L3: Probability models cont'd

Objectives:

1. Introduction to probability models.

2. Explore non-contractual probability models.

Luca Gaegauf | luca.gaegauf@access.uzh.ch

Agenda

1. Probability models for discrete non-contractual settings
2. Probability models for continuous non-contractual settings
3. References

Luca Gaegauf | luca.gaegauf@access.uzh.ch

Agenda

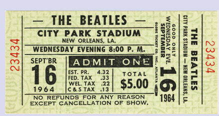
1. **Probability models for discrete non-contractual settings**
2. Probability models for continuous non-contractual settings
3. References

Probability models in a discrete non-contractual setting

Continuous non-contractual


Continuous contractual


Discrete non-contractual



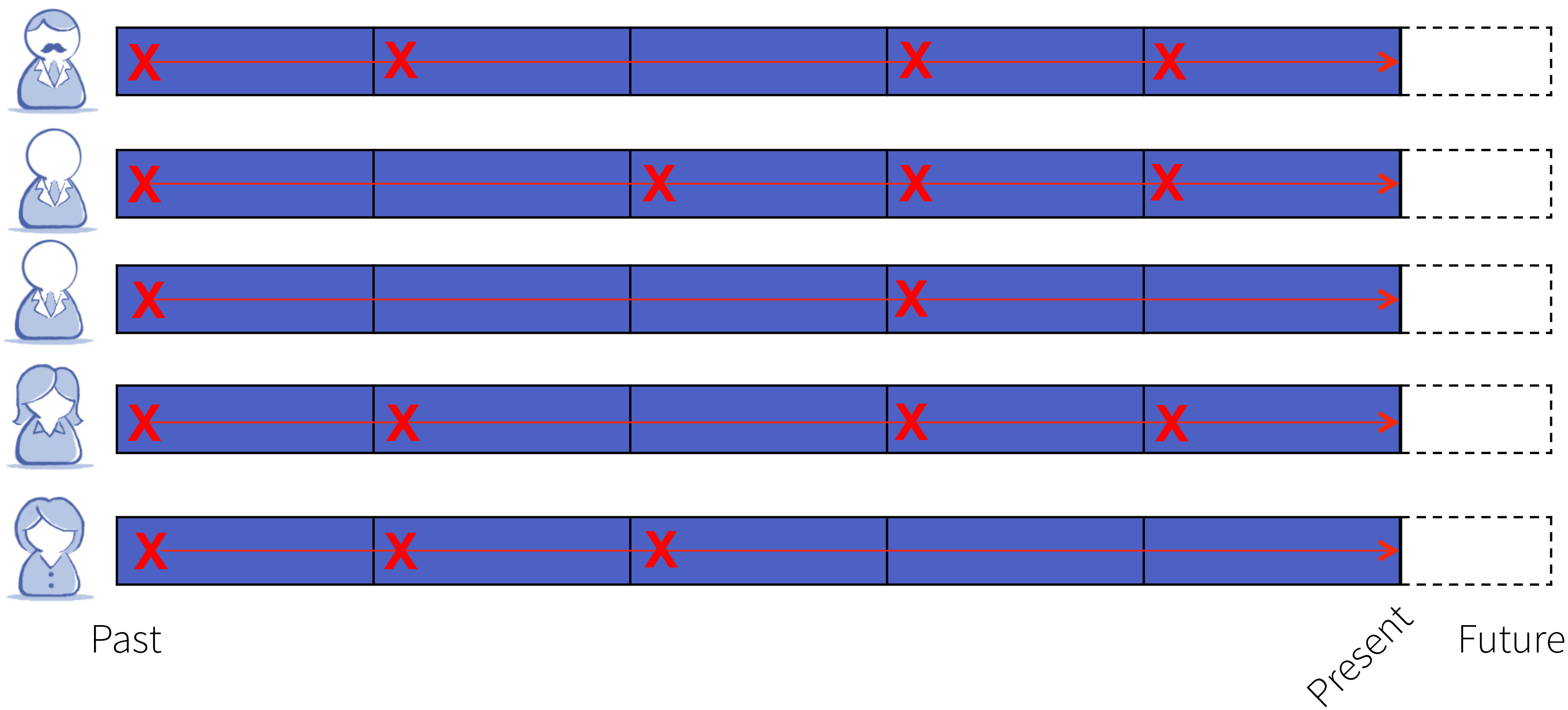
Discrete contractual


Fader & Hardie (2012)

Luca Gaegauf | luca.gaegauf@access.uzh.ch

How to think about the discrete non-contractual setting

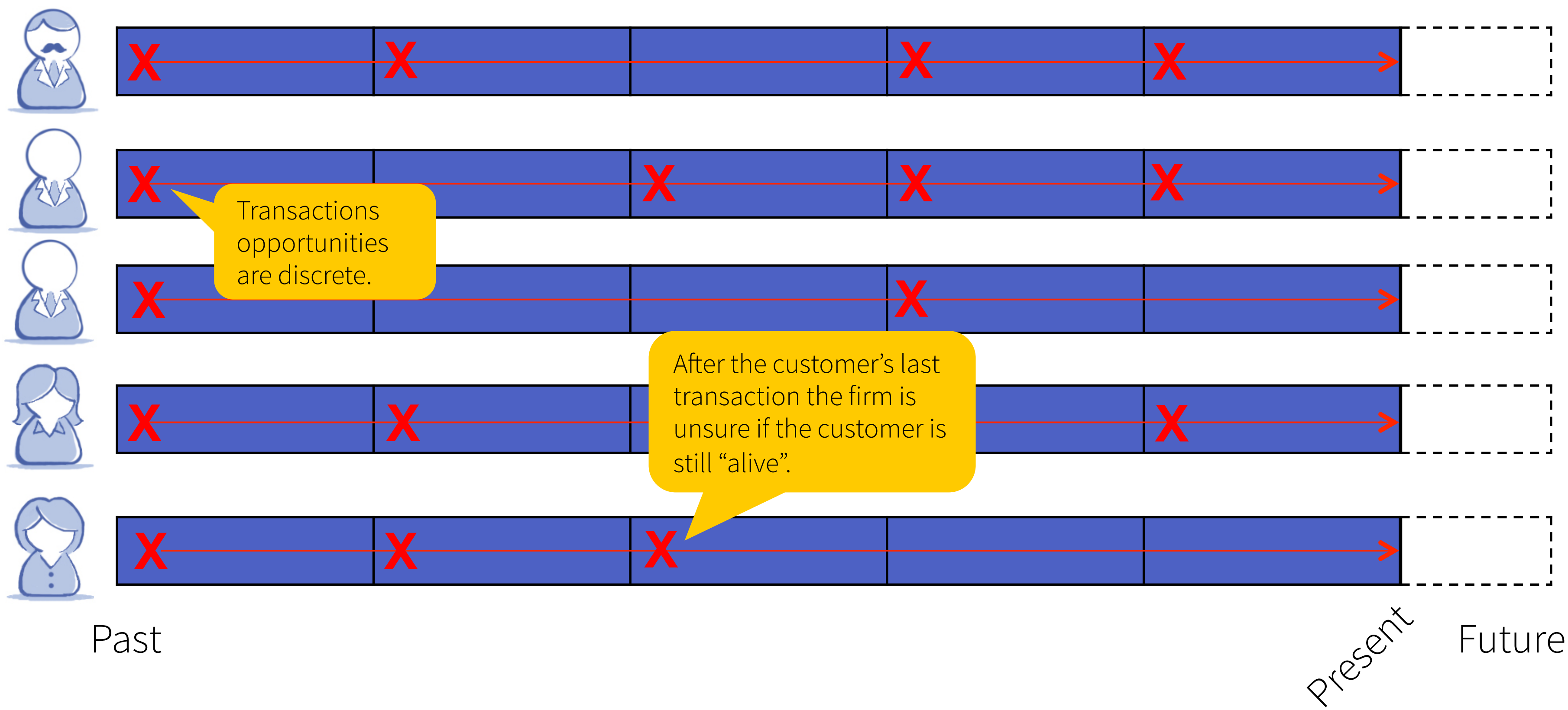
Transactions are conducted on set intervals.



 Luca Gaegauf | luca.gaegauf@access.uzh.ch

How to think about the discrete non-contractual setting

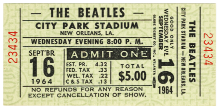
Transactions are conducted on set intervals.



We can differentiate between different discrete settings

- A transaction opportunity is:
  - A well-defined **point in time** at which a transaction can occur.
  - A well-defined **time interval** during which a transaction can occur.

"Necessarily discrete"

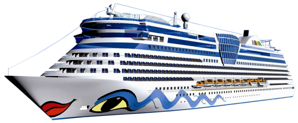


e.g. music event

"Generally discrete"

e.g. charity donations

Discretized by recording process



e.g. cruise vacation

Fader & Hardie (2012)

Luca Gaegauf | luca.gaegauf@access.uzh.ch

Discrete non-contractual vs. continuous non-contractual

The discrete non-contractual setting is a special case of the continuous non-contractual setting.

	Discrete non-contractual	**Continuous non-contractual**
Transaction process	Beta binomial	Negative binomial distribution
Dropout process	Beta geometric	Pareto type II/EG

Fader & Hardie (2012)

