A New Inventory Classification Criterion for Retail CRM

Hideo Tashiro*
Masataka Ban**
*Graduate School of Business Science, University of Tsukuba
**University of Tsukuba (ban.masataka.fn@u.tsukuba.ac.jp)

Jan. 20, 2020

(@ University of Tsukuba)

Contents

1．Introduction

A lot of retailers implements Reward Card, Loyalty program, and CRM program.

Customer relationship management (CRM)

=> corporate strategy that maximize long-term profits from current customers through activities to build long-term, good relationships with them.

Especially in a business environment having high competitiveness and difficulty in differentiation such as retail, maintaining current customers is more important than acquiring new ones.

=> Retailers that implement CRM find good customers within purchase history data, give them special treatment.

The long-term profitability of a customer is measured by customer lifetime value (LTV) index.

customer LTV… sum of cumulated future cash flows from a customer.

Kumar and Peterson (2012)
retailer which implements CRM has to make decisions in order to maximize customer LTV.

To increase the customer LTV,
determine adequate marketing mix,
promote multi-channel shopping options to customers,
promote up-selling and cross-selling to customers,
and so on.

=> Optimal resource allocation are needed for LTV management

Retail marketing mix ("7P" concept)

Retail mix…managerial decision tool of retailer.

- Product : Assortment, Categories, Brands, Service,…
- Place : Location of stores, Operating hours, Space,…
- Price : Margins, Price emphasis, Mark-down policy,…
- Promotion : Advertising, Display, Personal selling,…
- Process : Uniformity of offering, Service delivery,…
- People : Internal marketing, Organization culture,…
- Physical evidence : Facilities, Infrastructure, …

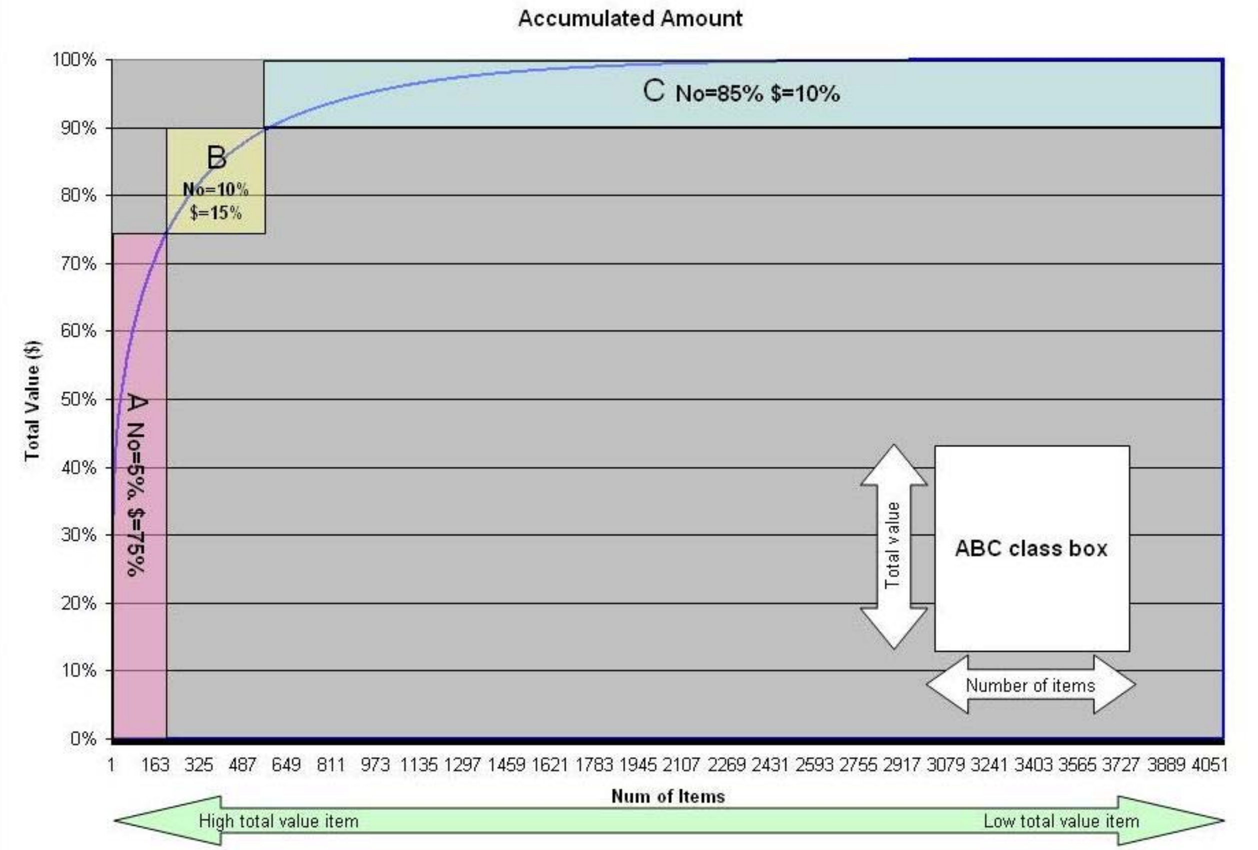
=> maximize the sales, profit, repeat, customer LTV.

=> minimize the (opportunity) loss.

It's necessary to evaluate the marketing strategy by such metrics.

Inventory goods classification

<u>ABC analysis</u>…divides inventory items into three categories. according to a priority of some criteria, such as sales volume, average unit cost, annual dollar usage, lead time and so forth.



(From Wikipedia "ABC analysis".)

Inventory classification criteria

Li et.al.(2019, International Transactions In Operational Research.) ,
"Multicriteria ABC inventory classification using acceptability analysis".

