Lecture note for Discrete Choice Models #1

At
Advanced Transportation Planning

Tetsuro HYODO

History of Discrete Choice Models on transportation demand forecasting

In 1960's, popular modal split model was aggregate modal split curve such as:

$$P_1 = \frac{1}{1 + e^{\beta_1 time_1 + \beta_2 cost_1}}$$

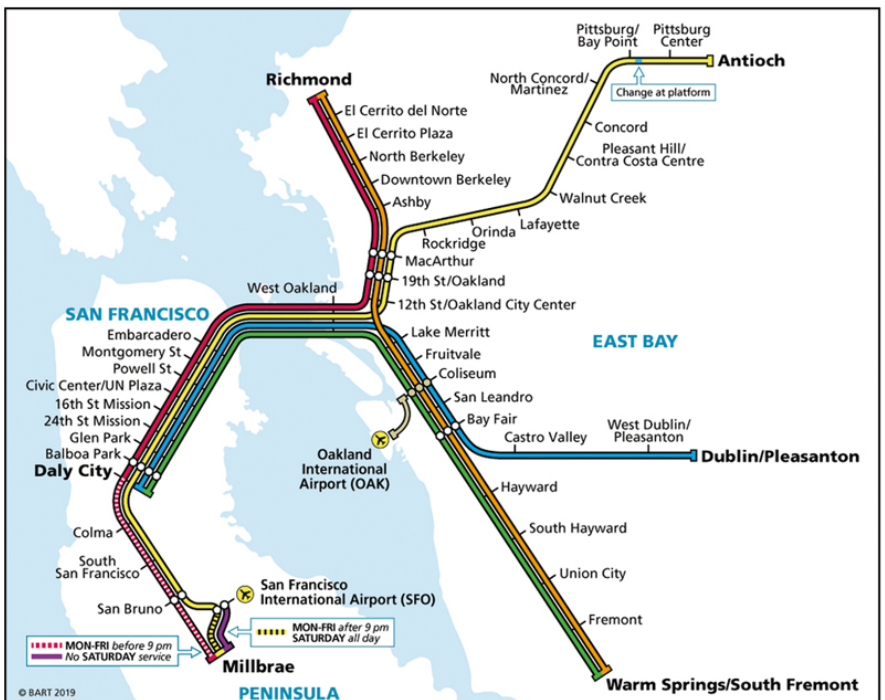
$$P_2 = (1 - P_1)\frac{1}{1 + e^{\beta_1 time_2 + \beta_2 cost_2}}$$

$$P_3 = (1 - P_1 - P_2)\frac{1}{1 + e^{\beta_1 time_3 + \beta_2 cost_3}}$$

There were no individual information, so this method was called "aggregate model"

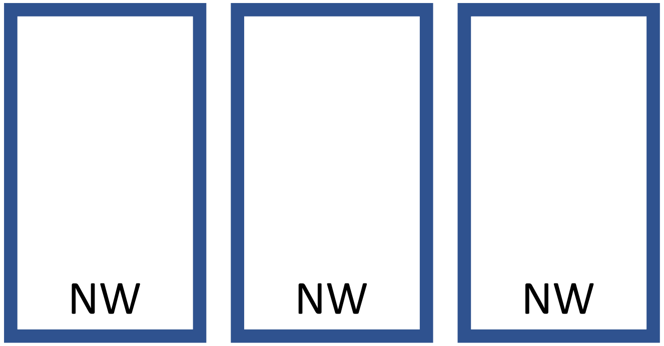
In early 1970's, the demand forecasting for BART (Bay Area Rapid Transit) applied discrete choice model, Logit Model.

Professor at UC Berkeley, department of ecomomics, Dr.Daniel McFadden developed this model.



The discrete choice model was successful to improve the accuracy of demand forecasting model. The error was under 5%, quite high predictability.