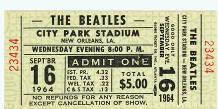
Agenda

Probability models in a continuous non-contractual setting

Continuous non-contractual


Continuous contractual


Discrete non-contractual

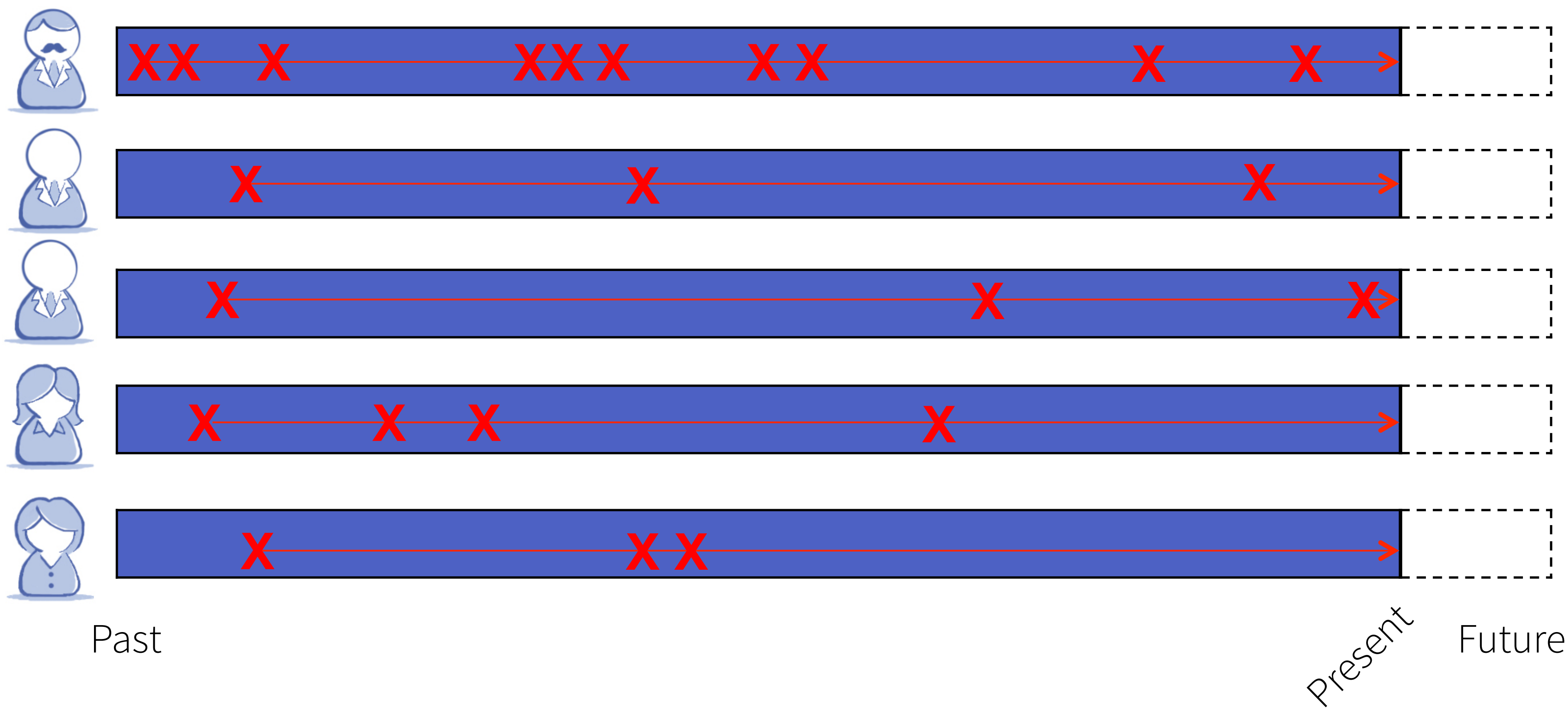


Discrete contractual


Fader & Hardie (2012)

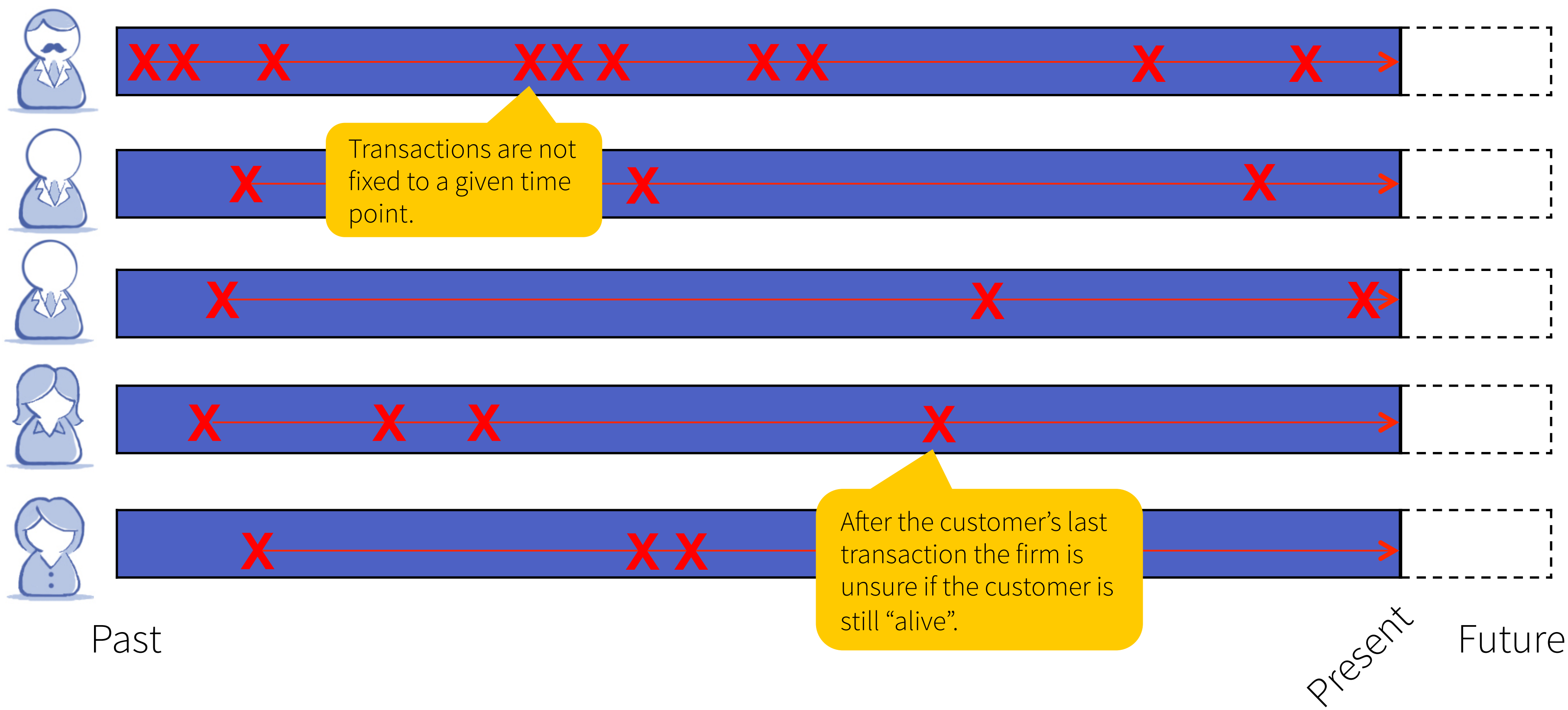
How to think about the continuous non-contractual setting

Transactions are continuous and the company is not notified when customers leave.



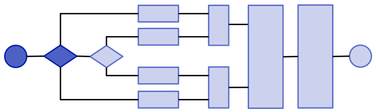
How to think about the continuous non-contractual setting

Transactions are continuous and the company is not notified when customers leave.



Step 1: Determine the marketing decision problem / information needed

How high can we expect the transaction level to be in the next period?

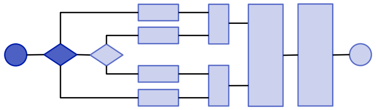


Step 2: Identify the observable individual-level behavior of interest

How high can we expect the transaction level to be in the next period?

Transaction process

Dropout process



Step 2: Identify the observable individual-level behavior of interest

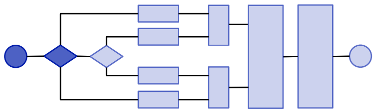
How high can we expect the transaction level to be in the next period?

Transaction process:

How often each customer purchases.

Dropout process:

The propensity of a customer to dropout.



In order to build the Pareto/NBD model we need to know…

- … the total number of past transactions of each customer (**frequency**).
- … the time passed since the last transaction (**recency**).
- Information about the exact transaction time is not needed.

Transforming the transaction data to build the Pareto/NBD model

Customer ID	Date	Sales ($)
ba366-lk	01.09.14	50.30
dq722-nZ	31.08.14	43.10
hW866-Qe	31.08.14	120.45
wp63i-Zu	31.08.14	1.25
ue109-qc	31.08.14	5.00
wp63i-Zu	30.08.14	5.20
EJ370-tG	30.08.14	81.90
wp63i-Zu	29.08.14	132.70
CC434-pv	29.08.14	21.20
dq722-nZ	29.08.14	
eh52a-vo	29.08.14	
...	...	...

Transaction data

Note: Transaction dates occur on a daily basis.

Transforming the transaction data to build the Pareto/NBD model

Customer ID	Date	Sales ($)
ba366-lk	01.09.14	50.30
dq722-nZ	31.08.14	43.10
hW866-Qe	31.08.14	120.45
wp63i-Zu	31.08.14	1.25
ue109-qc	31.08.14	5.00
wp63i-Zu	30.08.14	5.20
EJ370-tG	30.08.14	81.90
wp63i-Zu	29.08.14	132.70
CC434-pv	29.08.14	21.20
dq722-nZ	29.08.14	36.55
eh52a-vo	29.08.14	68.30
...	...	...

Transaction data

Customer ID	Number of transactions	Time of last transaction	Time interval (weeks)
ba366-lk	2	21.7	39
dq722-nZ	4	21.5	39
hW866-Qe	7	21.5	39
wp63i-Zu	10	21.5	39
ue109-qc	2	21.5	39
jw272-zF	0	0.00	39
EJ370-tG	1	21.4	39
iI151-ze	0	0.00	39
CC434-pv	1	21.3	39
PJ291-jL	0	0.00	39
eh52a-vo	1	21.3	39
...	...	...	...

Transformed data

Transforming the transaction data to build the Pareto/NBD model

Customer ID	Date	Sales ($)
ba366-lk	01.09.14	50.30
dq722-nZ	31.08.14	43.10
hW866-Qe	31.08.14	120.45
wp63i-Zu	31.08.14	1.25
ue109-qc	31.08.14	5.00
wp63i-Zu	30.08.14	5.20
EJ370-tG	30.08.14	81.90
wp63i-Zu	29.08.14	132.70
CC434-pv	29.08.14	21.20
dq722-nZ	29.08.14	36.55
eh52a-vo	29.08.14	68.30
...	...	...

Transaction data

Customer ID	Number of transactions	Time of last transaction	Time interval (weeks)
ba366-lk	2	21.7	39
[illegible]	4	21.5	39
[illegible]	7	21.5	39
wp63i-Zu	10	21.5	39
ue109-qc	2	21.5	39
jw272-zF	0	0.00	39
EJ370-tG	1	21.4	39
iI151-ze	0	0.00	39
CC434-pv	1	21.3	39
P[illegible]	[illegible]	0.00	39
e[illegible]	[illegible]	21.3	39
...	...	...	...

Frequency: number of transactions in the holdout period.

Recency: time of last transaction within the holdout period.

Transformed data

Luca Gaegauf | luca.gaegauf@access.uzh.ch

Transforming the transaction data to build the Pareto/NBD model

Customer ID	Date	Sales ($)
ba366-lk	01.09.14	50.30
dq722-nZ	31.08.14	43.10
hW866-Qe	31.08.14	120.45
wp63i-Zu	31.08.14	1.25
ue109-qc	31.08.14	5.00
wp63i-Zu	30.08.14	5.20
EJ370-tG	30.08.14	81.90
wp63i-Zu	29.08.14	132.70
CC434-pv	29.08.14	21.20
dq722-nZ	29.08.14	36.55
eh52a-vo	29.08.14	68.30
...	...	...

