A Discrete Choice Framework for Modeling and Forecasting The Adoption and Diffusion of New Transportation Services
Feras El Zarwi, Akshay Vij, Joan Walker

EE 290-O / IEOR 290
10/05/2017

Presentation by Cheng-Ju Wu
Outline

• Adoption and Diffusion of New Technologies: Bass model
• Discrete Choice Theory: Logit Model
• A New Model for New Transportation Service

Adoption and Diffusion of New Technologies

- In economics, people study the modeling of **adoption** and **diffusion** of new technologies.
- **Diffusion** is a macro process by which a new idea or new product is accepted by the consuming public.
- **Adoption** is a micro process that focuses on individual consumer in the stage of deciding to accept or reject a new product.
Bass model

- Basic assumption (Bass 1967): “the probability $P_t$ that an initial purchase will be made at $t$ given that no purchase has yet been made is a linear function of the number of previous buyers.”

$$P_t = p + \frac{q}{M} Y(t) \quad Y(0) = 0$$

- $p$: Coefficient of innovation;
- $q$: Coefficient of imitation;
- $M$: Total potential market for the technology;
- $Y(t)$: Cumulative number of individuals that adopted the new technology by time $t$ (number of previous buyers.)

Bass model

\[ P_t = p + \frac{q}{M} Y(t) \]

The probability \( P_t \) is computed as:

\[ P_t = \frac{f(t)}{1 - F(t)} = p + \frac{q}{M} Y(t) \quad f(t) \text{ is the likelihood of purchase at } t \]

\[ F(t) = \int_0^t f(\tau) d\tau , F(0) = 0 \]

Let \( Y(t) = MF(t) \), we can write

\[ f(t) = \frac{d}{dt} F(t) = [p + qF(t)][1 - F(t)] \]

adopted

adopt by innovation

adopt by imitation

not yet adopted
Bass model

\[ P_t = p + \frac{q}{M} Y(t) \]

The probability \( P_t \) is computed as:

\[ P_t = \frac{f(t)}{1 - F(t)} = \frac{Mf(t)}{M - MF(t)} \]

\( f(t) \) is the likelihood of purchase at \( t \)

\[ F(t) = \int_0^t f(\tau) d\tau, F(0) = 0 \]

Let \( S(t) = Mf(t) \) and \( Y(t) = MF(t) \), we have

\[ P_t = \frac{S(t)}{M - Y(t)} = p + \frac{q}{M} Y(t) \]

Then,

\[ S(t) = pM + (q - p)Y(t) - \frac{q}{M} Y^2(t) \]

\( S(t) \) is the sales of a product over time.
Bass model

Since \( F(t) = \int_0^t f(\tau)d\tau \), \( F(0) = 0 \) and \( S(t) = Mf(t) \)

Take derivative of \( Y(t) \) yields

\[
\frac{d}{dt} Y(t) = M \frac{d}{dt} F(t) = M f(t) = S(t) = Y'(t)
\]

\[
S(t) = pM + (q - p)Y(t) - \frac{q}{M} Y^2(t)
\]

The dynamics of \( Y(t) \) satisfied the following **Riccati equation**:

\[
Y'(t) = pM + (q - p)Y(t) - \frac{q}{M} Y^2(t)
\]
The S-shaped curve, $Y(t)$, shows that social imitation growth with time. We can estimate $p$, $q$ and $M$ by short term market data in the beginning, then use the model to forecast long term market behavior.
Discrete Choice Theory

• A decision maker named \( n \) chooses between \( J \) alternatives.
• Each decision would earn \( n \) a utility of \( U_{nj} \) for \( j = 1...J \).

\[
U_{nj} = V_{nj} + \epsilon_{nj}
\]

- utility
- representative utility
- a know distribution

\[
V_{nj} = V(x_{nj}, s_n)
\]

- Observed features of the alternatives
- features of the decision maker

• We want to compute the probability that \( n \) chooses \( i \)

\[
P_{ni} = \Pr(U_{ni} > U_{nj} \text{ for all } j \neq i)
= \Pr(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj} \text{ for all } j \neq i)
= \Pr(\epsilon_{nj} - \epsilon_{ni} < V_{ni} - V_{nj} \text{ for all } j \neq i)
\]

Reference: Roy Dong, EE 290O/IEOR 290 Lecture Note 6
Logit Model

• The probability user $n$ chooses option $i$ over all alternatives is given by

$$P_{ni} = \Pr(V_{ni} + \epsilon_{ni} > V_{nj} + \epsilon_{nj} \text{ for all } j \neq i)$$

• Suppose that $\epsilon_{ni}$ is given and it is iid, Gumbel distribution with $\mu = 0$ and $\beta = 1$, $P_{ni}$ is obtained by integration

$$P_{ni} = \int \left( \prod_{j \neq i} e^{-e^{(V_{ni}+\epsilon_{ni}-V_{nj})}} \right) e^{-\epsilon_{ni}} e^{-\epsilon_{ni}} d\epsilon_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$$

Reference: Roy Dong, EE 290O/IEOR 290 Lecture Note 6
Motivation

• New Transportation Services contains: autonomous vehicles, electric vehicles, ridesharing, carsharing, and many other new technologies.

• However, current travel demand models are unable to predict long-range trends in travel behavior as they do not entail a mechanism that projects membership and market share of new modes of transport (Uber, Lyft, etc).