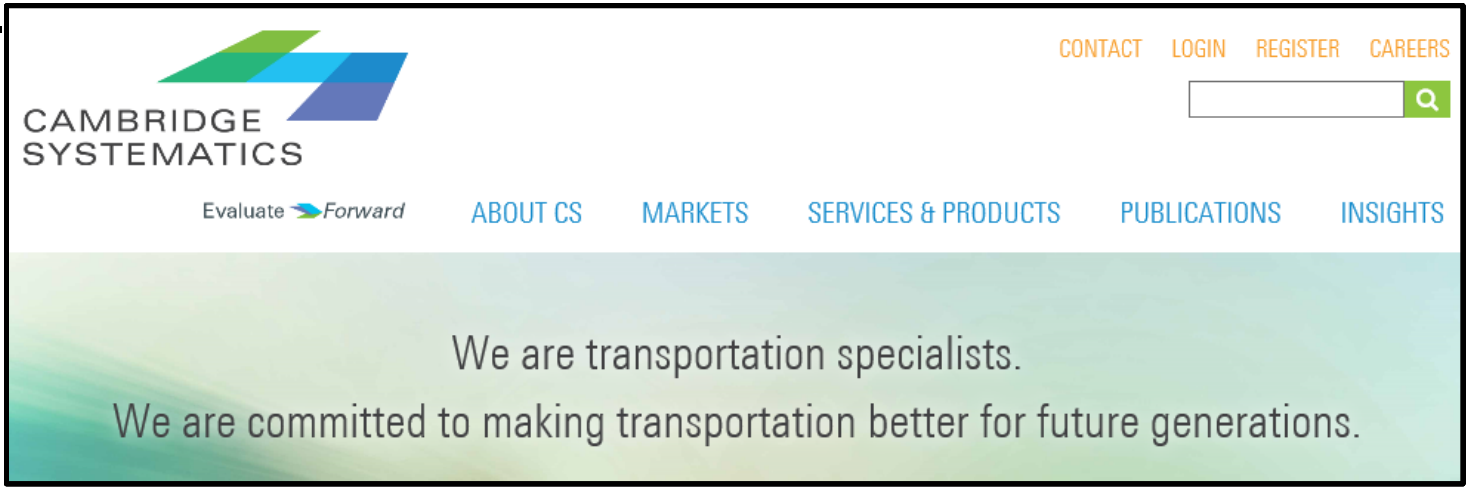
Because of the achievement of developing discrete choice models, Professor Daniel McFadden won the Nobel Prize in 2000.



This transportation demand forecasting "System" was sophisticated by professor Moshe Ben-Akiva, MIT with other professors.

They established "Cambridge Systematics", private company specialized for transportation demand forecasting.

Now Cambridge Systematics is one of the major consultant companies, and many graduates from MIT are working at this company.



Discrete choice model in Japan:

Professor Shigeru Morichi, my supervisor, stayed at MIT in 1979 when he was 36 years old. He studied not only theoretical mattes but also “Model System” developed by MIT team.
He returned to Japan and held a seminar on “Travel Demand Modelling by disaggregate model” by inviting three major MIT professors.

It gained popularity and became popular in Japanese academic people.

Disaggregate Logit model was applied demand forecasting for Tokyo Metropolitan Railway network. The first case was “Future master plan for 2000” which was conducted in 1985.

In Tokyo Area, every 15 years, the railway master plan was updated.
The 1985 plan involved “No.12 subway line”. This line is current “OHEDO line” which is the first ring subway line in Japan.

The opening day was December 12th in 2000 (Heisei 12th year) → 12/12/12.

Gumbel distribution (Generalized Extreme Value distribution Type-I)