Lolli, Ishizuka, Gamberini (2014, Int. J. Production Exonomics),
"New AHP-based approaches for multi-criteria inventory classification".

Ramanathan (2006, Comuters & Operations research),
"ABC inventory classification with multiple-criteria using weighted linear optimization".

=>Many previous studies in this field are concerned with classification methods (not the criteria or metrics).

Typical criteria showed in Ramanthan (2006)..
sales volume, average unit cost, annual dollar usage, lead time,…

We believe that some long-term evaluation criterion of goods inventory are needed for retail stores,
such as customer LTV.

Customer Lifetime Value (LTV)

customer LTV… sum of cumulated future cash flows
from a customer.

Some methods to estimate it find in the literature,
using accounting data, aggregated sales per customer, …

As for retail CRM with non-contractual setting (which is typical)

-> they can get a contract day of their CRM program with a customer,
but cannot observe the dropout (churn) day of same customer.

->In order to measure LTV, it is necessary to predict the day of dropout
in his/her whole life.

-> studies that measure individual customer LTV from purchase history data
with non-contractual setting
Schmittlein, et al. (1987)… Parate/NBD model
Abe (2009, and a series of his study)… Hierarchical Bayes Model

Objectives

To propose a new criterion/metric to classify the goods.
It is consistent to customer LTV.

In more detail,
to generate such the metric, we construct a model
that the customer LTV is distributed to their purchased goods
via purchase topics (purchase motivation).

Then, regression coefficients in the model indicate,
how much does a goods increase a customer LTV.

Assumption…
a goods that a profitable customer purchases
leads to a high performance for store profit.

2. Model

a.overview (data and models)

Two kinds of data are derived from a scan panel data with customer ID (from a Japanese pharmacy retail store)

・Scan Panel data with customer ID
=> customer individual purchase history data

~~・DTM(Document Term Matrix)~~
・CGM(Customer Goods Matrix)
=> Latent Dirichlet Allocation model (Topic model)
-> purchase topics (purchase motivation) are generated

・RFM (Recency, Frequency, Monetary data by customer)
=> estimate customer LTV by the model of Abe(2009)
-> customer LTV

We use purchase topic data estimated by LDA model as an explanatory variable in LTV model.

・Scan panel data with ID

ID	Store	YMD	Cash Desk	Receipt	Time zone	Item code	Volume	Price	unit price
1134	20	20070711	103	3211	11	4901301463111	1	298	298
1134	20	20070729	108	2443	10	4901301463111	1	298	298
1134	20	20070729	108	2443	11	4901301463111	1	298	298
1134	20	20070729	108	2443	15	4901301732149	3	780	260
1134	20	20070804	104	8537	12	4901301463111	1	298	298
1134	20	20070804	104	8537	12	4901301463111	1	298	298
1134	20	20070804	104	8537	12	4901301281463	2	2040	1020
1134	20	20070907	108	1113	14	4901301281463	1	298	298
1134	20	20070918	103	7823	19	4901301732149	1	260	260
1147	20	20071014	107	7856	14	4901301237132	1	228	228
112	20	20061105	105	7172	20	4901301463111	1	378	378
112	22	20061231	107	5838	13	4903301020981	1	298	298
112	20	20070103	104	3167	20	4901301463111	1	398	398
112	20	20070103	104	3167	20	4901301761958	2	210	105
112	20	20070103	104	3167	20	4903301020981	4	2120	530
112	20	20070103	104	3167	20	4901301281463	1	398	398
112	22	20070206	106	9088	11	4901301281463	1	298	298
112	20	20070210	103	1199	20	4901301281463	1	398	398

=> When, (where,) how many, how much
does a customer purchase it?

・RFM data : extract three types of variables from scan panel.

ID	Latest	Freq	Monetary
1	2008/6/9	4	4,572
2	2008/9/25	9	19,181
3	2008/12/10	7	25,508
4	2008/11/27	9	19,544
5	2008/12/26	24	46,486
6	2008/11/30	3	13,760
7	2008/12/28	9	12,085
8	2008/12/24	8	19,335
9	2008/12/29	24	51,772
10	2008/10/16	2	2,317
11	2008/12/30	73	249,084
12	2008/12/30	10	29,398
13	2008/1/3	13	94
14	2008/12/21	15	77,936
15	2008/12/26	107	96,820
16	2008/12/16	12	15,967

Recency (R)…# of days since last purchase
Frequency (F)…# of times of purchase
Monetary (M)…average spending for one purchase

=>find profitable customer by clustering
based on these metrics.
=>depends on time period of analysis.
=>not future profitability of customer.

・CGM data : purchase rate in a period of analysis

Item core

ID	4902430203111	4901301463111	4901301732149	4901301281463	4901301237132	4903301020981	4901301761958
1	0.031	0.007	0.125	0.122	0.000	0.020	0.027
2	0.000	0.049	0.000	0.000	0.134	0.000	0.131
3	0.000	0.000	0.000	0.000	0.076	0.000	0.015
4	0.000	0.000	0.000	0.000	0.000	0.065	0.000
5	0.135	0.000	0.000	0.000	0.025	0.000	0.000
6	0.070	0.000	0.000	0.000	0.116	0.011	0.000
7	0.095	0.000	0.000	0.000	0.016	0.132	0.000
8	0.000	0.000	0.120	0.000	0.000	0.129	0.000
9	0.128	0.000	0.003	0.000	0.035	0.111	0.000
10	0.032	0.098	0.000	0.000	0.084	0.040	0.000
11	0.120	0.000	0.000	0.000	0.000	0.070	0.000
12	0.000	0.021	0.000	0.068	0.000	0.103	0.000
13	0.063	0.012	0.048	0.000	0.000	0.000	0.125
14	0.000	0.000	0.054	0.109	0.000	0.000	0.000
15	0.098	0.000	0.000	0.000	0.000	0.074	0.116
16	0.102	0.122	0.000	0.000	0.066	0.000	0.051

=> customers ×
(purchased) goods matrix.
=> massive zero

b. purchase topic model (Latent Dirichlet Allocation)

Topic model…traditionally, used as a text mining tool, to discover hidden semantic structures in a text body.

