Econométrie 3

Discrete Choice and Choice Theory

Iragaël Joly

Grenoble-Inp, Génie Industriel - Gael

iragael.joly@grenoble-inp.fr

Nov. 2019

Table of contents

Discret Choice Model

Underlying behavioural decision

Economic decisions and DCM

Economists are interested in market demand but,

relationship between market and individual demand is complicated :

- Individual consumption decisions are based on individual needs and environmental factors,
- Individual decisions are complex
  - Economic rationality and associated assumption of utility maximisation permit
  - Unobserved attributes of individuals varying over a population
  - They obscure the implications of the individual behaviour model

From the individual behavior to the market demand

Specification, estimation and inference for discrete choice models (1)

The classical theory of consumer behavior

- A representative consumer with preferences :
- utility function defined over the consumption of a vector of goods, $U(d)$
- assumed to maximize this utility
- subject to a budget constraint, $w'd \leq I$, where w : vector of prices ; I : income (or total expenditure).
- Assuming the necessary continuity and curvature conditions : complete set of demand equations, $d^* = d(w, I)$

Specification, estimation and inference for discrete choice models

The classical theory of consumer behavior

- From individual choice to observed market data
- Demand system is assumed to hold at the aggregate level
- Random elements are introduced to account for measurement error or optimization errors

Specification, estimation and inference for discrete choice models

The classical theory of consumer behavior

1. Classical theory has little to say about the discrete choices
   - The classical calculus : margins of consumption will comment 'how large'
   - but not 'what' good, or what brand of good
2. Introduction of random elements : less comfortable at the individual level than in market demands.
   - Consider carefully the appropriate sources and form of random variation in individual models of discrete choice.
   - RUM of discrete choice provides the most general platform for the analysis of discrete choice.
3. Extension of the classical theory of utility maximization to the choice among multiple discrete alternatives provides a straightforward framework for analyzing discrete choice in probabilistic, statistical, ultimately econometric, terms.

Multinomial Model

The econometric model

We need for a general model

- objects of choice and sets of alternatives available
- observed attributes of decision makers (and combination rule)
- model of individual choice and behavior
- distribution of behavior patterns in the population

The actual choice for an individual

- randomly drawn from the population
- described by some levels of common set of attributes (X)
- facing a set of available alternatives (J)
- can be defined as a draw from a multinomial distribution with selection probabilities

The probabilities of each alternative : $P(Y|X, J), \forall Y \in J$

Choice Theory (1)

Choice : outcome of a sequential decision-making process

- defining the choice problem
- generating alternatives
- evaluating alternatives
- making a choice,
- executing the choice

Choice Theory (2)

Building the theory

- who (or what) is the decision maker,
- what are the characteristics of the decision maker,
- what are the alternatives available for the choice,
- what are the attributes of the alternatives, and
- what is the decision rule that the decision maker uses to make a choice.

Choice Theory (3)

Decision maker

- observation unit
- a person
- a group of persons (internal interactions are ignored) : household, family, firm, government agency
- notation : n

Choice Theory (4)

Characteristics of the decision maker : S_n

- Disaggregate models
  - Individuals
  - face different choice situations
  - have different tastes
- Characteristics
  - income
  - sex
  - age
  - level of education
  - household/firm size
  - etc.

Choice Theory (5)

Alternatives : $\mathcal{C}_n$

- Choice set
  - Non empty finite and countable set of alternatives
  - Universal : $\mathcal{C}$
  - Individual specific : $\mathcal{C}_n \subseteq \mathcal{C}$
  - Availability, awareness
- Example : Choice of a transportation model
  - $\mathcal{C} =$ car, bus, metro, walking
  - If the decision maker has no driver license, and the trip is 20km long
  - $\mathcal{C}_n =$ bus,metro

Choice Theory (6)

Continuous and Discrete Choice Set

Microeconomic demand analysis

Commodity bundle

- q_1 : quantity of good 1
- q_2 : quantity of good 2
- q_3 : quantity of good 3
- Unit price : p_i, $i = 1, 2, 3$
- Budget : I

Discrete choice analysis

List of alternatives

- Brand A
- Brand B
- Brand C

Choice Theory (7)