The distribution of maximum values
→ earthquake, flood, typhoon…

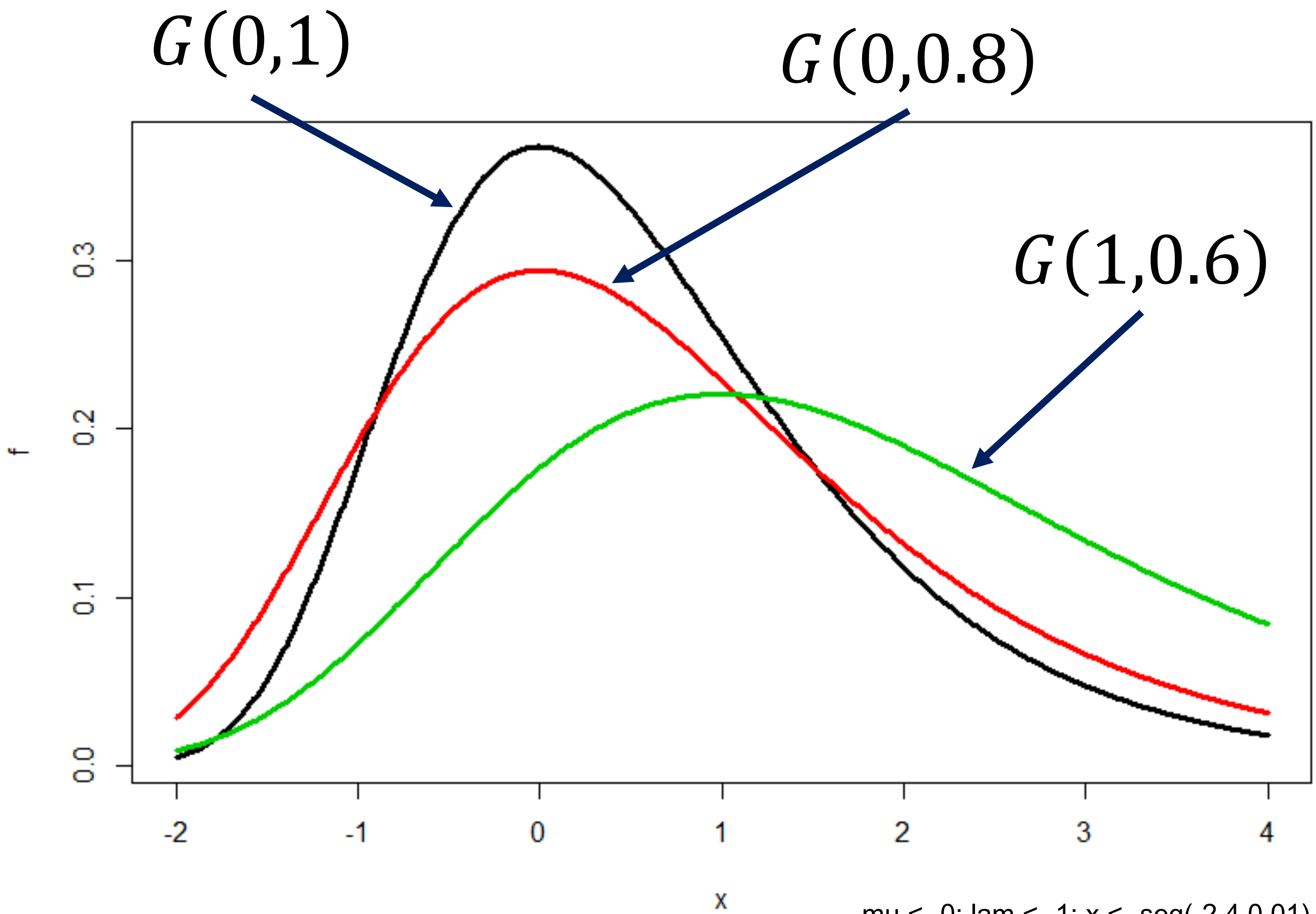
Gumbel distribution:
$G(\mu,\lambda)\ \mu: mode, \lambda: dispersion\ parameter$

$$F(x) = e^{-e^{-\lambda(x-\mu)}}$$

$$f(x) = \frac{\partial F(x)}{\partial x} = \lambda e^{-\lambda(x-\mu)} e^{-e^{-\lambda(x-\mu)}}$$

$$mean = \mu + \frac{\gamma}{\lambda}\ , \gamma \cong 0.577\ \ldots \leftarrow Euler's\ constant$$

$$variance = \frac{\pi^2}{6\lambda^2}$$




```
mu <- 0; lam <- 1; x <- seq(-2,4,0.01)
f <- lam*exp(-lam*(x-mu))*exp(-exp(-lam*(x-mu)))
  plot(x,f,type='l',lwd=3)
mu <- 0; lam <- 0.8; x <- seq(-2,4,0.01)
f <- lam*exp(-lam*(x-mu))*exp(-exp(-lam*(x-mu)))
  lines(x,f,lwd=3,col=2)
mu <- 1; lam <- 0.6; x <- seq(-2,4,0.01)
f <- lam*exp(-lam*(x-mu))*exp(-exp(-lam*(x-mu)))
  lines(x,f,lwd=3,col=3)
```