Transaction data

Customer ID	Number of transactions	Time of last transaction	Time interval (weeks)
ba366-lk	2	21.7	39
dq722-nZ	4	21.5	39
hW866-Qe	7	21.5	39
wp63i-Zu	10	21.5	39
ue109-qc	2	21.5	39
jw272-zF	0	0.00	39
EJ370-tG	[illegible]	[illegible]	39
iI151-ze	[illegible]	[illegible]	39
CC434-pv	[illegible]	[illegible]	39
PJ291-jL	[illegible]	[illegible]	39
eh52a-vo	1	21.3	39
...	...	...	...

Time interval is the duration of the holdout period. (This will be excluded on the further slides to save space.)

Transformed data

Step 2: Identify the observable individual-level behavior of interest: Transaction process

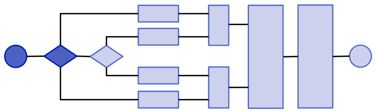
How high can we expect the transaction level to be in the next period?

Transaction process:

How often each customer purchases.

Dropout process:

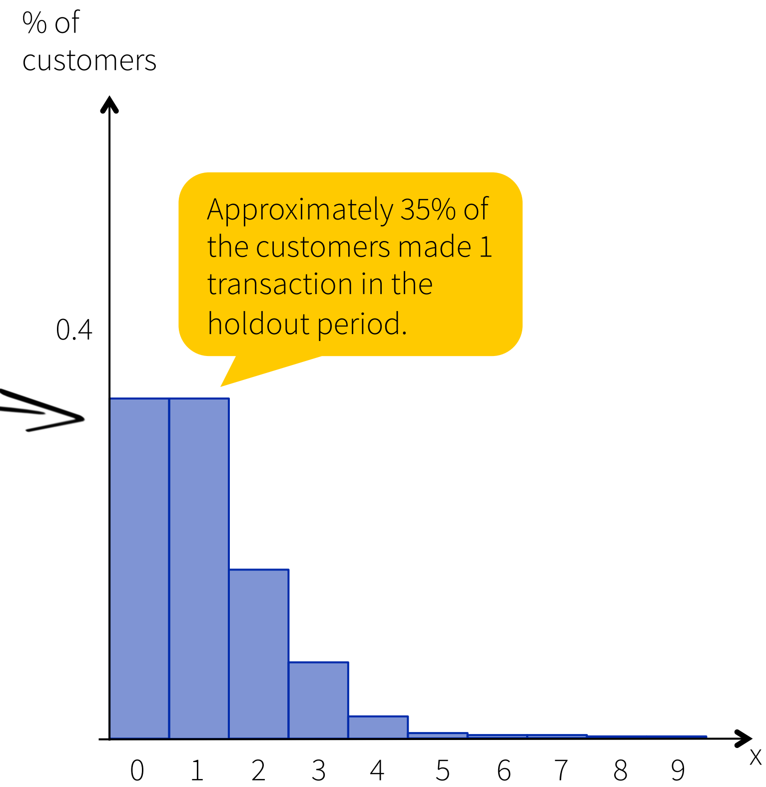
The propensity of a customer to dropout.



Step 2: Identify the observable individual-level behavior of interest: Transaction process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
iI151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data

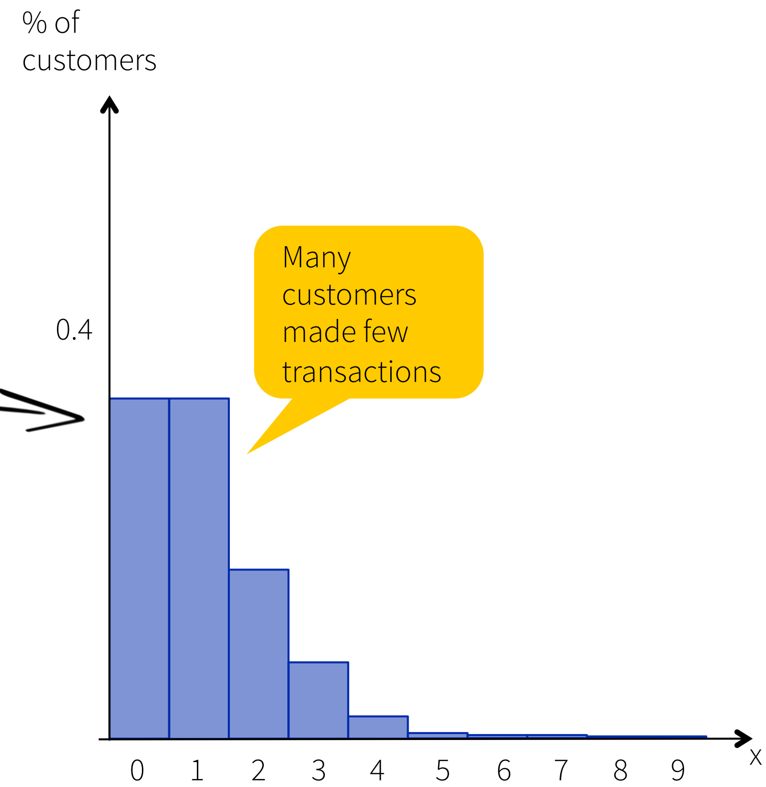


Cohort view

Step 2: Identify the observable individual-level behavior of interest: Transaction process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
iI151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data



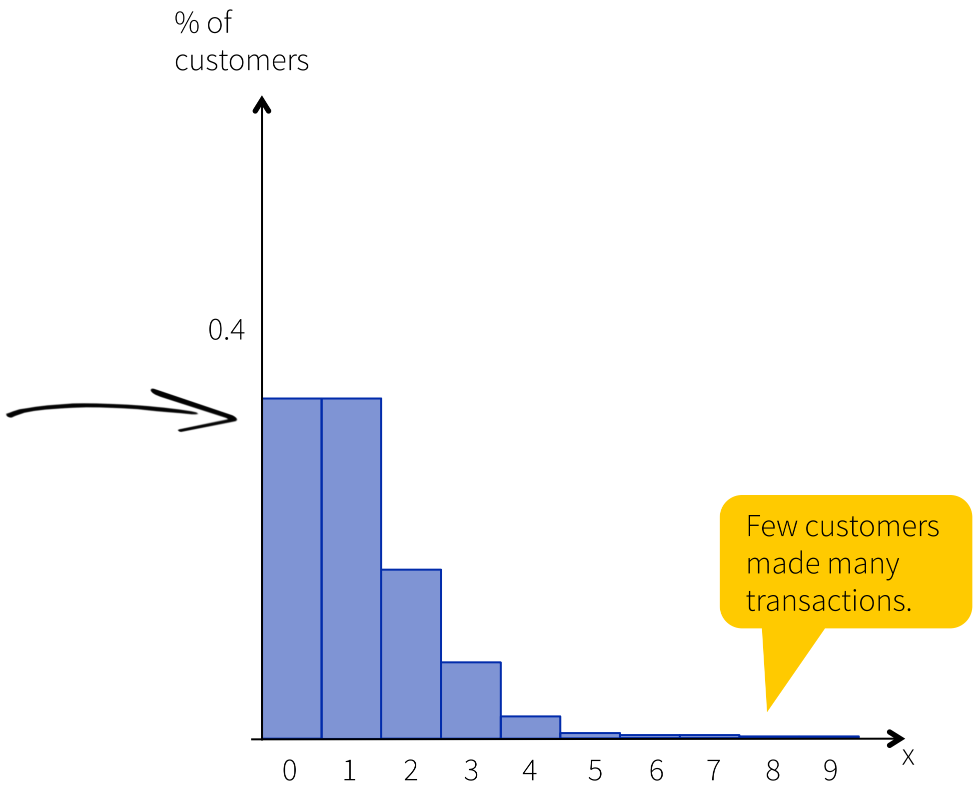
Cohort view

 Luca Gaegauf | luca.gaegauf@access.uzh.ch

Step 2: Identify the observable individual-level behavior of interest: Transaction process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
il151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data



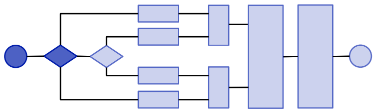
Cohort view

Luca Gaegauf | luca.gaegauf@access.uzh.ch

Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	(Continuous) **Timing** (Discrete)	**Choice**
How many?	When? How long?	Whether or not? Which?

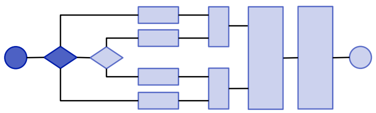


Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	(Continuous) **Timing** (Discrete)	Choice
How many?	When? How long?	Whether or not? Which?

"How many transactions were made during the observation period?"



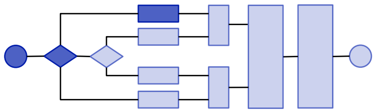
Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	(Continuous) **Timing** (Discrete)	Choice
How many?	When? How long?	Whether or not? Which?

"How many transactions were made during the observation period?"

Phenomenon	Individual-level	Heterogeneity	Model
Counting e.g. 1 product, 2 products, …	Poisson	gamma	NBD (negative binomial distribution)



Step 3: Select a probability distribution that characterizes this individual-level behavior

- λ is the mean transaction rate of the customer.
- The distribution of the transaction rate is described by:

$$P(X = x \mid \lambda) = \frac{\lambda^{x} e^{-\lambda}}{x!}, x = 0, 1, 2, \ldots$$

The transaction rate is characterized by a **Poisson distribution**.

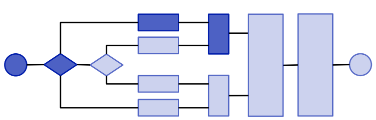
Fader & Hardie (2012)

Luca Gaegauf | luca.gaegauf@access.uzh.ch

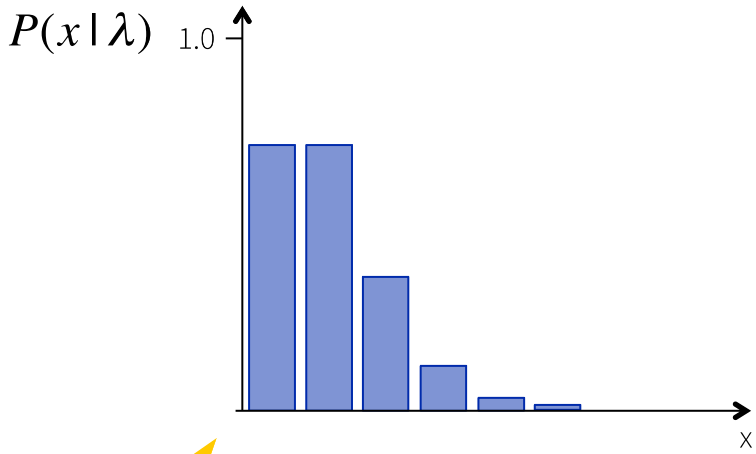
Step 4: Specify the distribution of the latent trait variables

Intuition: While a customer is alive, he purchases “randomly” around his mean transaction rate.