Alternative attributes : Z_{in}

- Characterize each alternative i for each individual n :
  ex : price, travel time, comfort, etc.
- Nature of the variables :
  - Discrete and continuous
  - Generic and specific
  - Measured or perceived

Decision rule

Homo economicus

Rational and self-interested and optimizing her outcome

Utility

$U_n : \mathcal{C}_n \to \mathbb{R} : a \to U_n(a)$

- captures the attractiveness of an alternative
- measure to be optimized by decision maker

Behavioral assumption

- the decision maker associates a utility with each alternative
- the decision maker is a perfect optimizer
- the alternative with the highest utility is chosen

Microeconomic consumption theory (1)

Utility, preferences and rationality

- Rationality : Preferences are complete, transitive and continuous
- Utility is consistent with rationality
- Utility is unique up to an order-preserving transformation

Microeconomic consumption theory (2)

Consumer problem

- Maximize Utility
- Subject to budget constraints
- Defining the demand functions
- Remember the Cobb-Douglas case :
  Maximize $U(x_1, x_2) = Ax_1^{\alpha}x_2^{\beta}$
  subject to $p_1 x_1 + p_2 x_2 = I$
  $x_1^* = \frac{I}{p_1}\frac{\alpha}{\alpha+\beta}$ $\quad x_2^* = \frac{I}{p_2}\frac{\beta}{\alpha+\beta}$
- increase with I, decrease with price, only dependant on its own price

Microeconomic consumption theory (3)

Indirect Utility

- Utility at the optimal demand $U(I, p_1, p_2) = U(x_1^*, x_2^*)$
- Maximum utility that is achievable for a given set of prices and income

In discrete choice

- only the indirect utility is used
- simply referred to as "utility"

Microeconomics outputs (1)

Roy's identity

- Derive the demand function from the indirect utility

$$x_i = -\frac{\partial U(I, p_i, ...)/\partial p_i}{\partial U(I, p_i)/\partial I}$$

Microeconomics outputs (2)

Elasticities

- Direct price elasticity Percent change in demand resulting form a 1 % change in price

$$E_{p_i}^{x_i} = \frac{\text{\% of change in} x_i}{\text{\% of change in} p_i} = \frac{\Delta x_i}{\Delta p_i} = \frac{\Delta x_i}{\Delta p_i} / \frac{x_i}{p_i}$$

- Asymptotically

$$E_{p_i}^{x_i} = \frac{p_i}{x_i(I, p_i, ...)} \frac{\partial x_i(I, p_i, ...)}{\partial p_i}$$

- Cross price elasticity

$$E_{p_j}^{x_i} = \frac{p_j}{x_i(I, p_i, ...)} \frac{\partial x_i(I, p_i, ...)}{\partial p_j}$$

Microeconomics outputs (1)

Consumer surplus

- Difference between what a consumer is willing to pay for a good and what she actually pays for that good.
- Area under the demand curve and above the market price

Discrete Goods (1)

Model for individual n

Simplifications

- We cannot estimate a set of parameters for each individual n
- Therefore, population level parameters are interacted with characteristics of the decision-maker S_n
- Prices of the continuous goods are neglected p_n
- Income is considered as another characteristic and merged into S_n
- alternative cost c_i is considered as another attribute and merged into Z_n

$$max_i U_{in} = U(Z_{in}, S_n, \beta)$$

Simple Example (1)

Attributes

	Attributes	
Alternatives	Travel time (t)	Travel cost (c)
Car (1)	t_1	c_1
Bus (2)	t_2	c_2

Simple Example (2)

Utility

$$U = U(y_1, y_2)$$

where we impose the restrictions that, for $i = 1, 2$

$$y_i = \begin{cases} 1 & \text{if travel alternative i is chosen,} \\ 0 & \text{otherwise;} \end{cases}$$

and that only one alternative is chosen : $y_1 + y_2 = 1$.

Simple example : mode choice (1)

Utility functions

$$
\begin{aligned}
U_1 &= -\beta_t t_1 - \beta_c c_1, \\
U_2 &= -\beta_t t_2 - \beta_c c_2
\end{aligned}
$$

where $\beta_t > 0$ and $\beta_c > 0$ are parameters.