Topic model in marketing science field

Jacobs, B.J.D., B. Donkers and D. Fok (2016)
"Model-Based Purchase Predictions for Large Assortment",
Marketing Science, 35 (3), 389-404.

=> applied extended LDA model to purchase history data

Toubia, O., et al.(2019)
"Extracting Features of Entertainment Products: A Guided Latent Dirichlet Allocation Approach Informed by the Psychology of Media Consumption",
Journal of Marketing Research, 56, 18-36.

Tirunillai, S. and G.J. Tellis (2014)
"Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation"
Journal of Marketing Research, 51, 463-479.

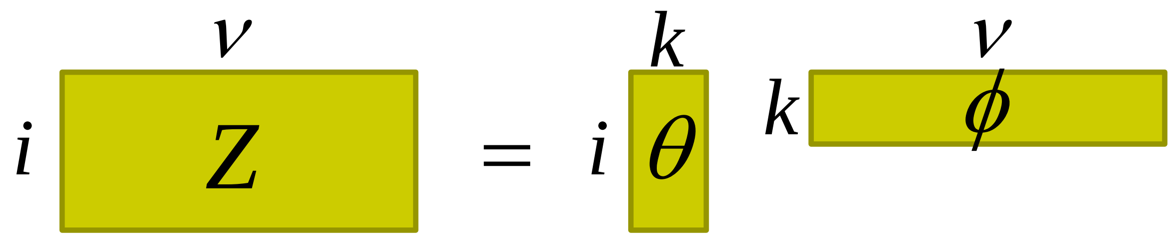
=> applied the model to text data

Z : CGM data consisting of occurrence probabilities of goods per customer in his/her purchase history.

A probability of goods *v*'s occurrence in purchase data of customer *i* is

$$p(v \mid i) = \sum_{k=1}^{K} p(v \mid k) p(k \mid i) = \sum_{k=1}^{K} \phi_{v|k} \theta_{k|i}$$

$$\left\{ \begin{array}{ll} p(v \mid k) = \phi_{v|k} & \text{: probability of goods } v\text{'s occurrence in topic } k \ (k = 1,\dots,K) \\ p(k \mid i) = \theta_{k|i} & \text{: probability of topic } k\text{'s occurrence in purchase data of customer } i \end{array} \right.$$



c.Customer LTV model (Abe 2009, Marketing science)

RFM analysis

=> depends on time period of analysis.
=> not the future profitability of a customer.

Research Motivation

=> not depend on time period of analysis.
=> evaluate future profitability of a customer.

Overview of the model

R (recency)
F (frequency)
M (monetary value)

Model

P (purchase rate)
L (lifetime)
S (spending)

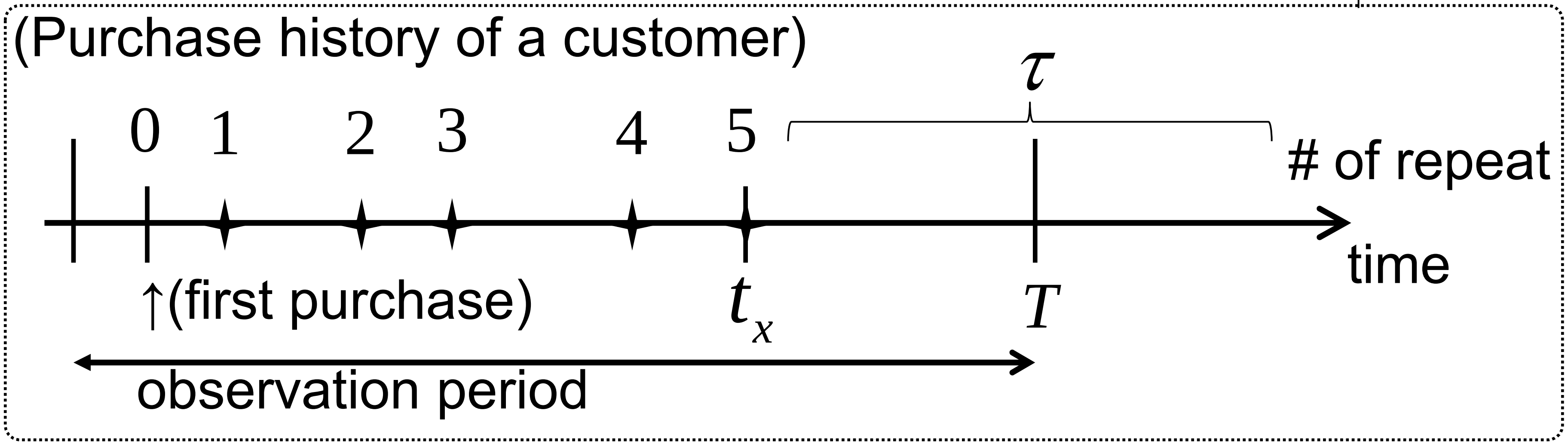
Customer LTV

Observations
(depend on time period)

Behavior Traits
(attribute)

Data and notations

X : RFM data consisting of three variables, "Recency", "Frequency", and "Monetary" data per customer



x : # of repeat purchase (Frequency)

T : end of the observation period

t_x : day of last purchase in the period

$T - t_x$: elapsed days since last purchase (Recency)

S_x : spending of each purchase (Monetary)

τ : unobserved customer lifetime (parameter)

Model assumption

Individual behavior :

A1 While active, each customer makes purchases according to a Poisson process.