Binary Logit Model: Random Utility Model (RUM)

$$
\begin{aligned}
P_1 &= Pr[V_1 + \varepsilon_1 \geq V_2 + \varepsilon_2] \\
&= Pr[\varepsilon_1 = x] \times Pr[\varepsilon_2 \geq x + V_1 - V_2] \\
&= \int_{-\infty}^{\infty} f(x) \left(\int_{-\infty}^{x} f(y + V_1 - V_2) dy \right) dx \\
&= \int_{-\infty}^{\infty} f(x) F(x + V_1 - V_2) dx \\
&= \int_{-\infty}^{\infty} \lambda e^{-\lambda x} e^{-e^{-\lambda x}} e^{-e^{-\lambda(x+V_1-V_2)}} dx
\end{aligned}
$$

Utility:
1) Weight of different goods
2) Diminishing marginal utility
（限界効用逓減）

$Here,$

$$y = e^{-e^{-\lambda x}} e^{-e^{-\lambda(x+V_1-V_2)}} = e^{-e^{-\lambda x}\left(1+e^{-\lambda(V_1-V_2)}\right)}$$

$$\frac{\partial y}{\partial x} = \lambda e^{-\lambda x} e^{-e^{-\lambda x}\left(1+e^{-\lambda(V_1-V_2)}\right)} \left(1 + e^{-\lambda(V_1-V_2)}\right) = \lambda e^{-\lambda x} y \left(1 + e^{-\lambda(V_1-V_2)}\right)$$

$variable\ conversion\ between$ 'y' and 'x'

$$P_1 = \int_{-\infty}^{\infty} \lambda e^{-\lambda x}\, y dx = \int_{0}^{1} \frac{1}{1 + e^{-\lambda(V_1-V_2)}} dy$$

$$= \frac{e^{\lambda V_1}}{e^{\lambda V_1} + e^{\lambda V_2}}$$

```
> y <- expression( exp(- exp(-lam*x) ) * exp( -exp(-lam*(x+V1-V2))) )
> D(y, "x")
> exp(-exp(-lam*x)) * exp(-lam*x)*lam * exp(-exp(-lam*(x + V1 - V2))) + exp(-exp(-lam*x)) *
(exp(-exp(-lam*(x + V1 - V2))) * (exp(-lam * (x+V1-V2))*lam))
```

Multinomial Logit Model (MNL):

One of the important properties of Gumbel distribution:

$$E\left[\max_j G\left(V_j, \lambda\right)\right] = \frac{1}{\lambda} ln\left(\sum_j e^{\lambda V_j}\right)$$

In case of four alternatives:

$$E\left[\max_{j=2,3,4}[U_2, U_3, U_4]\right] = \frac{1}{\lambda} ln\left(e^{\lambda V_2} + e^{\lambda V_3} + e^{\lambda V_4}\right)$$

$$P_1 = \frac{e^{\lambda V_1}}{e^{\lambda V_1} + e^{\lambda E\left[\max_{j=2,3,4}[U_2,U_3,U_4]\right]}}$$

$$= \frac{e^{\lambda V_1}}{e^{\lambda V_1} + e^{ln\left(e^{\lambda V_2}+e^{\lambda V_3}+e^{\lambda V_4}\right)}}$$

$$= \frac{e^{\lambda V_1}}{e^{\lambda V_1} + e^{\lambda V_2} + e^{\lambda V_3} + e^{\lambda V_4}}$$

Parameter estimation by Maximum Likelihood (ML)

In case of 5 samples, 2 alternatives and 1 variable

$$P_{11} = \frac{e^{\beta x_{11}}}{e^{\beta x_{11}} + e^{\beta x_{12}}}, P_{12} = \frac{e^{\beta x_{12}}}{e^{\beta x_{11}} + e^{\beta x_{12}}}$$

The choice result is as follows,

sample	1	2	3	4	5
Alt.1	1	0	0	1	0
Alt.2	0	1	1	0	1

Likelihood is…

$$L^* = P_{11}^1 \cdot P_{12}^0 \cdot P_{21}^0 \cdot P_{22}^1 \cdot P_{31}^0 \cdot P_{32}^1 \cdot P_{41}^1 \cdot P_{42}^0 \cdot P_{51}^0 \cdot P_{52}^1$$