Utility functions

$$
U_1 = -(\beta t/\beta c) t_1 - c_1 = -\beta t_1 - c_1
$$
$$
U_2 = -(\beta t/\beta c) t_2 - c_2 = -\beta t_2 - c_2
$$

where $\beta > 0$ is a parameter.

Simple example : mode choice (2)

Choice

- Alternative 1 is chosen if $U_1 \geq U_2$.
- Ties are ignored.

Choice

Alternative 1 is chosen if

$$-\beta t_1 - c_1 \geq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \geq c_1 - c_2$$

Alternative 2 is chosen if

$$-\beta t_1 - c_1 \leq -\beta t_2 - c_2$$

or

$$-\beta(t_1 - t_2) \leq c_1 - c_2$$

Dominated alternative

If $c_2 > c_1$ and $t_2 > t_1$, $U_1 > U_2$ for any $\beta > 0$

If $c_1 > c_2$ and $t_1 > t_2$, $U_2 > U_1$ for any $\beta > 0$

Simple example : mode choice (3)

Trade-off

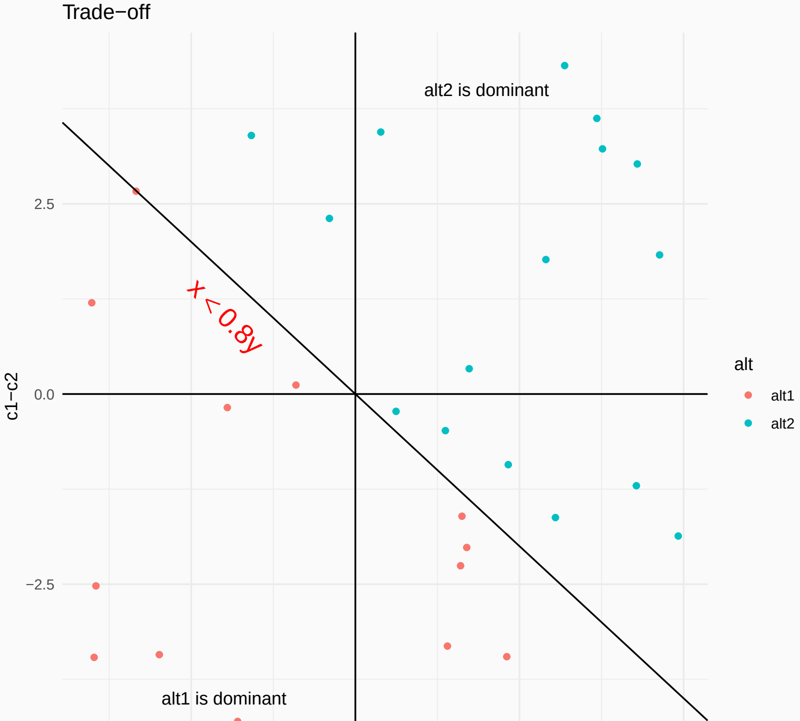
Assume $c_2 > c_1$ and $t_1 > t_2$.

Is the traveler willing to pay the extra cost $c_2 - c_1$ to save the extra time $t_1 - t_2$?

Alternative 2 is chosen if $-\beta(t_1 - t_2) \leq (c_1 - c_2)$

$$\beta \geq \frac{(c_1 - c_2)}{(t_1 - t_2)}$$

β is called the willingness to pay or value of time


Trade−off
alt2 is dominant
x < 0.8y
alt1 is dominant
2.5
0.0
−2.5
c1−c2
alt
alt1
alt2

Probabilistic choice theory

Behavioral validity of the utility maximization ? (1)

Assumptions

Decision-makers

- are able to process information
- have perfect discrimination power
- have transitive preferences
- are perfect maximizer
- are always consistent

Relax the assumptions

- Use a probabilistic approach : what is the probability that alternative i is chosen ?