->Purchase randomly occurs independently of past purchases.

A2 Each customer remains active for a lifetime, which has an exponentially distributed duration.

->Dropout because of store switch, moving house, death,.. is randomly occurs.

Heterogeneity across customers:

A3 Individual's above parameters follow a multivariate log-normal distribution.

->It has traditionally attracted great interest in marketing.

The LTV model for a customer

Purchase rate (P) : # of purchase until dropout day τ

$$p(x|\lambda)=\begin{cases}\dfrac{(\lambda T)^x}{x!}e^{-\lambda T} & \text{if } \tau > T,\\[2ex] \dfrac{(\lambda \tau)^x}{x!}e^{-\lambda \tau} & \text{if } \tau \le T,\end{cases} \qquad x=0,1,2,\ldots,$$

λ : purchase rate

Lifetime (L) : the time until a customer dropout

$$f(\tau)=\mu e^{-\mu\tau} \qquad \tau \ge 0,$$

μ : dropout rate

Spending (S) : expected spending per purchase

$$\log(S_x) \sim N(\log(\eta),\omega^2), \qquad S_x > 0$$

We can get PLS which is not dependent on observation period, although RFM is dependent on it.

LTV value of customer *i*

Customer specific parameters

λ_i : purchase rate

μ_i : dropout rate

η_i : expected spending per purchase

(get a value by data augmentation in MCMC)

τ_i : (unobserved) dropout day

LTV: LTV of customer *i* (d is annual discount rate)

$$\mathrm{LTV}_i = \frac{\lambda_i \eta_i e^{\omega^2/2}}{\mu_i + \log(1+d)}$$

Hierarchical model

Three parameters (PLS) are regressed on customer attributes and purchase topics.

$$\begin{array}{l} \text{P:} \\ \text{L:} \\ \text{S:} \end{array} \begin{bmatrix} \log(\lambda_i) \\ \log(\mu_i) \\ \log(\eta_i) \end{bmatrix} \sim \mathrm{MVN}(\beta d_i, \Gamma_0),$$

$$d_i = (\text{intercept}, \text{customer attributes}, \text{purchase topics})$$

$$\theta_i^T = \{\theta_{k|i}\}$$

(*T* means transposed matrix)

Now, decompose β to three parts corresponding to the data.

$$\beta = (\beta_{\text{intercept}}, \beta_{\text{attributes}}, \beta_{\text{topics}})$$

$\beta_{\text{topics}}\theta_i^T$: customer *i*'s PLS that stems from his/her purchased goods

$\underline{Z^T}\theta\beta_{\text{topics}}^T$: allocate customer PLS to goods via topics (*goods LTV*, use this for inventory evaluation)

CGM

d. Bayesian estimation

Bayes' theorem $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$,

A, B : some probabilistic events

$P(A|B)$: conditional probability of A occurring given B.

$P(A)$: probability of observing A

A : a model (parameter) which represents an event

B : data of the event

$P(A|B)$: the likelihood of the model A under the data B
=> posterior probability, posterior distribution

$P(B|A)$: the likelihood of the data B occurring under the model A => likelihood

$P(A)$: the likelihood of the model A without data B
=> prior probability, prior distribution
=> a belief about the model analyst has on ahead.

Utility of Bayesian estimation

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

In the Bayesian estimation, we evaluate the posterior distribution of the model, not the likelihood of it.

$P(A|B)$ $P(B|A)$

Even if we do not have enough sample size for the likelihood estimation, prior information provide us a stable inference of posterior. $P(A)$

RFM data is cross sectional data.

=> individual parameters are estimated by only one sample.

e.Model summary

Scan panel data (from a retail store)

Z :CGM(Customer Goods Matrix)

- LDA model

  θ :customer-topic co-occurance probability matrix

X :RFM data

- customer LTV model

  λ, μ, η :customer PLS and LTV

- hierarchical model (prior)

  β :effectiveness of purchase topics to PLS

- allocate customer CLV(PLS) to goods via topics

  $Z^T \theta \beta^T_{\text{topics}}$: goods-PLS matrix

- calculate goods LTV from goods PLS

f. Markov chain Monte Carlo (MCMC) strategy

M_1: Latent Dirichlet Allocation model (Topic model)
=> collapsed Gibbs sampling (by "LDA" R package)

M_2: customer LTV model (Abe, 2009 and subsequent papers)
=> Hamiltonian Monte Carlo (by STAN)

What we really want to do…

$$p(M_1, M_2) = p(M_1 \mid M_2)\, p(M_2)$$

Today, we will show empirical results in the following slides…

$$p(M_1 \mid E(M_2)), \quad p(M_2)$$

3. Empirical study

a.Data (scan panel data with ID from a retail store)

・pharmacy store chain.
・not only marketed drugs, but also daily necessities, grocery, cosmetics, and so on.
・operate a reward card system.
・Period: one year (from Jan. 1st to Dec. 31th in 2008, daily data)

・Number of customers :400 (households)
・Number of goods : 525 goods (SKU)
=> The size of CGM matrix is 400 × 525.