Phenomenon	**Individual-level**	**Heterogeneity**	**Model**
Counting e.g. 1 product, 2 products, …	Poisson	gamma	NBD (negative binomial distribution)

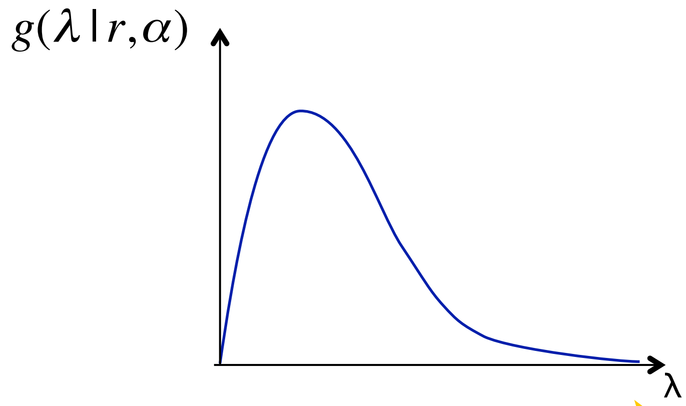


Step 5: Derive the corresponding aggregate or observed distribution for the behavior of interest



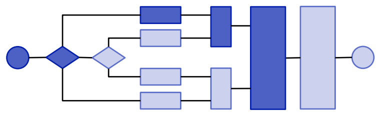
Poisson distribution

The transaction rate is characterized by a Poisson distribution.



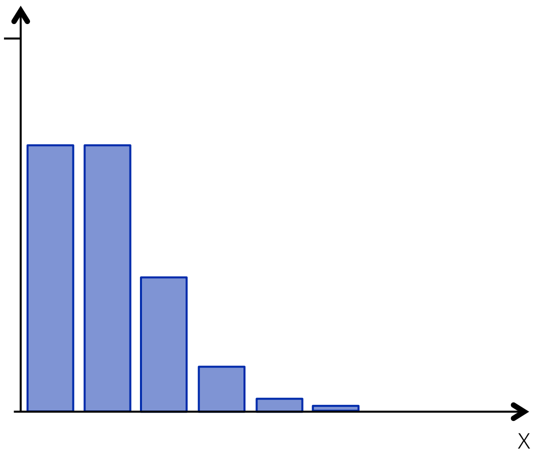
Gamma distribution

The latent traits are characterized by a gamma distribution .



Step 5: Derive the corresponding aggregate or observed distribution for the behavior of interest

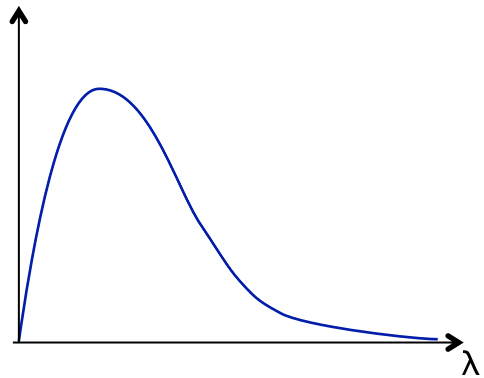
$P(x \mid \lambda)$



Poisson distribution

+

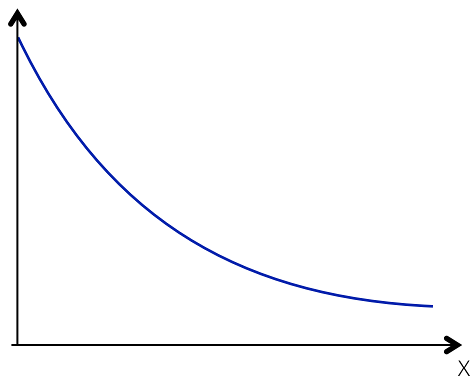
$g(\lambda \mid r, \alpha)$



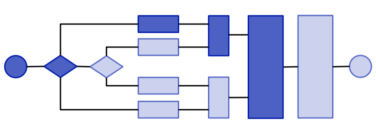
Gamma distribution

$$P(x \mid r, \alpha) = \int P(x \mid \lambda) g(\lambda \mid r, \alpha) d\lambda$$

=



Negative binomial distribution



Step 2: Identify the observable individual-level behavior of interest: Dropout process

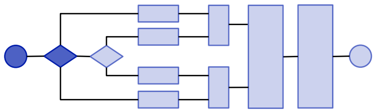
How high can we expect the transaction level to be in the next period?

Transaction process:

How often each customer purchases.

Dropout process:

The propensity of a customer to dropout.



Step 2: Identify the observable individual-level behavior of interest: Dropout process

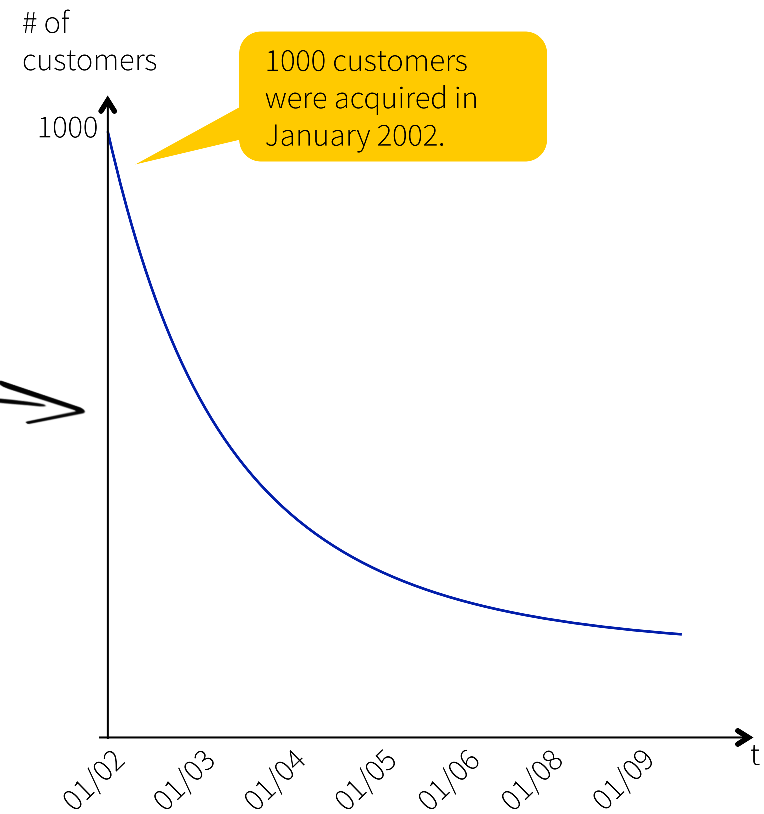
Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
iI151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data

Step 2: Identify the observable individual-level behavior of interest: Dropout process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
iI151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data

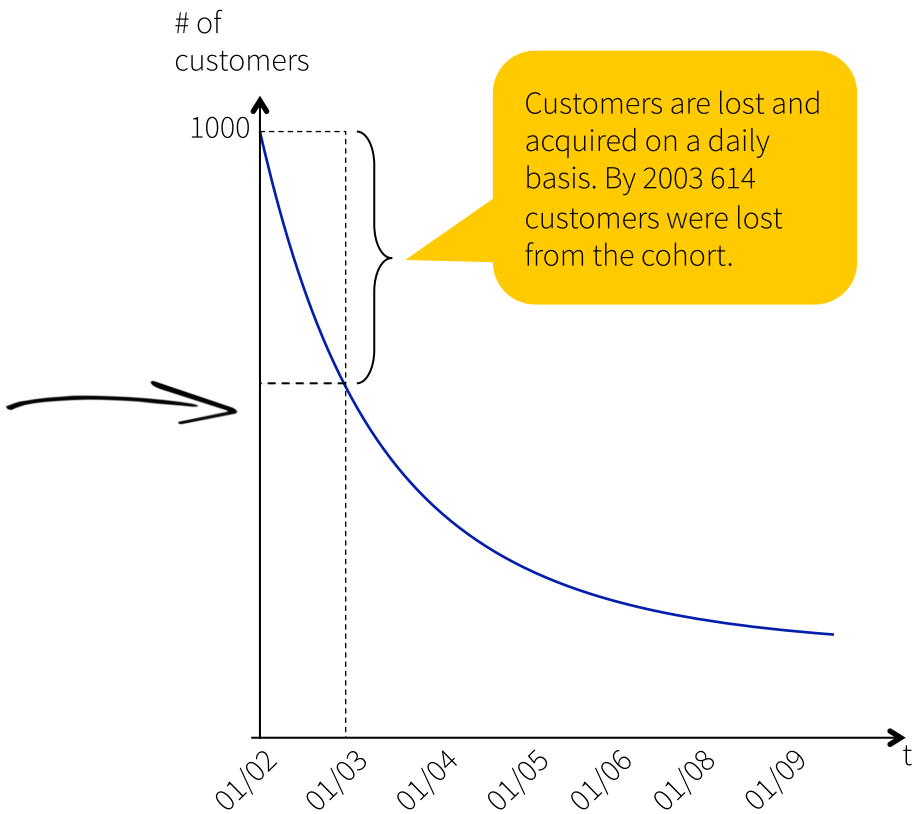


Cohort view

Step 2: Identify the observable individual-level behavior of interest: Dropout process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
il151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data

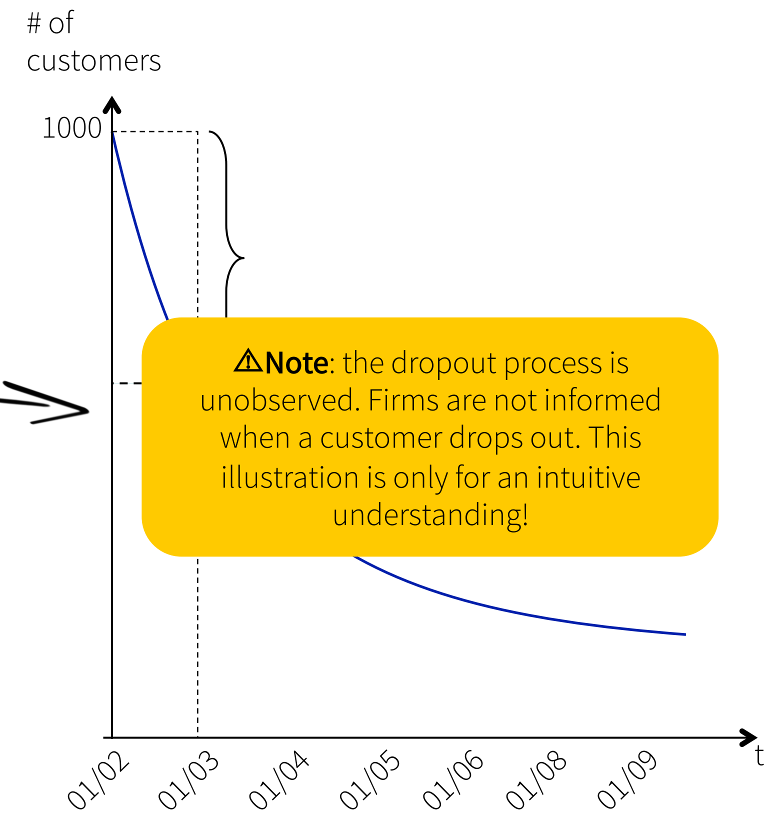


Cohort view

Step 2: Identify the observable individual-level behavior of interest: Dropout process

Customer ID	Number of transactions	Time of last transaction
ba366-lk	2	21.7
dq722-nZ	4	21.5
hW866-Qe	7	21.5
wp63i-Zu	10	21.5
ue109-qc	2	21.5
jw272-zF	0	0.00
EJ370-tG	1	21.4
iI151-ze	0	0.00
CC434-pv	1	21.3
PJ291-jL	0	0.00
eh52a-vo	1	21.3
...	...	...

