Exercise for the Estimation of Binary Logit Model

Imagine that there is railway station "x" and region "X". The access transport modes from X to x are only bicycle and bus. Let's make the binary logit model expressing the modal choice of people living in X.

We assume that the factors for the choice are only travel time *t* and cost *c*. The systematic term of utility fuction of mode *j*(1=bicycle, 2=bus) for individual *i* is expressed as follows,

$$V_{i1} = \beta_0 + \beta_1 t_{i1} + \beta_2 c_{i1} \quad \cdots\cdots (1)$$

$$V_{i2} = \beta_1 t_{i2} + \beta_2 c_{i2} \quad \cdots\cdots (2)$$

where β_k(k=0,1,2) are parameter to be estimated. The choice probability of mode *j* for individual *i* is expressed as follows,

$$P_{ij} = \frac{\exp(V_{ij})}{\exp(V_{i1})+\exp(V_{i2})} \quad \cdots\cdots (3)$$

We conduct the survey on the modal choice, travel time (minutes) and cost (yen) for 10 people. The result is in the following table.

Table 1. Survey result

Individual	Choice	Time (bicycle)	Cost (bicycle)	Time (bus)	Cost (bus)
1	1	10	100	8	160
2	2	20	110	12	160
3	2	14	130	16	160
4	1	15	100	14	200
5	2	30	150	20	240
6	2	22	140	18	200
7	1	8	100	14	160
8	1	14	120	23	200
9	1	20	150	15	200
10	2	18	130	12	160

1) Now, assume the variable δ_{ij} which is 1 if *i* chooses *j*, otherwise 0. The joint choice probability (likelihood) for 10 samples *L* is expressed by equation (4). Fill variables in (A) to (D).

$$L = \prod_{i=1}^{10} \boxed{\text{(A)}}^{\boxed{\text{(B)}}} \boxed{\text{(C)}}^{\boxed{\text{(D)}}} \quad \cdots\cdots (4)$$

2) Estimators $\hat{\beta}_k$ are the value of β_k that maximizes the log-likelihood function *l* (=ln*L*). That is, our task is to solve the following simultaneous equations.

$$\nabla l = \begin{pmatrix} \frac{\partial l}{\partial \beta_0} \\ \frac{\partial l}{\partial \beta_1} \\ \frac{\partial l}{\partial \beta_2} \end{pmatrix} = 0 \quad \cdots\cdots (5)$$

Prove that ∇l is expressed by equation (6).

$$\nabla l = \begin{pmatrix} \sum_{i=1}^{10}(\delta_{i1} - P_{i1}) \\ \sum_{i=1}^{10}(\delta_{i1} - P_{i1})(t_{i1} - t_{i2}) \\ \sum_{i=1}^{10}(\delta_{i1} - P_{i1})(c_{i1} - c_{i2}) \end{pmatrix} \cdots\cdots (6)$$

3) Equation (5) is non-linear equation and can be solved by optimization methods like Newton-RaphsonMethod. Derive the following Hessian matrix ∇^2 by using P_{ij}, t_{ik} ,and c_{ik}.

$$\nabla^2 l = \begin{pmatrix} \frac{\partial^2 l}{\partial \beta_0^2} & \frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 l}{\partial \beta_0 \partial \beta_2} \\ \frac{\partial^2 l}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 l}{\partial \beta_1^2} & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} \\ \frac{\partial^2 l}{\partial \beta_0 \partial \beta_2} & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 l}{\partial \beta_2^2} \end{pmatrix} \cdots\cdots (7)$$

4) The general operation of Newton-Raphson Method is summirized in the following. We start from the initial β_k as $\boldsymbol{\beta}^{(0)} = (0,0,0)^t$. Considering that ∇l and $\nabla^2 l$ are the function of $\boldsymbol{\beta}$, we obtain $\boldsymbol{\beta}^{(1)}$ in the first step.

$$\boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(0)} - \left\{\nabla^2 l\left(\boldsymbol{\beta}^{(0)}\right)\right\}^{-1} \nabla l\left(\boldsymbol{\beta}^{(0)}\right) \cdots\cdots (8)$$

We obtain $\boldsymbol{\beta}^{(m+1)}$ by the following equation in the itteration process.

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} - \left\{\nabla^2 l\left(\boldsymbol{\beta}^{(m)}\right)\right\}^{-1} \nabla l\left(\boldsymbol{\beta}^{(m)}\right) \cdots\cdots (9)$$

The itteration may be stopped when $\boldsymbol{\beta}^{(m+1)} - \boldsymbol{\beta}^{(m)}$ become smaller than some criteria, and regard $\boldsymbol{\beta}^{(m)}$ as estimators $\hat{\beta}_k$. Find $\hat{\beta}_k$ using the data in Table 1.

5) Explain your estimation results from several viewpoints.