・Average Recency: 6 days ago (max 29 days ago)
・Average times of shopping:37.47 per customer
・Average spending per shopping: 2,679 yen (about $25)
=>they may be profitable customers…

・Sex: 89.5% female
・Average age: 48.07

b.LDA model for CGM data

As a prior analysis, *K* (# of topics)=19 is estimated by hierarchical Dirichlet process - LDA model.

topic

goods

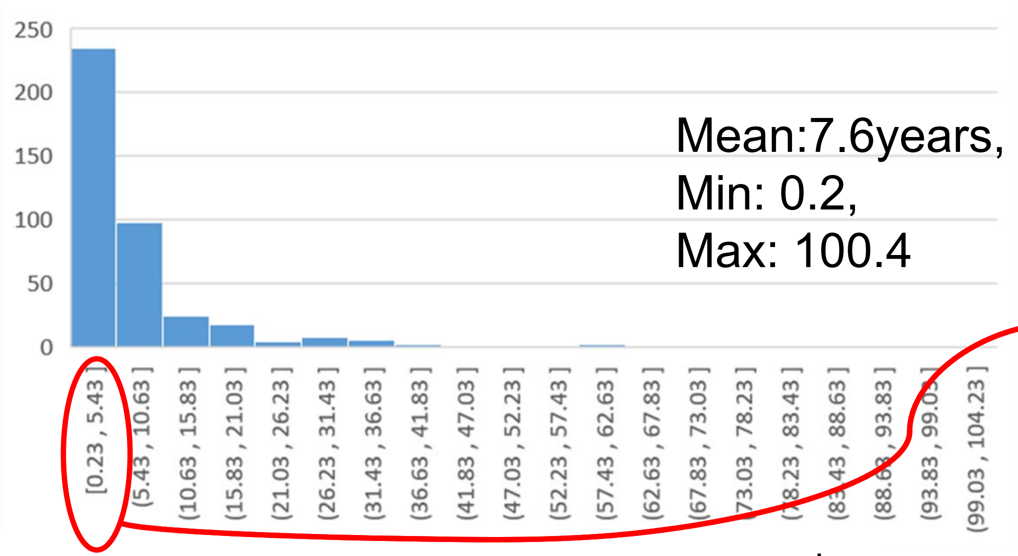
goods code	topic1	topic2	topic3	topic4	topic5	topic6	topic7	topic8	topic9	topic10	topic11	topic12	topic13	topic14	topic15	topic16	topic17	topic18	topic19
0000049716955	0.0000	0.0119	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4901008047072	0.0000	0.0119	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4901451161394	0.0000	0.0000	0.0000	0.0047	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4901872344130	0.0000	0.0000	0.0000	0.0071	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4903333051991	0.0000	0.0000	0.0000	0.0071	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4954097702407	0.0000	0.0000	0.0118	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987138131256	0.0000	0.0000	0.0118	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987305031914	0.0000	0.0119	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4997090154323	0.0000	0.0000	0.0118	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4902951700014	0.0000	0.0326	0.0000	0.0000	0.0061	0.0000	0.0000	0.0000	0.0213	0.0469	0.0151	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4901033731120	0.0000	0.0000	0.0059	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987167035723	0.0000	0.0000	0.0059	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987241130191	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0141	0.0000	0.0000	0.0000
4902105030592	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0085	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4984142219031	0.0000	0.0089	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4902667149428	0.0000	0.0000	0.0059	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0413	0.0000	0.0000	0.0000	0.0000
4901372401517	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0064	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4903347070018	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0064	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987035085614	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0064	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4571243118013	0.0000	0.0000	0.0000	0.0000	0.0000	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4958181028572	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0129	0.0000	0.0000
4987164120606	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0168	0.0000
4987188123126	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0042	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4975497805635	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0088	0.0000	0.0000	0.0000
4972422023508	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0140	0.0000
4902430962940	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0020	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0046
4987336506979	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0168	0.0000
4902380135845	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0069
4511413302422	0.0000	0.0000	0.0000	0.0000	0.0000	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4901070711758	0.0000	0.0000	0.0000	0.0000	0.0000	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987774030005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0052	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4902471052815	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0036	0.0000	0.0000	0.0000	0.0000
4902508160575	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0036	0.0000	0.0000	0.0000	0.0000
4903301020981	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0068	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4902522661966	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0032	0.0000	0.0000
4901234210813	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0070	0.0000	0.0000	0.0000
4903326130405	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4987067281503	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4987072069912	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4987138322043	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0112	0.0000
4987234171040	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4987426002015	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4991820120459	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0056	0.0000
4903110166429	0.0000	0.0000	0.0000	0.0000	0.0000	0.0105	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4987045056109	0.0000	0.0000	0.0000	0.0000	0.0000	0.0105	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
0400900000368	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0081	0.0000	0.0000

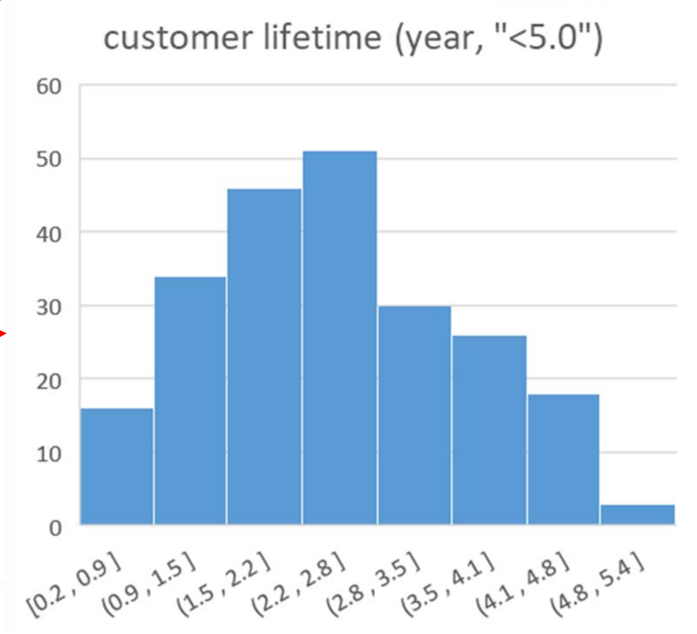
Red colored: over 1/525 (1/ # of goods)