Log-Likelihood is…

$$L = \ln P_{11} + \ln P_{22} + \ln P_{32} + \ln P_{41} + \ln P_{52}$$

So, unknown parameter β can be estimated by

$$\max_{\beta} L$$

The value of variables x_{nj} are as follows;

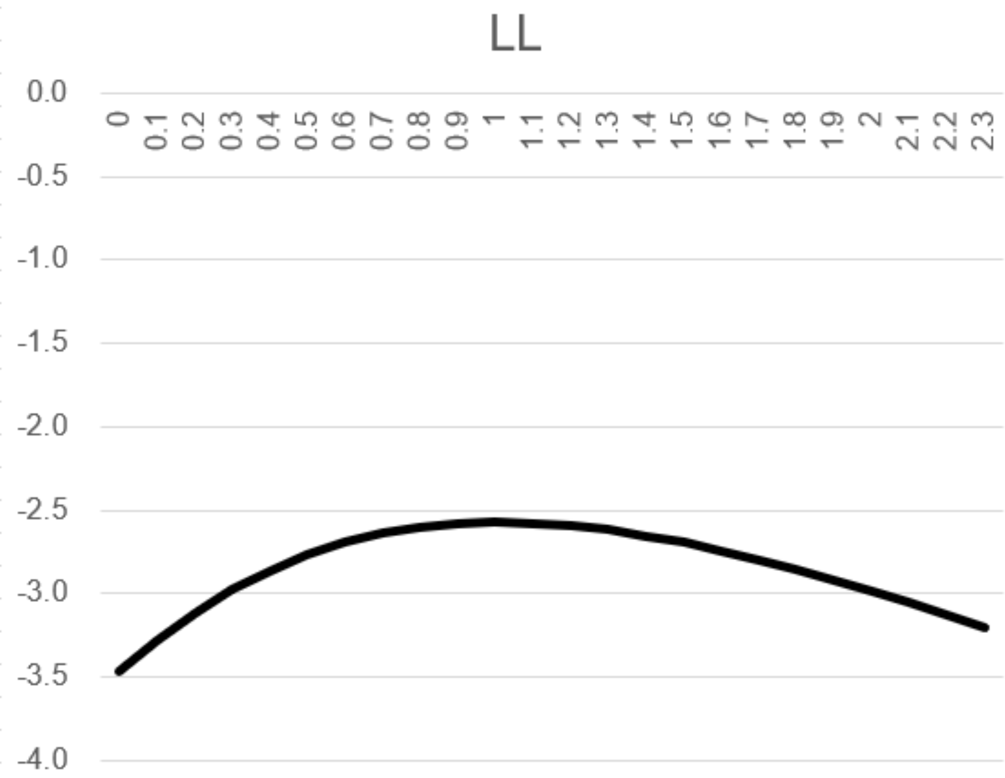
Microsoft Excel
ワークシート

sample	1	2	3	4	5
Alt.1	2	1	1	2	1
Alt.2	1	1	3	3	3

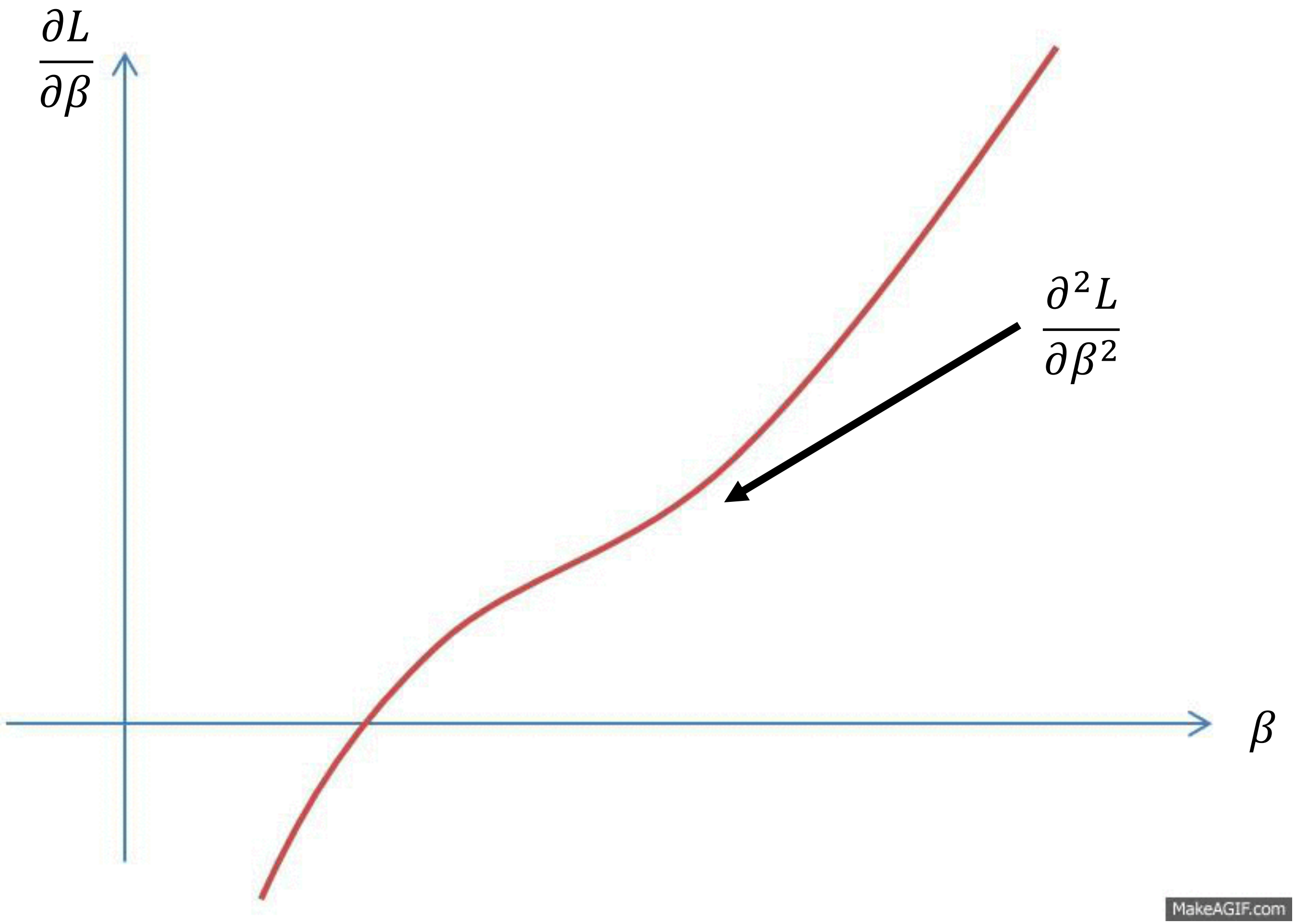
sample	1	2	3	4	5
Alt.1	2	1	1	2	1
Alt.2	1	1	3	3	3

L(0)=	-3.3651
rh0^2=	0.2352

beta	LL
0	-3.46574
0.1	-3.27822
0.2	-3.11546
0.3	-2.97683
0.4	-2.86138
0.5	-2.76782
0.6	-2.69469
0.7	-2.64035
0.8	-2.60315
0.9	-2.58141
1	-2.57353
1.1	-2.57798
1.2	-2.59338
1.3	-2.61845
1.4	-2.65205
1.5	-2.69315
1.6	-2.74086
1.7	-2.79438
1.8	-2.85302
1.9	-2.91617
2	-2.9833
2.1	-3.05395
2.2	-3.12772
2.3	-3.20424



Actually, Newton-Raphson method is applied.



$$L = \sum_{n=1}^{N} \sum_{j=1}^{J} \delta_{nj} \ln P_{nj} \quad \to Max$$

Index of goodness of fit

a) McFadden's rho-squared

$$\bar{\rho}^2 = 1 - \frac{LL}{L(0)}$$

LL: Final log-likelihood, *L(0)*: Initial log-likelihood

Previous example, *LL* = -2.574

L(0) = 2*ln(0.4) + 3*ln(0.6) = -3.365

$$\bar{\rho}^2 = 1 - \frac{LL}{L(0)} = 1 - \frac{-2.574}{-3.365} = 0.2352$$

The value of rho-squared is usually very low comparing R-squared of regression model. Over 0.3 is very good.