Introducing probability (1)

Random utility

- Decision-maker are rational maximizers
- Analysts have no access to the utility used by the decision-maker
- Utility becomes a random variable

Random utility model (1)

Probability model

$$P(i|C_n) = P(U_{in} \geq U_{jn}, \forall j \in C_n)$$

Random utility

$$U_{in} = V_{in} + \varepsilon_{in}$$

Random utility model

$$P(i|C_n) = P(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n)$$

or

$$P(i|C_n) = P(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n)$$

Derivation (1)

Joint distributions ε_n

Assume that $\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{J_n n})$ is a multivariate random variable with CDF $F_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n})$ and pdf $f_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n})$

Derive the model for the first alternative

$P_n(1|C_n) = P(V_{2n} + \varepsilon_{2n} \leq V_{1n} + \varepsilon_{1n}, \ldots, V_{Jn} + \varepsilon_{Jn} \leq V_{1n} + \varepsilon_{1n})$

or

$P_n(1|C_n) = Pr(\varepsilon_{2n} - \varepsilon_{1n} \leq V_{1n} - V_{2n}, \ldots, \varepsilon_{Jn} - \varepsilon_{1n} \leq V_{1n} - V_{Jn})$

Change of variables

$\xi_{1n} = \varepsilon_{1n},$

$\xi_{in} = \varepsilon_{in} - \varepsilon_{1n},\ i = 2, \ldots, J_n,$

Derivation (2)

Model in ξ

$P_n(1|C_n) = P(\xi_{2n} \leq V_{1n} - V_{2n}, \ldots, \xi_{J_n n} \leq V_{1n} - V_{J_n n})$

Note :

The determinant of the change of variable matrix is 1, so that ε and ξ have the same pdf

Derivation (1)

Derivation

$$
\begin{aligned}
& P_n(1|C_n) \\
&= P(\xi_{2n} \leq V_{1n} - V_{2n}, \ldots, \xi_{J_n n} \leq V_{1n} - V_{J_n n}) \\
&= F_{\xi_{1n},\xi_{2n},\ldots,\xi_{J_n}}(+\infty, V_{1n} - V_{2n}, \ldots, V_{1n} - V_{J_n n}) \\
&= \int_{\xi_1=-\infty}^{+\infty} \int_{\xi_2=-\infty}^{V_{1n}-V_{2n}} \ldots \int_{\xi_{J_n}=-\infty}^{V_{1n}-V_{J_n n}} f_{\xi_{1n},\xi_{2n},\ldots,\xi_{J_n n}}(\xi_1, \xi_2, \ldots, \xi_{J_n}) d\xi \\
&= \int_{\varepsilon_1=-\infty}^{+\infty} \int_{\varepsilon_2=-\infty}^{V_{1n}-V_{2n}+\varepsilon_1} \ldots \int_{\varepsilon_{J_n}=-\infty}^{V_{1n}-V_{J_n n}+\varepsilon_1} f_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n n}}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{J_n}) d\varepsilon \\
&= \int_{\varepsilon_1=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}-\varepsilon_{Jn}}}{\partial \varepsilon_1}(\varepsilon_1, V_{1n} - V_{J_n n} + \varepsilon_1, \ldots, V_{2n} - V_{J_n n} + \varepsilon_1) d\varepsilon_1
\end{aligned}
$$

The random utility model :

$$
P_n(i|C_n) = \int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}-\varepsilon_{Jn}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon
$$

Derivation (2)

- The general formulation is complex.
- We will derive specific models based on simple assumptions.
- We will then relax some of these assumptions to propose more advanced models.

Binary Choice Model

Binary choice example (1)

Binary choice

- A simple model is binary choice model
- Data are individual choices between two options
- Model : two utilities (one for each option)
- Our questions :
  - To identify and evaluate determinants of choice among alternative, individual and context variables
  - To estimate and predict economic outputs : marginal effects, elasticities, willingness to pay...

Binary choice example (2)

Binary choice

- Binary choice : two alternatives (Auto and Transit)
- Simplification : only travel time (alternative attribute), no individual variables
- No 'opt-out' issue
- one choice per individual / one observation (iid)

Binary choice example (3)

Specification of the utilities

$$U_C = \beta_1 T_C + \varepsilon_C$$

$$U_T = \beta_1 T_T + \varepsilon_T$$

where T_C is the travel time with car (min) and T_T the travel time with transit (min).