Interpretation of the topics is so difficult…

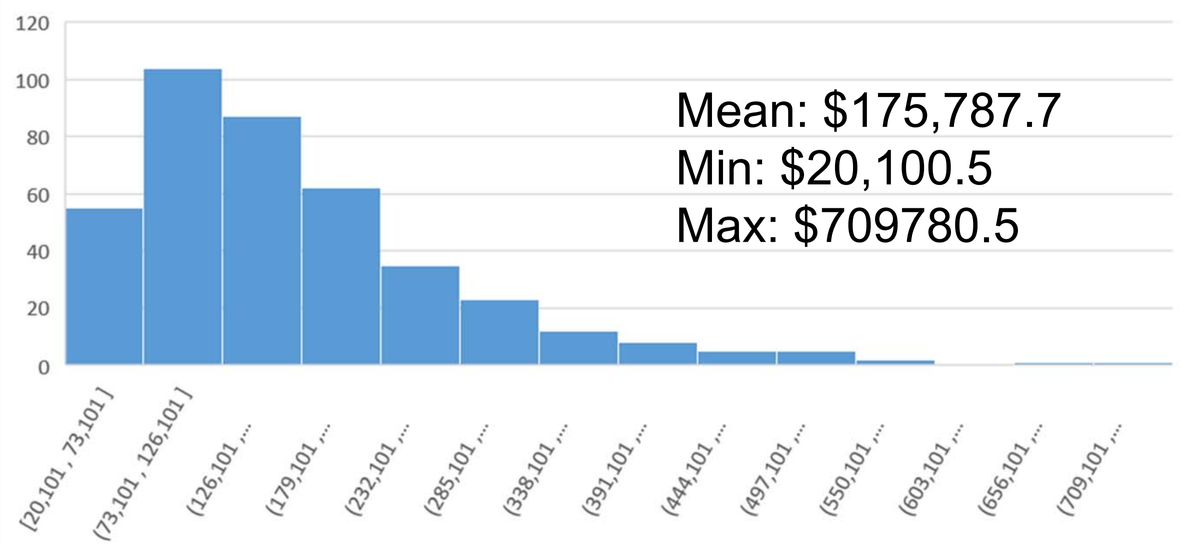
c.Customer LTV model for RFM data

- Histogram of customer Lifetime $(1/\mu_i)$ (year)





- Histogram of customer LTV ($)



Estimation result of $\beta = (\beta_{\text{intercept}}, \beta_{\text{attributes}}, \beta_{\text{topics}})$

(P) purchase rate log(λ)		mean	X2.5	X97.5
	sex(female=1)	0.25	−0.08	0.59
	log(age)	0.00	−0.26	0.27
	topic1	−0.20	−1.49	1.04
	topic2	−0.58	−1.77	0.58
	topic3	−0.42	−1.55	0.70
	topic4	0.04	−1.14	1.21
	topic5	−0.27	−1.62	0.93
	topic6	−0.37	−1.51	0.74
	topic7	−0.32	−1.50	0.86
	topic8	−0.18	−1.37	1.01
	topic9	−0.46	−1.61	0.61
	topic11	0.00	−1.16	1.18
	topic12	−0.41	−1.52	0.65
	topic13	0.26	−0.95	1.42
	topic14	0.06	−1.24	1.33
	topic15	−0.38	−1.53	0.69
	topic16	−0.33	−1.50	0.76
	topic17	−0.28	−1.41	0.81
	topic18	−0.52	−1.61	0.51
	topic19	2.64	1.04	4.30
	intercept	−0.27	−1.45	0.82

(L) churn rate log(μ)		mean
	sex(female=1)	−0.41
	log(age)	0.02
	topic1	−0.67
	topic2	−0.74
	topic3	−1.59
	topic4	−0.93
	topic5	−1.34
	topic6	−1.40
	topic7	1.42
	topic8	−2.49
	topic9	−1.79
	topic11	−0.54
	topic12	−0.13
	topic13	−2.95
	topic14	−0.89
	topic15	−2.71
	topic16	−2.85
	topic17	−1.25
	topic18	−0.19
	topic19	−2.68
	intercept	−15.38

(S) Spending log(η)		mean
	sex(female=1)	0.03
	log(age)	0.01
	topic1	−0.18
	topic2	0.26
	topic3	0.07
	topic4	−0.03
	topic5	0.02
	topic6	−0.05
	topic7	0.01
	topic8	0.33
	topic9	−0.22
	topic11	−0.51
	topic12	0.05
	topic13	−0.60
	topic14	−0.04
	topic15	0.11
	topic16	0.23
	topic17	0.08
	topic18	0.04
	topic19	0.01
	intercept	7.42

=> A lot of estimates are not statistically significant.

=>Only topic 19th has positive effect to purchase rate.