• Construct a new travel demand models.
Features of New Model

• The proposed model trying to model the following aspects:
  - Technology adoption and use is influenced by socio-demographics,
  - Attributes of the new technology/service,
  - Spatial effect (or network effect) and finally social influences.
Model

• The probability that individual n during time period t after the new technology was available in the market adopted or chose to not adopt could be written as

\[ P(y_{ntj} \mid Z_{nt}, X_{ntj}, q_{ns}) \forall j \in \{0, 1 \mid y_{n(t-1)j}\} \]

\[ j=1 \text{ adopt} \]
\[ j=0 \text{ not adopt} \]

individual n during time period t chose to adopt or not.

characteristics of individual n

an individual n belonging to latent class s

attributes of the new technology
Now, evaluating the probability of adoption or non-adoption will be based on a binary logit formulation that transforms the utility specification into probabilities.

\[ U_{ntj|s} = V_{ntj|s} + \varepsilon_{ntj|s} = z_{nt}^T \beta_s + x_{ntj}^T \gamma_s + \varepsilon_{ntj|s} \]

Noise, i.i.d. Extreme Value Type I distribution with mean zero and variance \( \pi^2/6 \)

\[ V_{ntj|s} \text{ is the systematic utility that is observed by the analyst.} \]

\[
\begin{align*}
V_{\text{adopt}, n, t|s=\text{innovator}} &= z_{nt}^T \beta_1 + x_{ntj}^T \gamma_1 \\
V_{\text{non-adopt}, n, t|s=\text{innovator}} &= 0 \\
V_{\text{adopt}, n, t|s=\text{imitator}} &= z_{nt}^T \beta_2 + x_{ntj}^T \gamma_2 + \Delta_{(t-1)} \alpha_2 \\
V_{\text{non-adopt}, n, t|s=\text{imitator}} &= 0 \\
V_{\text{adopt}, n, t|s=\text{non-adopter}} &= \lambda \\
V_{\text{non-adopt}, n, t|s=\text{non-adopter}} &= 0
\end{align*}
\]
Model

• Add socio-demographic variables to utility function

\[ U_{ni} = V_{ni} + \varepsilon_{ni} = d_{ij} \beta + \ln(size_j) \alpha + Z_n \gamma + X_{nj} \theta + home_n \delta + \varepsilon_{ni} \]

- origin-destination \((i,j)\)
- attractions associated with destination \(j\)
- socio-demographic characteristics of decision-maker \(n\)
- attributes of the new technology at destination alternative \(j\) for individual \(n\)
- dummy variable which will be equal to one if decision-maker \(n\) resides within a certain proximity as his corresponding destination alternative

\(\alpha, \beta, \gamma, \theta, \delta\) are model parameters

• In order to assess the impact of the spatial/network effect of the new technology on the utility of adoption, define the level of accessibility

\[ Accessibility_{n,i,t} = \ln \left( \sum_{j=1}^{J_t} e^{V_{ni,j}} \right) \]

logsum utility function
Model

- The marginal probability $P(y)$ of observing a vector of choices $y$ for all decision-makers is:

$$P(y) = \prod_{n=1}^{N} \sum_{s=1}^{S} P(y_n|q_{ns}) P(q_{ns}|Z_n) = \prod_{n=1}^{N} \sum_{s=1}^{S} P(q_{ns}|Z_n) \prod_{t=1}^{T_n} \sum_{j \in C} P(y_{ntj}|Z_{nt}, X_{ntj}, q_{ns})^{y_{ntj}}$$

**Network effect model**

$$P(y_{ntj}|Z_{nt}, X_{ntj}, q_{ns}) = P(U_{ntj|s} \geq U_{ntj'|s} \forall j' \in C) = \frac{e^{v_{ntj|s}}}{\sum_{j'=1}^{I} e^{v_{ntj'|s}}}$$

**Class membership model**

$$P(q_{ns}|Z_n) = P(U_{ns} \geq U_{ns'} \forall s' = 1,2, \ldots, S) = \frac{e^{v_{ns}}}{\sum_{s'=1}^{S} e^{v_{ns'}}}$$

$$U_{ns} = V_{ns} + \varepsilon_{ns} = z_{nt}^{'} \tau_{s} + \varepsilon_{ns}$$

**Class specific adoption model**

$$P(y_n|q_{ns}) = \prod_{t=1}^{T_n} \prod_{j \in C} P(y_{ntj}|Z_{nt}, X_{ntj}, q_{ns})^{y_{ntj}}$$
The marginal probability $P(y)$ of observing a vector of choices $y$ for all decision-makers is:

$$P(y) = \prod_{n=1}^{N} \sum_{s=1}^{S} P(y_n|q_{ns}) P(q_{ns}|Z_n) = \prod_{n=1}^{N} \sum_{s=1}^{S} P(q_{ns}|Z_n) \prod_{t=1}^{T_n} \prod_{j \in C} P(y_{ntj}|Z_{nt}, X_{ntj}, q_{ns})^{y_{ntj}}$$

The model was estimated via the Expectation- Maximization (EM) algorithm.
Adoptions for New Transportation Service

Figure 3: Generalized Technology Adoption Model

Data set

- **User data**: the dataset consists of all individuals that have signed up for the carsharing service for a time period of 2.5 years after being launched.

**Figure 4: Cumulative Number of Adopters of Carsharing Service**

Data set

• **Facility data**: the growth in the number of pods/stations and on-street pick-up/drop-off locations for the 2.5-year time period.

![Bar chart showing growth in number of pods/stations and on-street parking](image)

**Figure 5: Growth in Number of Pods/Stations and On-Street Parking over Time**

Results

Figure 6: Cumulative Adoptions for New Transportation Service