Choice model

Specification of Utility with slope and constant

$$\begin{aligned}
P(C|C,T) &= P(U_C \geq U_T) \\
&= P(\beta_1 T_C + \varepsilon_C \geq \beta_1 T_T + \varepsilon_T) \\
&= P(\beta_1 (T_C - T_T) \geq \varepsilon_T - \varepsilon_C) \\
&= P(\varepsilon \leq \beta_1 (T_C - T_T)) \\
&= P(\varepsilon'_T - \varepsilon'_C \leq \beta_1 (T_C - T_T) + (\beta_C - \beta_T)) = \\
&= P(\varepsilon' \leq \beta_1 (T_C - T_T) + \beta_0)
\end{aligned}$$

Where $\varepsilon = \varepsilon_T - \varepsilon_C$

where $\beta_0 = \beta_C - \beta_T$ and

$\varepsilon' = \varepsilon'_T - \varepsilon'_C$

Estimation of the outcomes probabilities

Parametric estimation of these probabilities of outcome needs :

- Data

#	Time auto	Time transit	Choice	#	Time auto	Time transit	Choice
1	52.9	4.4	T	11	99.1	8.4	T
2	4.1	28.5	T	12	18.5	84.0	C
3	4.1	86.9	C	13	82.0	38.0	C
4	56.2	31.6	T	14	8.6	1.6	T
5	51.8	20.2	T	15	22.5	74.1	C
6	0.2	91.2	C	16	51.4	83.8	C
7	27.6	79.7	C	17	81.0	19.2	T
8	89.9	2.2	T	18	51.0	85.0	C
9	41.5	24.5	T	19	62.2	90.1	C
10	95.0	43.5	T	20	95.1	22.2	T
				21	41.6	91.5	C

- Error specification

Error specification

Three assumptions about the random variables ε_T and ε_C

- What's their mean ?
- What's their variance ?
- What's their distribution ?

The binary probit model (1)

Choice model

$$P(C|\{C,T\}) = P(\beta_1(T_C - T_T) + \beta_0 \geq \varepsilon) = F_\varepsilon(\beta_1(T_C - T_T) + \beta_0)$$

The binary probit model

$$P(C|\{C,T\}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1(T_C - T_T) - \beta_0} exp\{-\frac{1}{2}t^2\}dt$$

Not a closed form expression

The Logistic distribution : Logistic (η, μ)

Probability density function (pdf)

$$f(x) = \frac{\mu exp\{-\mu(x-\eta)\}}{(1+exp\{-\mu(x-\eta)\})^2}$$

Cumulative distribution function (CDF)

$$P(X \geq \varepsilon) = F(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{1+exp\{-\mu(c-\eta)\}}$$

with $\mu > 0$

The binary logit model

Choice model

$$P(C|\{C,T\}) = P(\beta_1(T_C - T_T) + \beta_0 \geq \varepsilon) = F_\varepsilon(\beta_1(T_C - T_T) + \beta_0)$$

The binary logit model

$$\begin{aligned} P(C|\{C,T\}) &= \frac{1}{1 + exp\{-(\beta_1(T_C - T_T) + \beta_0)\}} \\ &= \frac{exp\{\beta_1 T_C + \beta_0\}}{exp\{\beta_1 T_C + \beta_0\} + exp\{\beta_1 T_T\}} \\ &= \frac{exp\{V_C\}}{exp\{V_C\} + exp\{V_T\}} \end{aligned}$$

Binary Logit Application

Binary choice example (1)

Data

#	Time auto	Time transit	Choice	#	Time auto	Time transit	Choice
1	52.9	4.4	T	11	99.1	8.4	T
2	4.1	28.5	T	12	18.5	84.0	C
3	4.1	86.9	C	13	82.0	38.0	C
4	56.2	31.6	T	14	8.6	1.6	T
5	51.8	20.2	T	15	22.5	74.1	C
6	0.2	91.2	C	16	51.4	83.8	C
7	27.6	79.7	C	17	81.0	19.2	T
8	89.9	2.2	T	18	51.0	85.0	C
9	41.5	24.5	T	19	62.2	90.1	C
10	95.0	43.5	T	20	95.1	22.2	T
				21	41.6	91.5	C

Binary choice example

First individual

- Parameters

  Let's assume that $\beta_0 = 0.5$ and $\beta_1 = -0.1$

- Variables

  Let's consider the first observation :

  - $T_{C1} = 52.9$
  - $T_{T1} = 4.4$

- Choice

  - Choice = *transit* : $y_{\text{auto},1} = 0$ and, $y_{\text{transit},1} = 1$

What's the probability given by the model that this individual indeed chooses transit ?