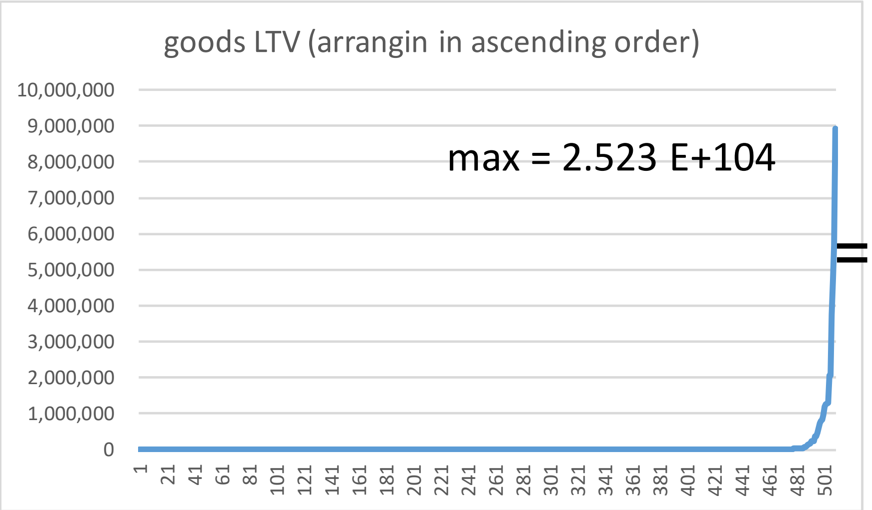
Topic no.19 : goods list (with high probability)

=>Goods in this topic induce high purchase rate.

topic	goods ID	probability	goods name	category
19	4901301211361	0.103	ハミング　濃縮タイプ　レギュラー　替　５４０ｍｌ	fablic softner liquid
19	4962307070024	0.069	ビタアルト３０００　１００ｍｌ×１０本	vitamin drink
19	4987306003491	0.067	大正製薬　リポビタンＤ　１００ｍｌ×１０	vitamin drink
19	4908522051202	0.046	ＮＩＤ　緑茶ティーバッグ　５ｇ×５０	instant tea
19	4987603426955	0.041	ポケットコール　６０包	injury treatment
19	4960085213305	0.037	東洋　キズテープ　Ｎｏ．４　１サイズ　１００枚	injury treatment
19	4904760010148	0.032	金ちゃん　焼そば　復刻版　カップ　７７ｇ	instant noodle
19	4582151171816	0.030	ジャスティス　エクシード　烏龍茶　２Ｌ	tea
19	4901820335517	0.030	パスコ　笑顔の食卓　８枚	bread

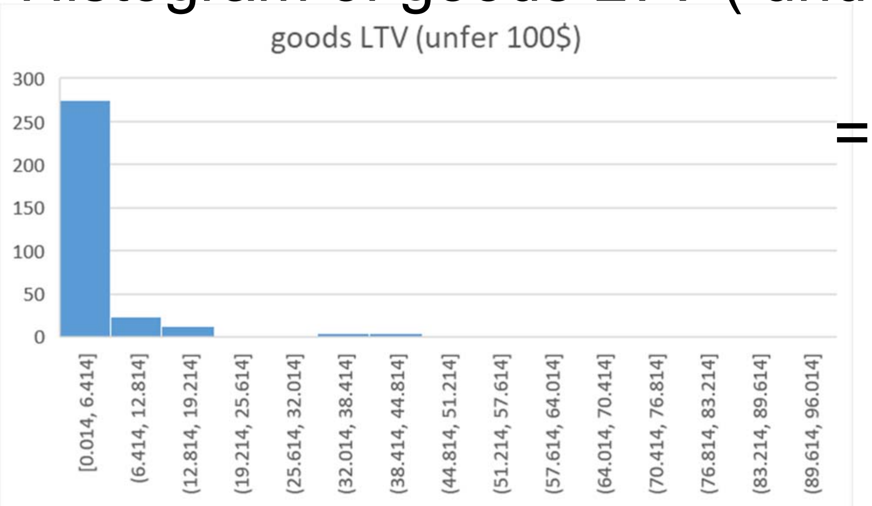
d. goods LTV (PLS) metrics

goods LTV = $Z^T \theta \beta^T_{\text{topics}}$



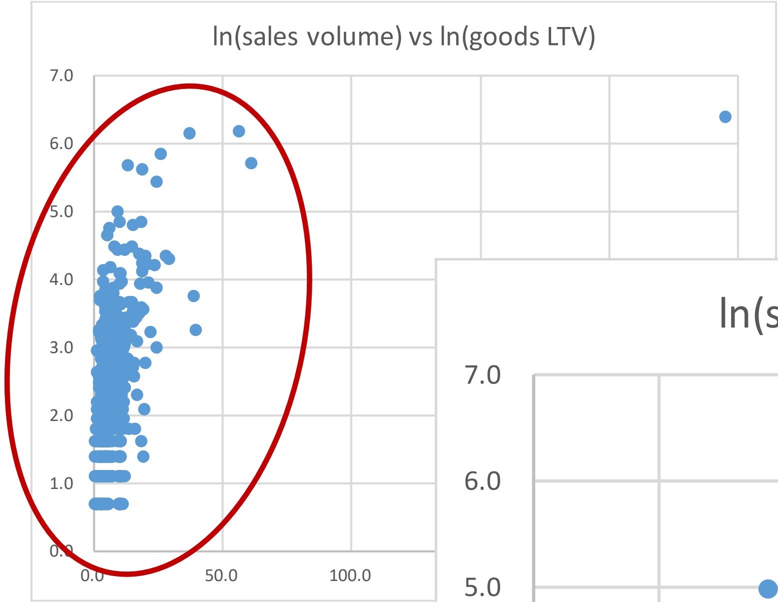
=>there are few extremely massive values.