Index of goodness of fit

b) Hit ratio

Choice result of n-th sample is the maximum choice probability alternative. If the alternative equals observed alternative, "hit" otherwise "miss".
Previous case…

	1	2	3	4	5
P1	0. 731	0. 500	0. 119	0. 269	0. 119
P2	0. 269	0. 500	0. 881	0. 731	0. 881

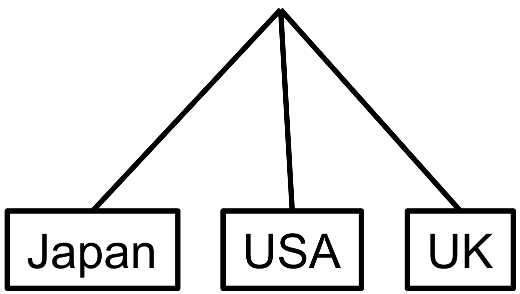
Hit table

		predict	
		1	2
observe	1	1	1
	2	0	3

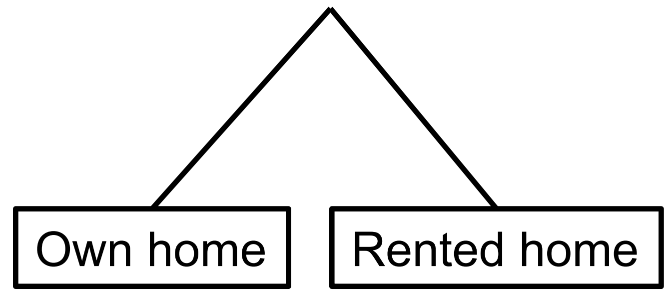
Hit ratio=4/5=80%

- Not only modal choice, destination choice, car own…

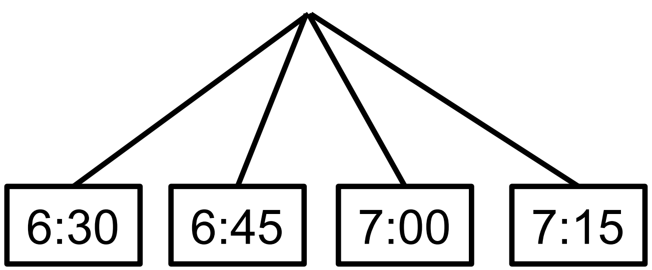
Destination choice model



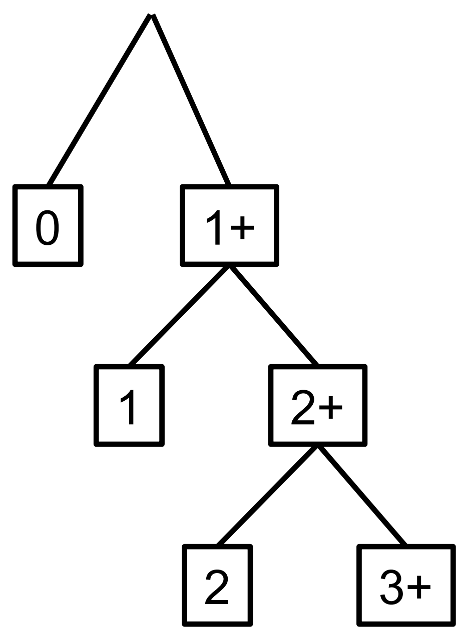
Housing type choice model



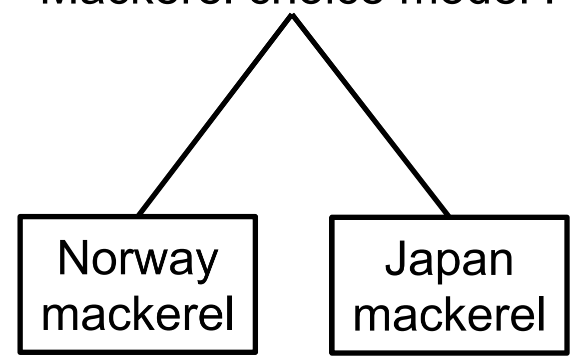
Departure time choice model



Trip generation choice model



Mackerel choice model !



Car ownership choice model