Binary choice example

First individual

- Utility functions $V_{C1} = \beta_1 T_{C1} = -5.29$
  $V_{T1} = \beta_1 T_{T1} + \beta_0 = 0.06$
- Choice model

$$P_1(\text{transit}) = \frac{exp\{V_{T1}\}}{exp\{V_{T1}\} + exp\{V_{C1}\}} = \frac{e^{0.06}}{e^{0.06} + e^{-5.29}} \simeq 1$$

- Comments
  - The model fits the observation very well
  - Consistent with the assumption that travel time is the only explanatory variable

Binary choice example

Second individual

- Parameters Let's assume that $\beta_0 = 0.5$ and $\beta_1 = -0.1$
- Variables
  - $T_{C2} = 4.1$
  - $T_{T2} = 28.5$
- Choice
  - Choice = *transit* : $y_{\text{auto},2} = 0$ and, $y_{\text{transit},2} = 1$

What's the probability given by the model that this individual indeed chooses transit ?

Binary choice example

Second individual

- Utility functions $V_{C2} = \beta_1 T_{C2} = -0.41$
  $V_{T2} = \beta_1 T_{T2} + \beta_0 = -2.35$
- Choice model

$$P_2(\text{transit}) = \frac{exp\{V_{T2}\}}{exp\{V_{T2}\} + exp\{V_{C2}\}} = \frac{e^{-2.35}}{e^{-2.35} + e^{-0.41}} \simeq 0.13$$

- Comments
  - The model fits the observation
  - But the assumption is that travel time is the only explanatory variable
  - Still, the probability is not small

Binary choice example

Likelihood

- Two observations The probability that the model reproduces both observations is

$$P_1(\text{transit})P_2(\text{transit}) = 0.13$$

- All observations The probability that the model reproduces all observations is

$$P_1(\text{transit})P_2(\text{transit})\ldots P_{21}(\text{auto}) = 4.6210^{-4}$$

- Likelihood of the sample

$$\mathcal{L} = \prod_n \left(P_n(\text{auto})^{y_{\text{auto,n}}} P_n(\text{transit})^{y_{\text{transit,n}}}\right.$$

where $y_{j,n}$ is 1 if individual n has chosen alternative j , 0 otherwise

Binary choice example

Likelihood

- Likelihood
  - Probability that the model fits all observations.
  - It is a function of the parameters.
- Examples

β_0	β_1	$\mathcal{L}$
0	0	4.5710^{-7}
0	-1	1.9710^{-30}
0	-0.1	4.110^{-4}
0.5	-0.1	4.6210^{-4}

Statistics (1)

Statistics on the parameters

```
dat <- read.table("Rscripts\\Data_Bin.txt",
                  header = T, sep=";")
head(dat)
dat$diff <- dat$transit-dat$auto
dat$Ch <- ifelse(dat$Choice=='T',1,0)
dat$Choice2 <- factor(dat$Choice)
table(dat$Ch)
#glm(data=dat, Choice2~diff, family=binomial("logit"))
glm(data=dat, Ch~diff, family=binomial("logit"))
# multinom(data=dat, Ch~diff )
print(logLik(M1))
# Restricted LogLik:
-length(dat$Ch)*log(2)
```

Statistics (1)

```
##   id auto transit Choice
## 1  1 52.9     4.4      T
## 2 11 99.1     8.4      T
## 3  2  4.1    28.5      T
## 4 12 18.5    84.0      C
## 5  3  4.1    86.9      C
## 6 13 82.0    38.0      C
##
##  0  1
## 10 11
```

Statistics (2)

```
##
## Call:  glm(formula = Ch ~ diff, family = binomial("logit"), data = dat
##
## Coefficients:
## (Intercept)         diff
##     0.23758     -0.05311
##
## Degrees of Freedom: 20 Total (i.e. Null);  19 Residual
## Null Deviance:     29.06
## Residual Deviance: 12.33  AIC: 16.33
```

```
## 'log Lik.' -6.166042 (df=2)
## [1] -14.55609
```

Summary statistics

- Likelihood $\mathcal{L}(\beta^*) = -6.166$
- $\mathcal{L}(0) = -14.556$

  $\mathcal{L}(0)$ sample log likelihood with all parameters are zero,
  that is a model always predicting $P(1|\{1,2\}) = P(2|\{1,2\}) = \frac{1}{2}$
  Purely a function of sample size : $\mathcal{L}(0) = \log((\frac{1}{2})^n) = -N\log(2)$

Summary statistics...