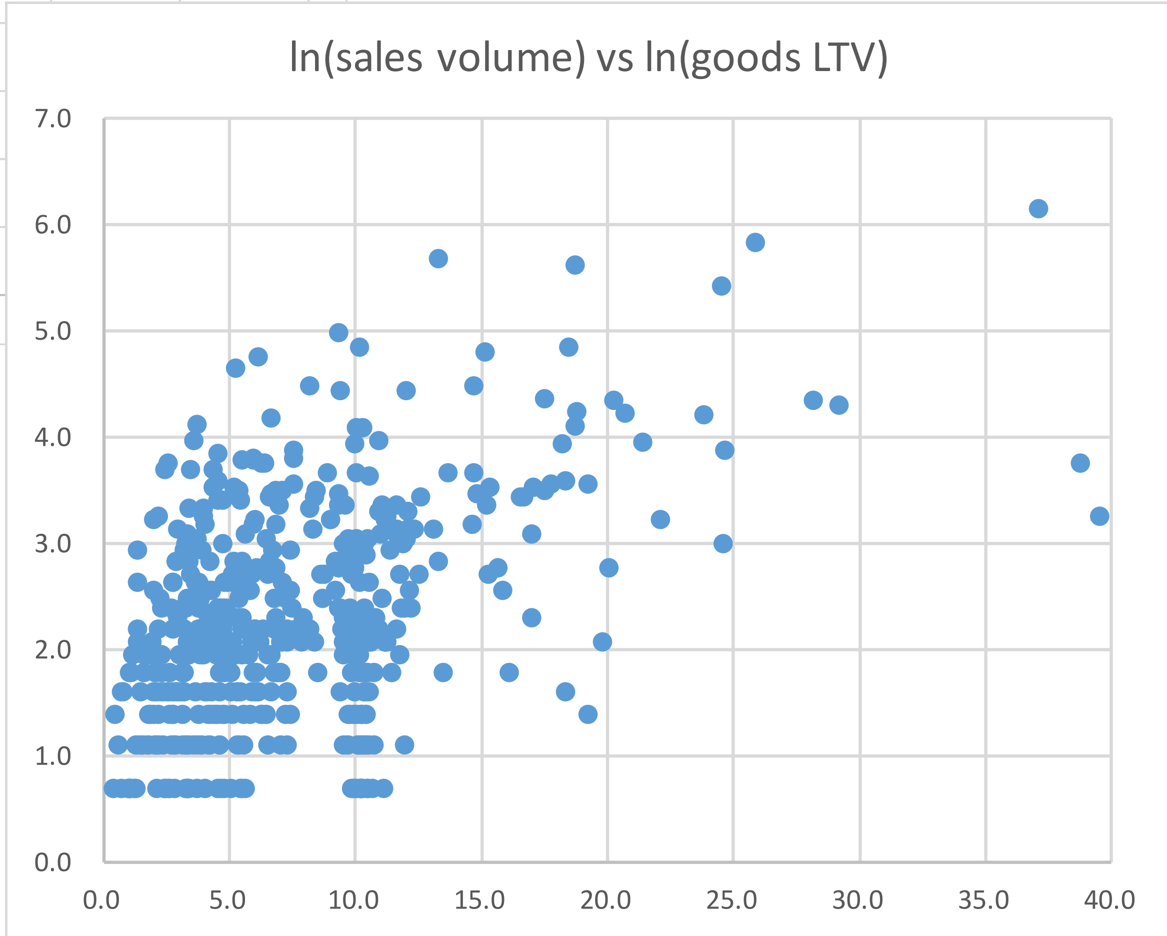
Histogram of goods LTV (under $100)



=>more than half of goods have under $100.

Relation between sales volume and goods LTV





4. Summary and Limitation

Summary

・Propose a new metric for inventory classification criteria.
・Extract CGM and RFM from scan panel data.
・LDA model and LTV model are applied to them respectively.
・Proposed model allocate customer LTV to goods via topics.
・Empirical results using pharmacy retail store data are shown.

Limitation

・We need a lot of improvements…
・(estimation)…Simultaneous estimation is needed.
・(LDA)…Refined model reflecting the structure of data is needed.
ex. goods hierarchy…category-brand-SKU

References：

Abe, M. (2009) “Counting Your Customers” One by One: Hierarchical Bayes Extension to the Pareto/NBD Model, *Marketing Science*, 28(3), 541-553.

Jacobs, B.J.D., B. Donkers and D. Fok (2016) “Model-Based Purchase Predictions for Large Assortment”, Marketing Science, 35 (3), 389-404.

Kumar, V. and J.A.Petersen(2012) *Statistical models in Customer Relationship Management*, First Edition . WILEY, U.K..

Li et.al.(2019) “Multicriteria ABC inventory classification using acceptability analysis”. International Transactions In Operational Research, 26, 2494-2507

Lolli, Ishizuka, Gamberini (2014) “New AHP-based approaches for multi-criteria inventory classification”, Int. J. Production Economics, 156, 62-74.

Ramanathan (2006) “ABC inventory classification with multiple-criteria using weighted linear optimization”, Comuters & Operations research, 32, 695-700.

Tirunillai, S. and G.J. Tellis (2014) “Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation”, Journal of Marketing Research, 51, 463-479.

Toubia, O., et al.(2019) “Extracting Features of Entertainment Products: A Guided Latent Dirichlet Allocation Approach Informed by the Psychology of Media Consumption”, Journal of Marketing Research, 56, 18-36.

Series of Abe’s customer LTV model (In Japanese)

阿部誠 (2008)「消費者行動理論にもとづいた個人レベルのRF分析ー階層ベイズによる Pareto/NBD モデルの改良ー, 日本統計学会誌, 第37巻, 239-259.

阿部誠(2011),「RFM指標と顧客生涯価値：階層ベイズモデルを使った 非契約型顧客関係管理における消費者行動の分析」,日本統計学会誌,第41巻, 第1号, 51-81.

Thank you for listening.