Transformed data



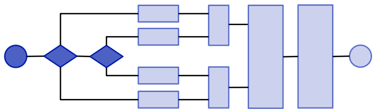
⚠**Note**: the dropout process is unobserved. Firms are not informed when a customer drops out. This illustration is only for an intuitive understanding!

Cohort view

Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	(Continuous) **Timing** (Discrete)	Choice
How many?	When? How long?	Whether or not? Which?

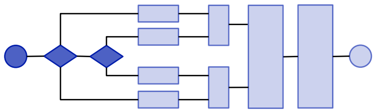


Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	**(Continuous)Timing** (Discrete)	**Choice**
How many?	When? How long?	Whether or not? Which?

"How long is the customers relationship with the firm?"



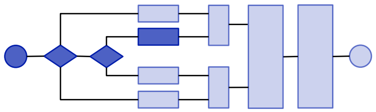
Step 3: Select a probability distribution that characterizes this individual-level behavior

What question do we want to answer?

Counting	(Continuous) Timing (Discrete)	Choice
How many?	When? How long?	Whether or not? Which?

"How long is the customers relationship with the firm?"

Phenomenon	Individual-level	Heterogeneity	Model
Timing (continuous) e.g. days	exponential	gamma	Pareto type II / EG



 Luca Gaegauf | luca.gaegauf@access.uzh.ch

Step 3: Select a probability distribution that characterizes this individual-level behavior

- A customer has an unobserved lifetime (length t) with the firm, which is distributed by the dropout rate μ.
- The duration of the customer's relationship with the firm is characterized by:

$$S(t \mid \mu) = e^{-\mu t}$$

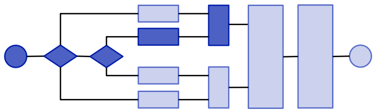
The dropout process characterized by the **exponential distribution**.

Fader & Hardie (2012)

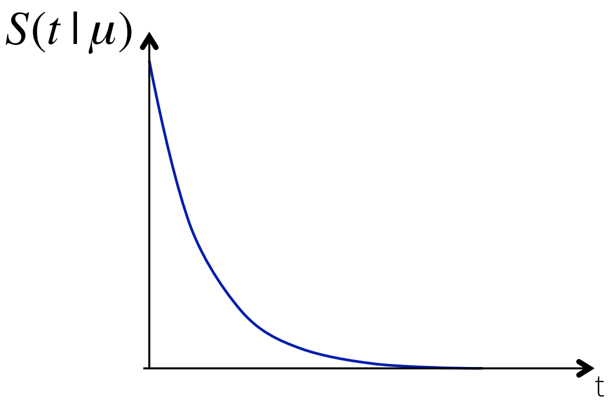
Step 4: Specify the distribution of the latent trait variables

The propensity of a customer to dropout varies between customers.

Phenomenon	Individual-level	Heterogeneity	Model
Timing (continuous) e.g. days	exponential	gamma	Pareto type II / EG

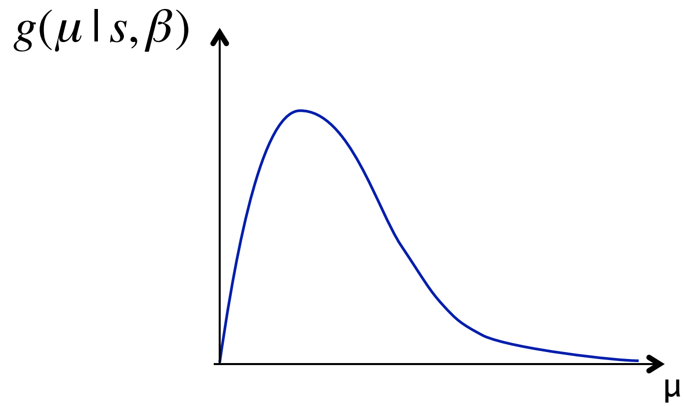


Step 5: Derive the corresponding aggregate or observed distribution for the behavior of interest



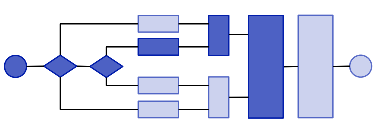
Exponential distribution

The dropout process characterized by the exponential distribution.

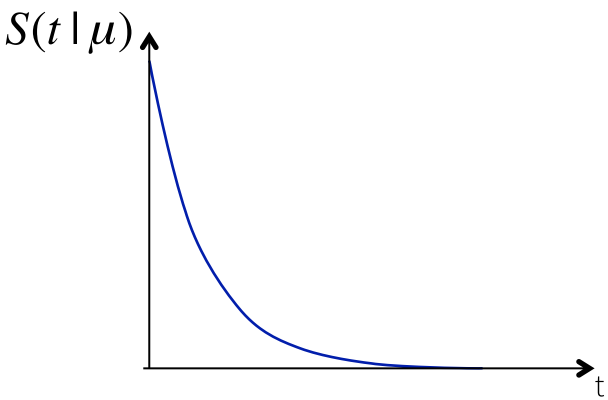


Gamma distribution

The latent traits are characterized by a gamma distribution .

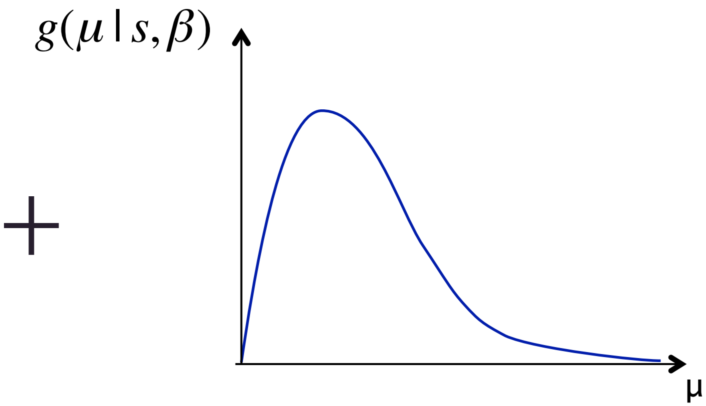


Step 5: Derive the corresponding aggregate or observed distribution for the behavior of interest



Exponential distribution

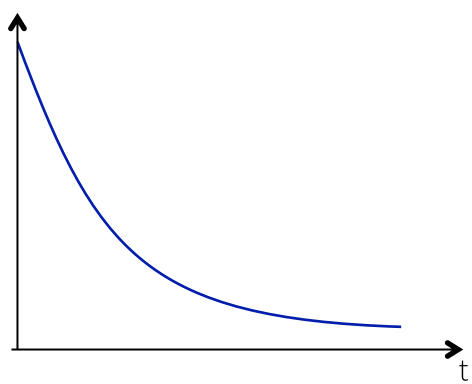
+



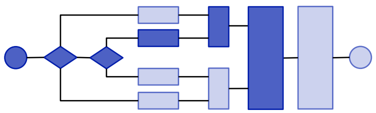
Gamma distribution

$$S(t \mid s, \beta) = \int S(t \mid \mu) g(\mu \mid s, \beta) \, d\mu$$

=



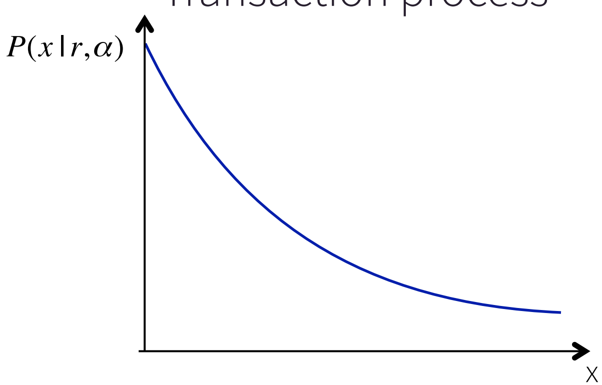
Pareto type II / exponential gamma



NBD, and Pareto type II model the transaction, and the dropout process

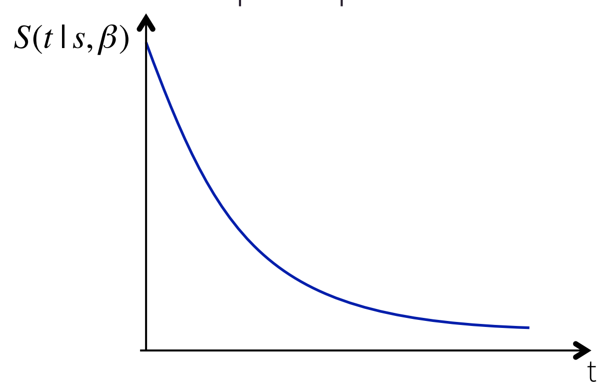
How high can we expect the transaction level to be in the next period?