- Likelihood ratio test
  $log(\frac{\mathcal{L}(0)}{\mathcal{L}(\beta^*)}) = log(\mathcal{L}(0)) - log(\mathcal{L}(\beta^*)) = \mathcal{L}(0) - \mathcal{L}(\beta^*)$
  $LR = -2(\mathcal{L}(0) - \mathcal{L}(\beta^*)) = 16.780$
  - Under H_0, LR is asymptotically distributed as χ^2 with K degrees of freedom.
  - Similar to the F test in regression models
  - H_0 : the two models are equivalent

The scale effect (1)

```r
# PROBIT Model
probit <- glm(data=dat, Ch~diff,
              family=binomial("probit"))
probit
probit$coefficients * pi / sqrt(3)
```

The scale effect (2)

```
##
## Call:  glm(formula = Ch ~ diff, family = binomial("probit"), data = da
##
## Coefficients:
## (Intercept)         diff
##     0.06443     -0.03000
##
## Degrees of Freedom: 20 Total (i.e. Null);  19 Residual
## Null Deviance:     29.06
## Residual Deviance: 12.33  AIC: 16.33
```

```
## (Intercept)        diff
##  0.11687104 -0.05441208
```

Table of contents

Outputs in DCM

Economic Outputs in DCM (1)

Behavioral model

$$P(i|x_n;\theta)$$

- individual single prediction is of little use in practice.
- Need for indicators about aggregate demand.
  - Aggregate market shares.
  - Willingness to pay
  - Elasticities

Population Prediction (1)

Aggregation

- Identify the population T of interest (in general, already done during the phase of the model specification and estimation).
- Obtain individual characteristics x_n and C_n for each individual n in the population.
- The number of individuals choosing alternative i is

$$N_T(i) = \sum_{n=1}^{N_T} P_n(i|x_n;\theta)$$

.

Population Prediction (2)

Aggregation

- The share of the population choosing alternative i is

$$W(i) = \frac{1}{N_T} \sum_{n=1}^{N_T} P_n(i|x_n;\theta) = E[P(i|x_n;\theta)]$$

.

Aggregation (1)

Distribution

Population distribution among alternatives can be estimated

Population	Alternatives				Total
	1	2	$\cdots$	J	
1	$P(1\|x_1;\theta)$	$P(2\|x_1;\theta)$	$\cdots$	$P(J\|x_1;\theta)$	1
2	$P(1\|x_2;\theta)$	$P(2\|x_2;\theta)$	$\cdots$	$P(J\|x_2;\theta)$	1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N_T	$P(1\|x_{N_T};\theta)$	$P(2\|x_{N_T};\theta)$	$\cdots$	$P(J\|x_{N_T};\theta)$	1
Total	$N_T(1)$	$N_T(2)$	$\cdots$	$N_T(J)$	N_T

Population Prediction (1)

- Assume the distribution of x_n is available. $x_n = (x_n^C, x_n^D)$ is composed of discrete and continuous variables.
- x_n^C distributed with pdf $p^C(x)$.
- x_n^D distributed with pdf $p^D(x)$.

Market shares

$$W(i) = \sum_{x_D} \int_{x_C} P_n(i|x^C, x^D)\dot{p}^C(x^C)p^D(x^d)dx^C = E[P_n(i|x_n; \theta)]$$

Confidence intervals (1)

Inference in Model

$$P(i|x_n, p_i; \theta)$$

- In reality, we use θ, the maximum likelihood estimate of θ
- Property : the estimator is normally distributed $\mathcal{N}(\hat{\theta}, \hat{\Sigma})$

Confidence intervals (2)

Calculating the confidence interval by simulation

- Draw R times $\tilde{\theta}$ from $\mathcal{N}(\hat{\theta}, \hat{\Sigma})$
- For each $\tilde{\theta}$, calculate the requested quantity (e.g. market share, revenue, etc.) using $P(i|x_n, p_i; \tilde{\theta})$
- Calculate the 5% and the 95% quantiles of the generated quantities.
- They define the 90% confidence interval.

