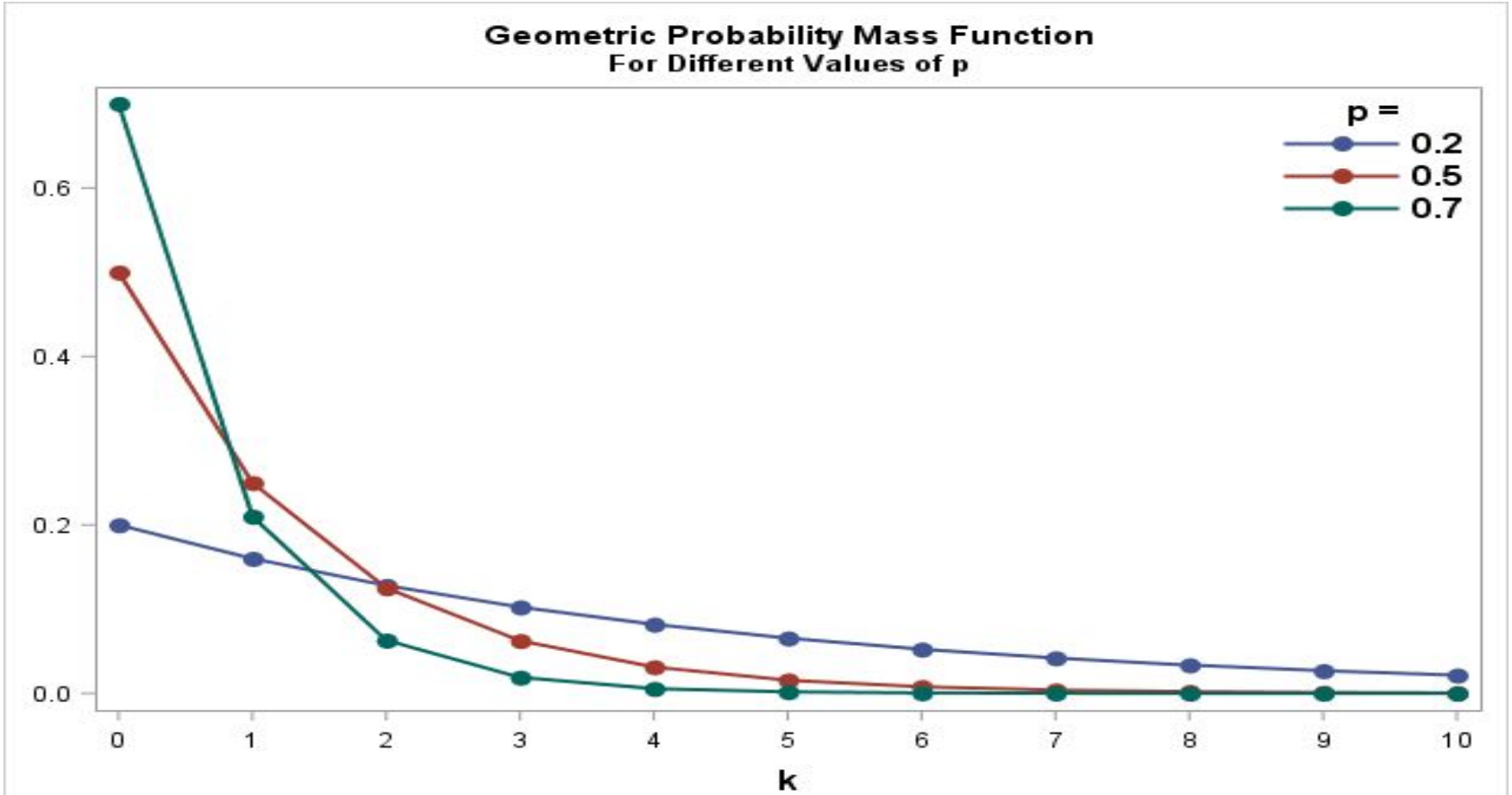
Poisson Distribution



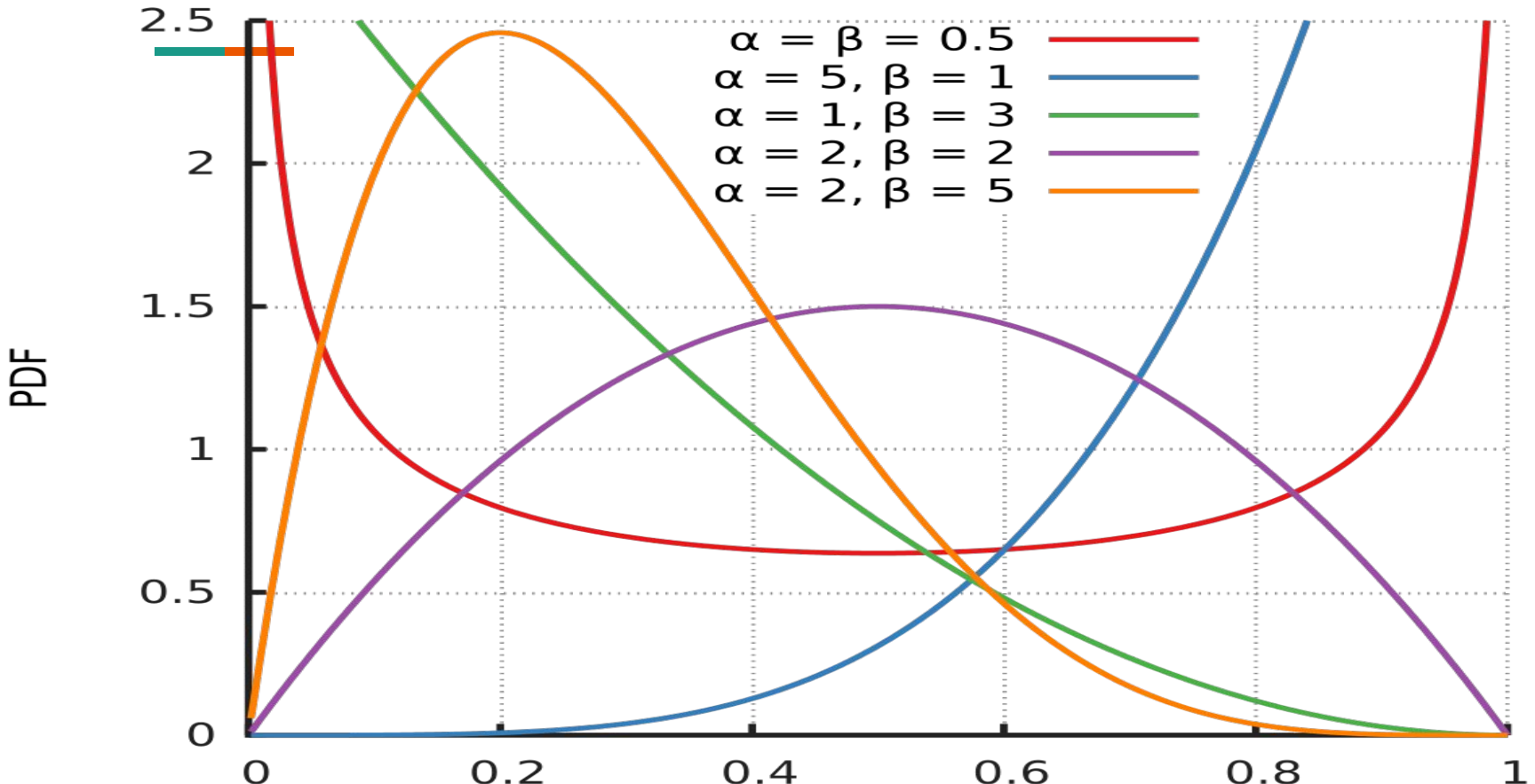
Gamma Distribution



Geometric Distribution



Beta Distribution



Definitions

Recency

- How recently a customer has made a purchase excluding the first purchase.(end of obser - last purchase)

Frequency

- Number of times the user made purchase in time period.(excluding the first transactions)

Monetary Value

- Average amount spent by customers excluding the first purchase. (total amount spent)

Age

- How long has it been since joining. (end of obs - first purchase)

KPIs

- Churn Prediction
- Future Transactions
- LTV