Transaction process



Negative binomial distribution

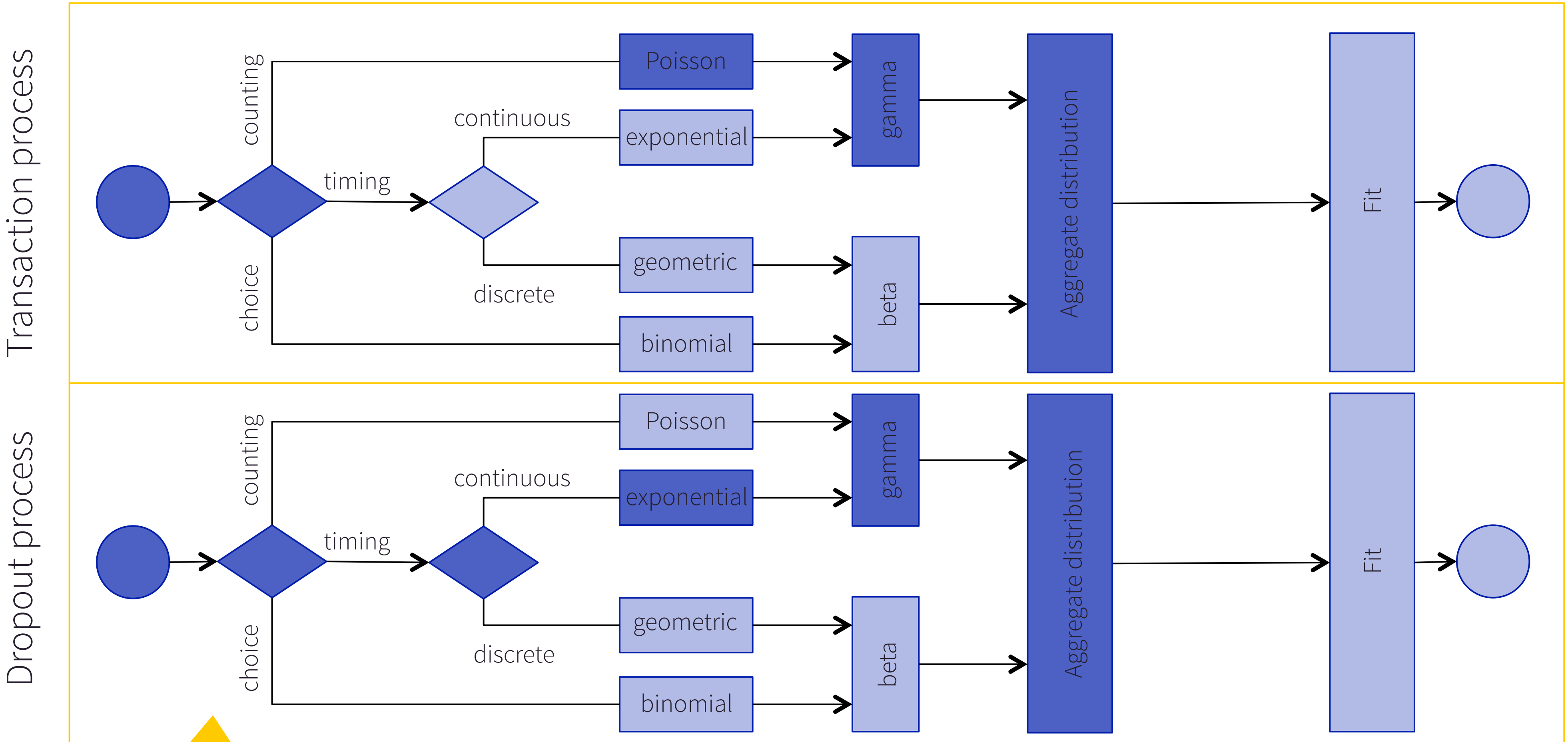
Dropout process



Pareto type II

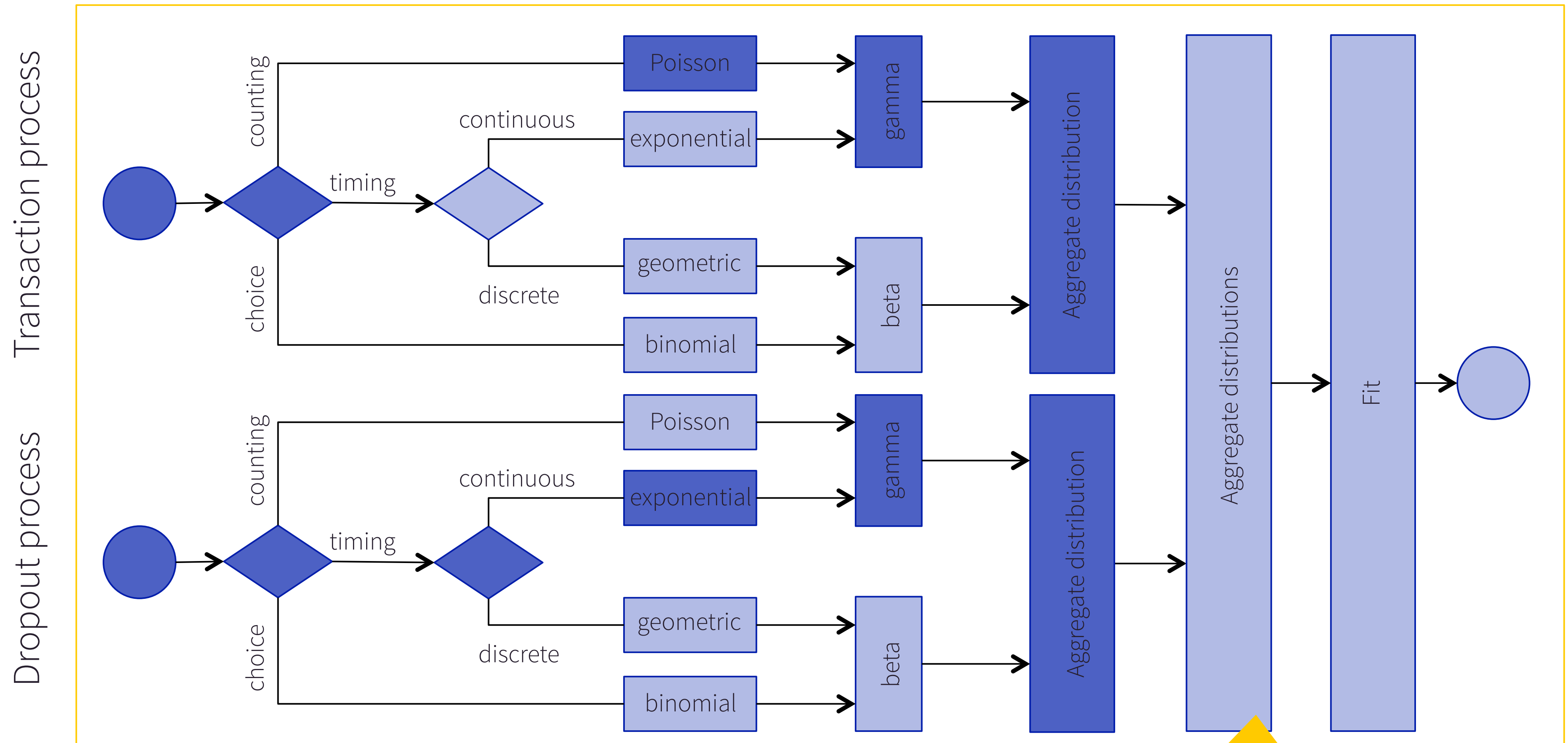
Luca Gaegauf | luca.gaegauf@access.uzh.ch

NBD, and Pareto type II model the transaction, and the dropout process…



We complete the steps for each model separately.

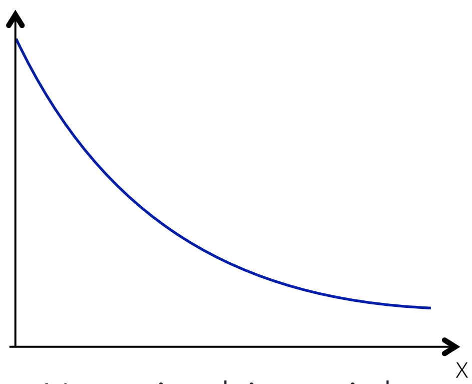
… combine the models to receive the Pareto/NBD



Before fitting the model, we aggregate the distributions.

Step 5.5: Aggregate the distributions to receive the Pareto/NBD model

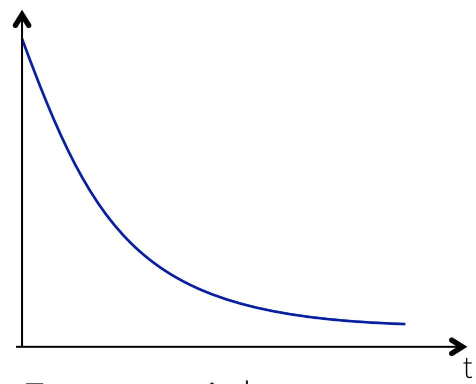
$P(x \mid r, \alpha)$



Negative binomial distribution

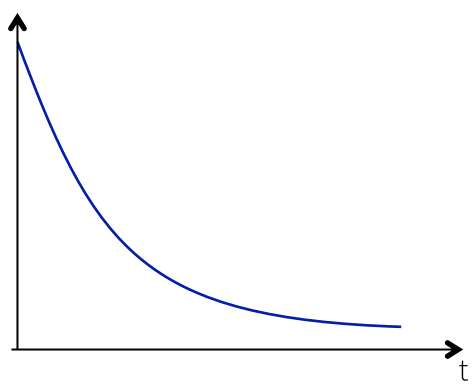
$+$

$S(t \mid s, \beta)$



Exponential gamma

$=$

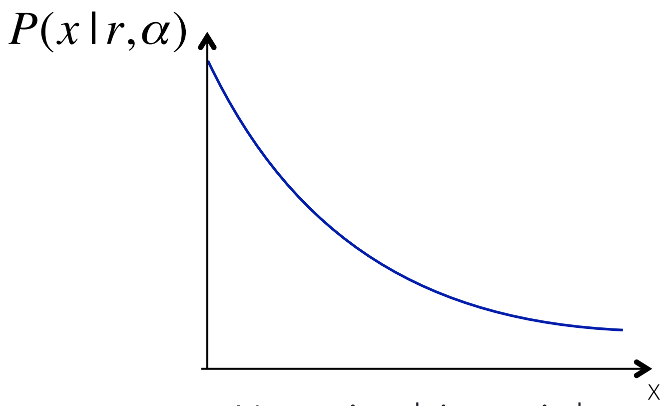


Pareto/NBD

$$P(active \mid r, \alpha, s, \beta, X = x, t_x, T)$$
$$= \int \int P(active \mid \lambda, \mu, X = x, t_x, T) S(\lambda, \mu \mid r, \alpha, s, \beta, X = x, t_x, T) \, d\lambda \, d\mu$$

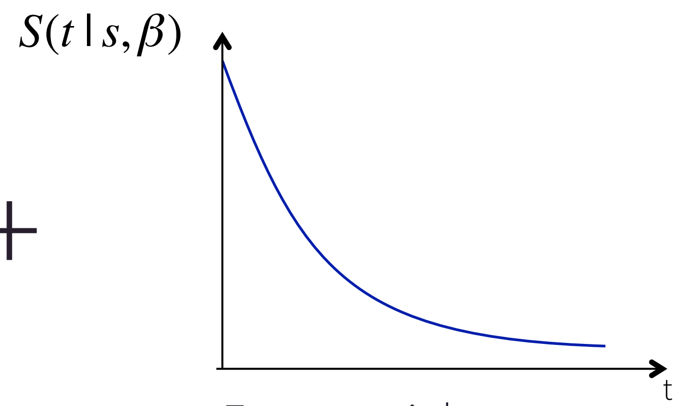
Guo, et al. (2013)