Willingness to pay (1)

Definition : Discrete variable

- Let c_{in} be the cost of alternative i for individual n.
- Let x_{in} be the value of another variable of the model
- Let $V_{in}(c_{in}, x_{in})$ be the value of the utility function.
- Consider a scenario where the variable under interest takes the value $x'_{in} = x_{in} + \delta^x_{in}$

Willingness to pay (2)

- We denote by δ^c_{in} the additional cost that would achieve the same utility, that is

$$V_{in}(c_{in} + \delta^c_{in}, x_{in} + \delta^x_{in}) = V_{in}(c_{in}, x_{in})$$

- The willingness to pay is the additional cost per unit of x, that is

$$\frac{\delta^c_{in}}{\delta^x_{in}}$$

Willingness to pay (3)

Continuous variable

- If x_{in} is continuous,
- if V_{in} is differentiable in x_{in} and c_{in},
- invoke Taylor's theorem :

$$
\begin{aligned}
V_{in}(c_{in}, x_{in}) &= V_{in}(c_{in} + \delta^c_{in}, x_{in} + \delta^x_{in}) \\
&\simeq V_{in}(c_{in}, x_{in}) + \delta^c_{in} \frac{V_{in}}{c_{in}}(c_{in}, x_{in}) + \delta^x_{in} \frac{V_{in}}{x_{in}}(c_{in}, x_{in})
\end{aligned}
$$

$$
\frac{\delta^c_{in}}{\delta^x_{in}} = -\frac{(\partial V_{in}/\partial x_{in})(c_{in}, x_{in})}{(\partial V_{in}/\partial c_{in})(c_{in}, x_{in})}
$$

Linear utility function

- If x_{in} and c_{in} appear linearly in the utility function, that is

$$V_{in}(c_{in}, x_{in}) = \beta_c c_{in} + \beta_x x_{in} + \ldots$$

- then the willingness to pay is

$$\frac{\delta_{in}^c}{\delta_{in}^x} = -\frac{(\partial V_{in}/\partial x_{in})(c_{in}, x_{in})}{(\partial V_{in}/\partial c_{in})(c_{in}, x_{in})} = \frac{\beta_{in}^x}{\beta_{in}^c}$$

Marginal Effects in Probability Models (1)

Partial Effect

- Partial effect are derivatives : $\frac{\partial P(y=j|X)}{\partial x}$
- depending on the model :
  - Logit : $\frac{\partial P(y=j|X)}{\partial x} = P(y=j|X) * (1 - P(y=j|X)) * \beta$
  - Probit : $\frac{\partial P(y=j|X)}{\partial x} = \text{normal density} * \beta$ P(y=j | X)

Disaggregate elasticities (1)

Direct and cross elasticities

- Varying variable x_{ink}
- Direct : $E^i_{x_{ink}} = \frac{\partial P_n(i)}{\partial x_{ink}} \frac{x_{ink}}{P_n(i)}$ or $\frac{\Delta P_n(i)}{\Delta x_{ink}} \frac{x_{ink}}{P_n(i)}$
- Cross : $E^i_{x_{jnk}} = \frac{\partial P_n(i)}{\partial x_{jnk}} \frac{x_{jnk}}{P_n(i)}$ or $\frac{\Delta P_n(i)}{\Delta x_{jnk}} \frac{x_{jnk}}{P_n(i)}$

Aggregate elasticity

It can be shown that :

$$E^{W(i)}_{x_{jk}} = \sum_{n=1}^{N_T} \frac{P_n(i)}{\sum_{n=1}^{N_T} P_n(i)} E^{P_n(i)}_{x_{jnk}}$$

Références

Train, K. (2009). *Discrete choice methods with simulation (2nd ed.).* UK : Cambridge University Press, Cambridge edition.