Step 5.5: Aggregate the distributions to receive the Pareto/NBD model



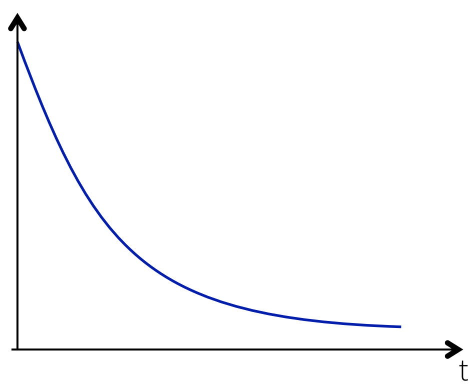
Negative binomial distribution

+



Exponential gamma

=



Pareto/NBD

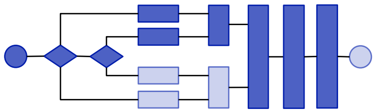
x: number of transactions observed in time interval (0, T].
t_x: time of last purchase.
(see slides 16-19)

$$P(active \mid r,\alpha,s,\beta,X=x,t_x,T)$$
$$=\int\int P(active \mid \lambda,\mu,X=x,t_x,T)\,S(\lambda,\mu \mid r,\alpha,s,\beta,X=x,t_x,T)\,d\lambda\, d\mu$$

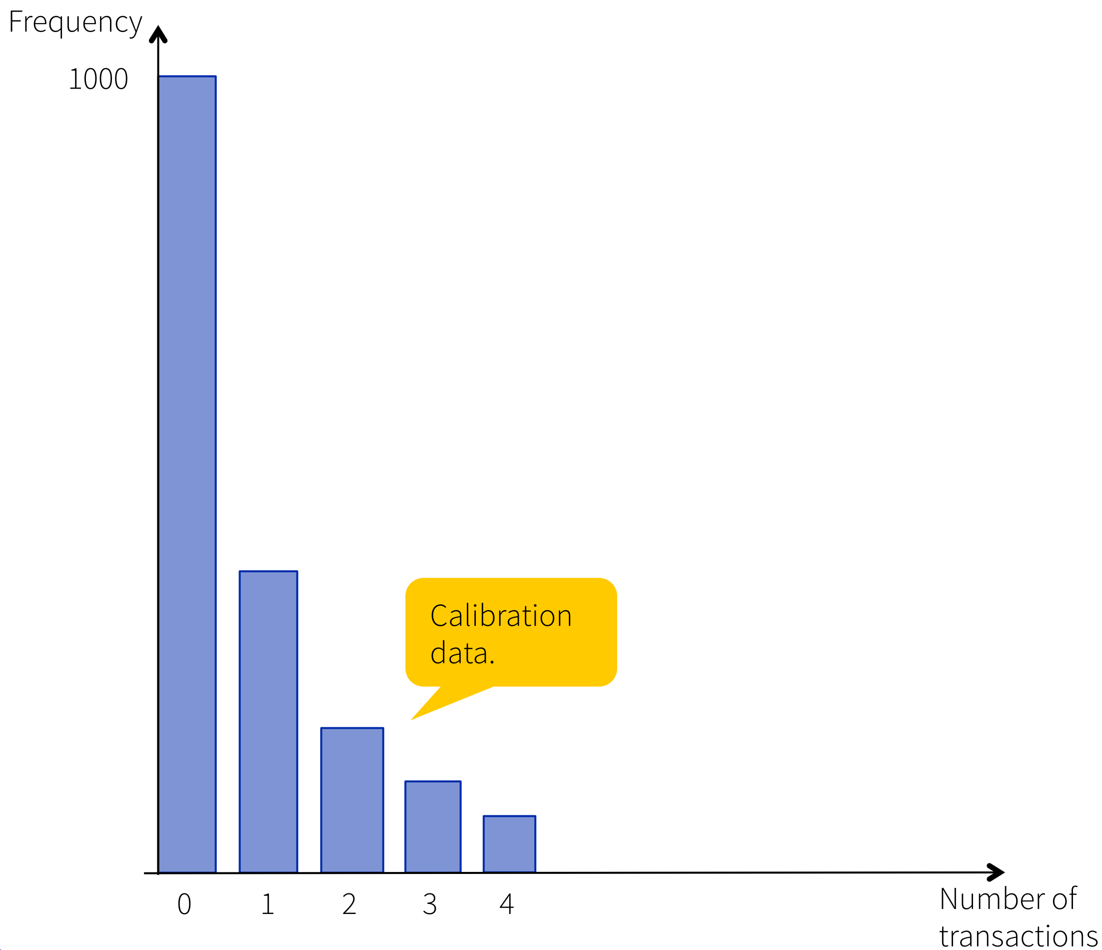
Guo, et al. (2013)

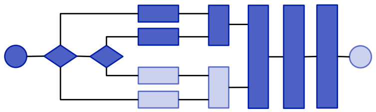
Step 6: Estimate parameters of the mixing distribution

- Estimate the parameters of the aggregated distribution by fitting it on the observed data.
- Using log-likelihood minimize the difference between the data and the curve.

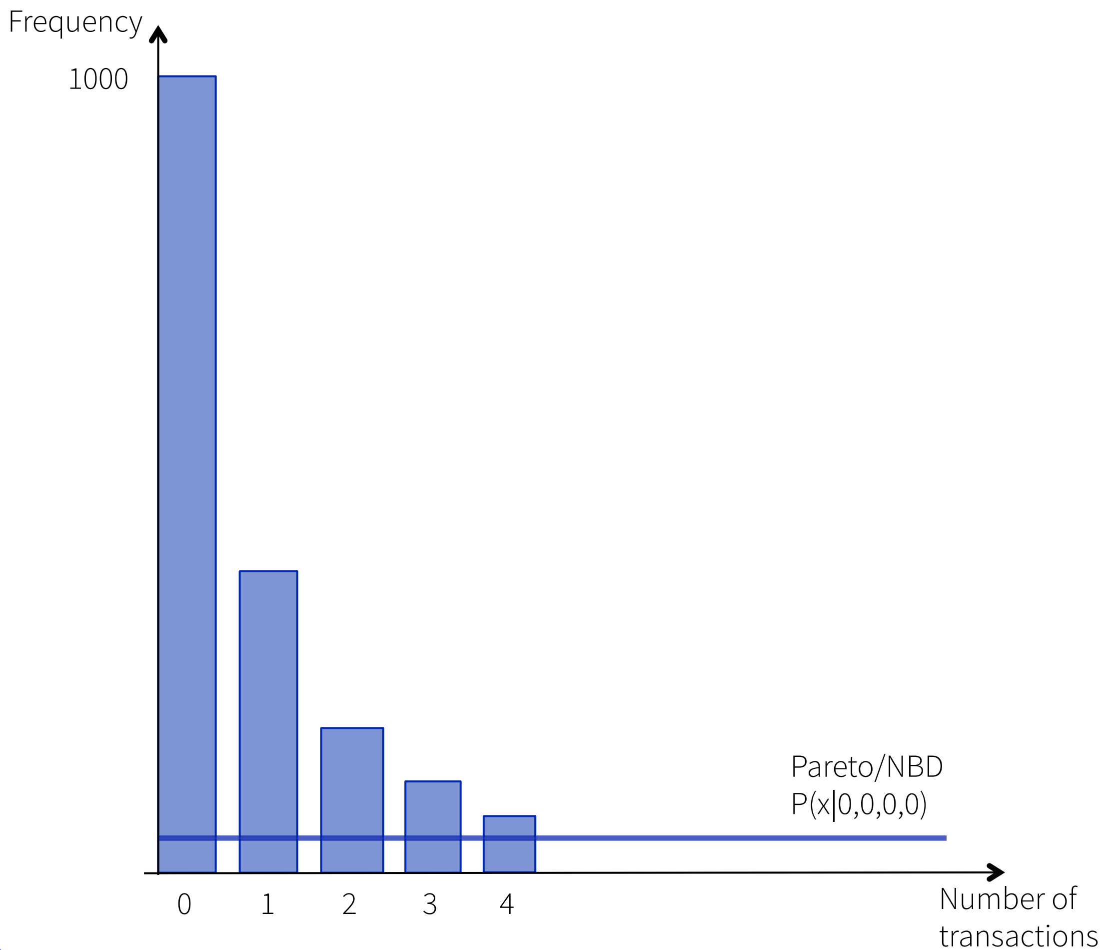


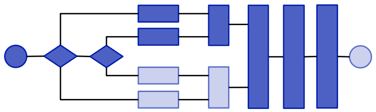
Step 6: Estimate parameters of the mixing distribution



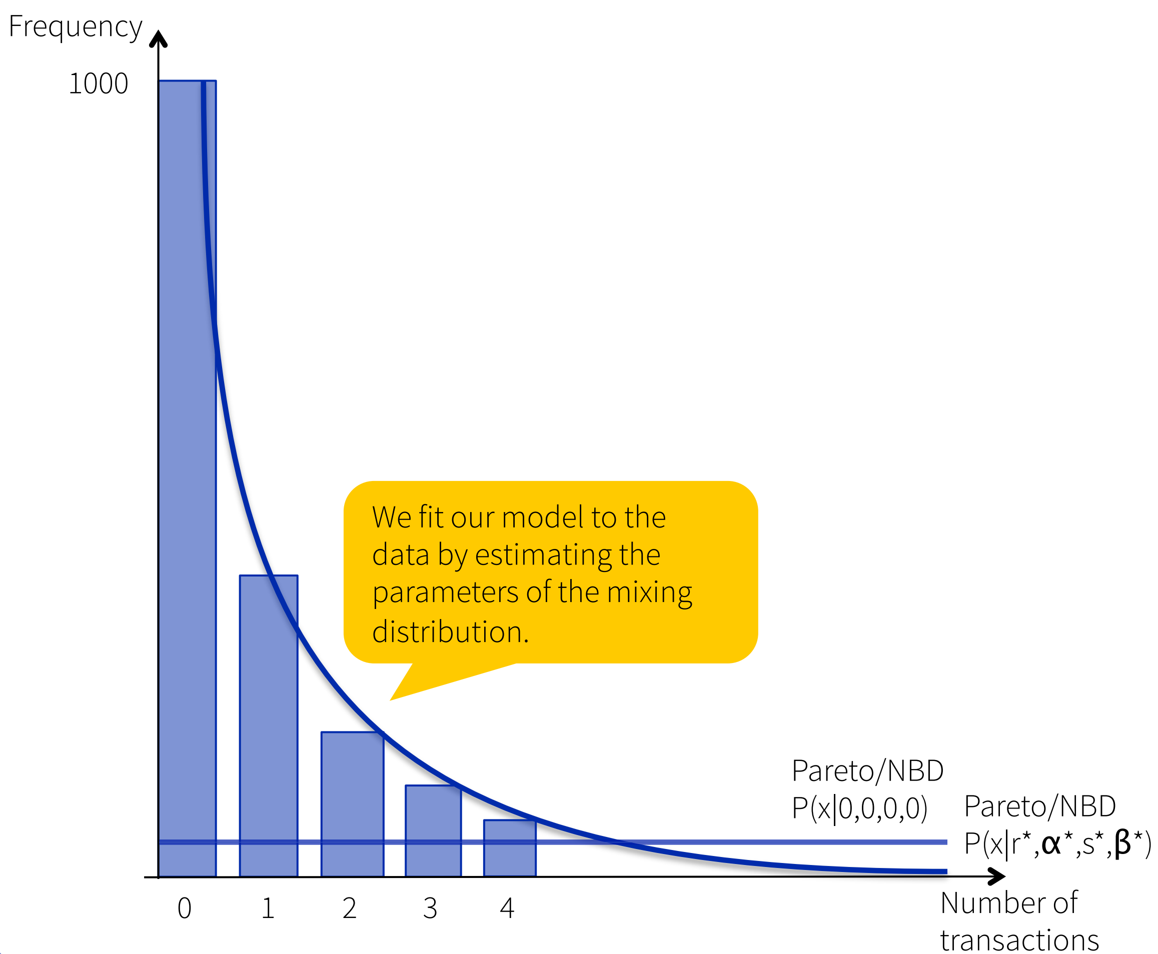


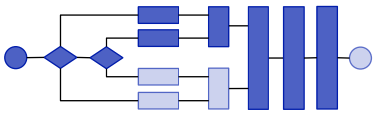
Step 6: Estimate parameters of the mixing distribution



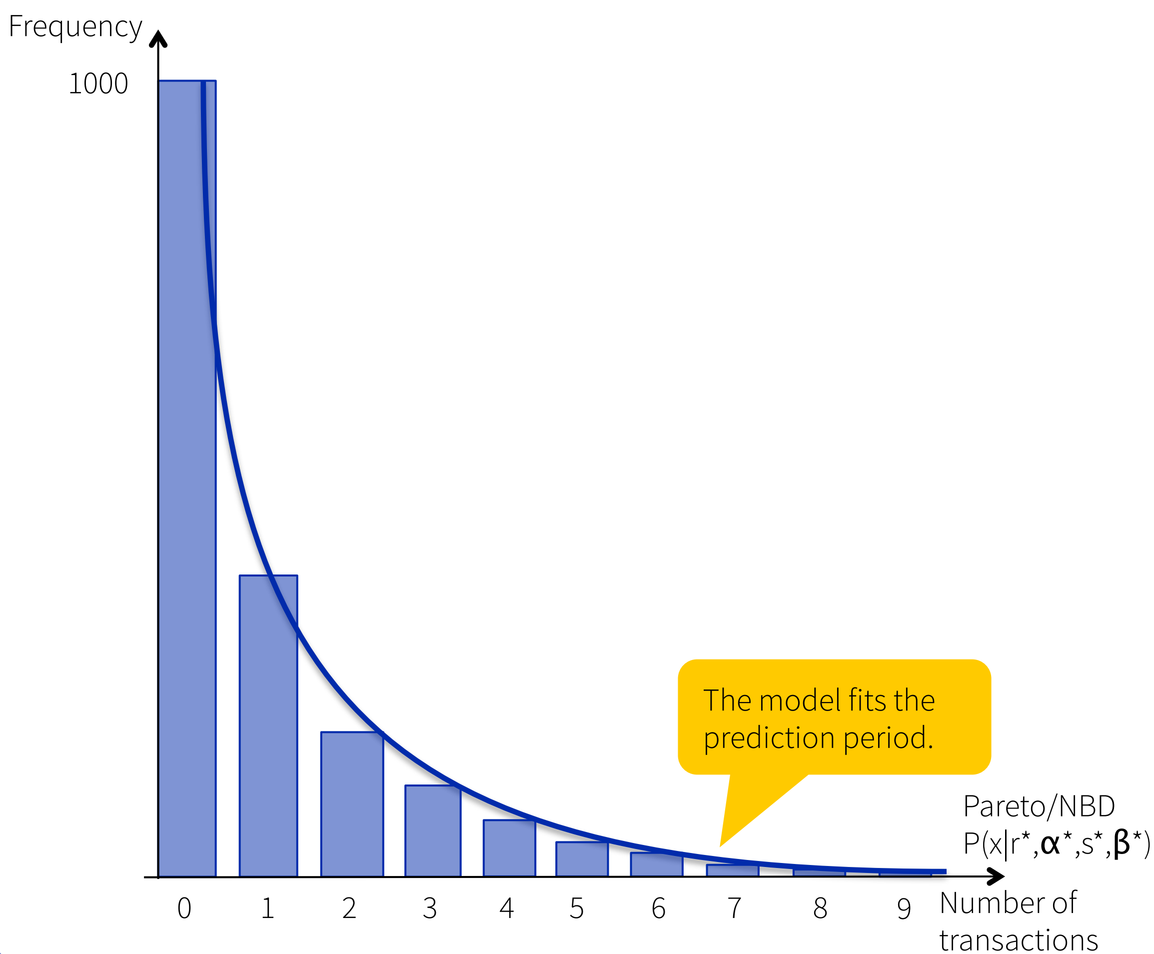


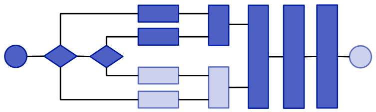
Step 6: Estimate parameters of the mixing distribution



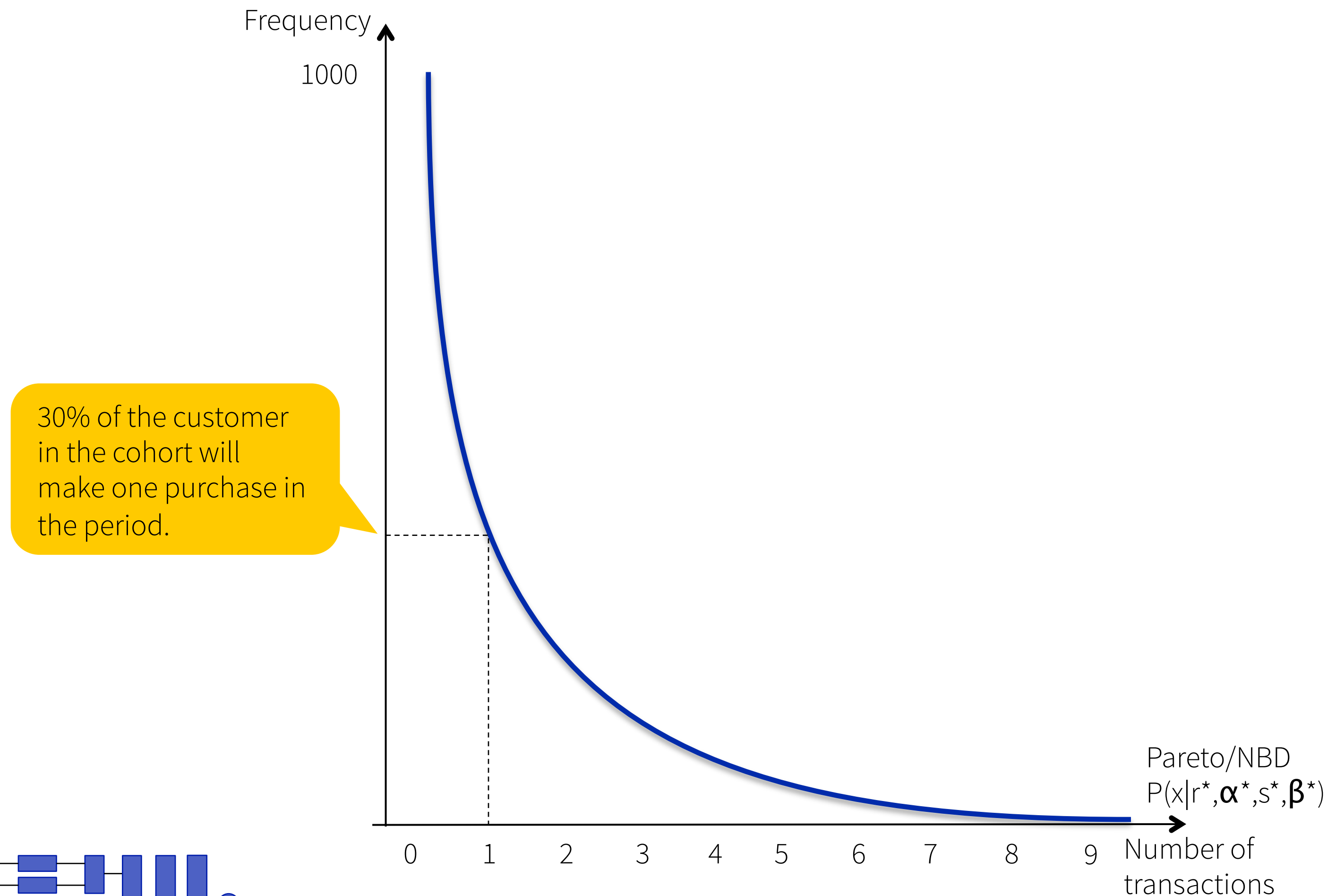


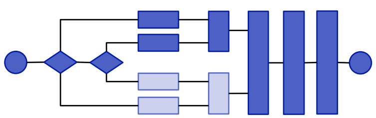
Step 6: Estimate parameters of the mixing distribution





Step 7: Solve marketing problem
Transaction level of future periods





Step 7: Solve marketing problem
Further applications

- Using the Pareto/NBD model determine:
  - The expected number of transaction in the time interval (0, T].
  - The probability that an individual with observed behavior (x, t_x, T) is still alive at time T.
  - The expected number of transaction in the future period (T, T+t] for an individual with observed behavior (x, t_x, T).

Fader & Hardie (2012)

Agenda

❶ Probability models for discrete non-contractual settings

❷ Probability models for continuous non-contractual settings

❸ **References**

References

- Fader, P. S., & Hardie, B. G. (2012). Reconciling and Clarifying CLV Formulas.
- Guo, Y., Wang, H., & Liu, W. (2013). Improved Pareto/NBD Model and Its Applications in Customer Segmentation based on Personal Information Combination. *International Journal of Database Theory & Application*, *6*(5).

Luca Gaegauf – Student

University of Zurich

☎ +41 79 725 60 63

🖂 E-Mail: luca.gaegauf@access.uzh.ch