1 Exercises in discrete choice (logit) theory

Jonas Eliasson/Daniel Jonsson/Tom Petersen 2003-01-12

1. We have the following discrete choice model (logit):

$$P_i = \frac{e^{\mu V_i}}{\sum_j e^{\mu V_j}} \tag{1}$$

   where V_i is interpreted as the measurable utility of alternative i and μ as inversely proportional to how well this utility measure can predict the choice of a randomly chosen individual.

   (a) Formulate a stochastic utility maximisation problem which leads to these choice probabilities.

   (b) What happens with the choice probabilities when $\mu \to 0$ and $\mu \to \infty$, respectively?

   (c) Suppose all observable utilities are changed from V_i to $V_i + \triangle V_i$. Express the new choice probabilities in the old ones (P_i) and the changes in the observable utilities $\triangle V_i$.

2. Show that a utility function in a discrete choice model never can be determined more than up to an additive constant.

3. The concept of elasticity is often used in economic theory. The own-elasticity of the "demand" P_i of an alternative i with respect to a variable c_i, in the utility function corresponding to this alternative, is defined as

$$\frac{\partial P_i / P_i}{\partial c_i / c_i} = \frac{\partial P_i}{\partial c_i} \frac{c_i}{P_i} \tag{2}$$

   and the cross-elasticities with respect to a variable c_j belonging to *another* alternative ($j \neq i$) is

$$\frac{\partial P_i / P_i}{\partial c_j / c_j} = \frac{\partial P_i}{\partial c_j} \frac{c_j}{P_i} \tag{3}$$

   (a) Interpret the definitions of the elasticities in words.

   (b) Derive the elasticities from the following discrete choice model, where the probability to choose alternative i is given by

$$P_i = \frac{e^{V_i}}{\sum_j e^{V_j}} \tag{4}$$

   with utility functions $V_i = \alpha_i + \beta c_i$.

4. The concept of value-of-time (VOT) plays a central role in many analyses of social welfare. It can be defined as follows: Let $P_i(c_i, t_i)$ be the probability that an individual chooses an alternative i (e.g. a mode of transport) given that the monetary cost is c_i and the time spent with this alternative is t_i. Let then Δt be a small increase in time spent. Then the value of Δt is the monetary compensation Δc that the individual needs for the probability to choose alternative i to remain constant. In short,

$$P_i(c_i, t_i) = P_i(c_i - \Delta c, t_i + \Delta t) \tag{5}$$

   and the VOT is defined as $\Delta c / \Delta t$.

   (a) Suppose P_i is described by a logit model where $V_i = \alpha_i - \beta c_i - \gamma t_i$ (α_i, β and γ are constant parameters). Show how the VOT can be expressed in these parameters.

   (b) Suppose instead that $V_i = f_i(c_i, t_i)$ is some (known) non-linear function. Show that it is still possible to define a VOT as above, but that this depends on the "starting point" values c_i and t_i.

5. It is sometimes desireable to insert income, Y, as an explanatory variable. Why are the following two suggestions not suitable?

$$V_i = \alpha_i - \beta(Y - c_i) - \gamma t_i \tag{6}$$

$$V_i = \alpha_i - \beta c_i - \gamma t_i + f(Y) \tag{7}$$

6. V_i is often specified as a linear function. A good argument for this is that the linear function can be seen as a Taylor approximation of an unknown, non-linear utility function. Suppose the individuals' "true" utility function is $V_i = f(Y - c_i, T - t_i)$, where f is an unknown function, Y is total income and T is total available time. Show by Taylor expansion around (Y, T) that V_i can be approximated by $-\beta c_i - \gamma t_i$. Which sign can we expect on the parameters β and γ?

7. A decision maker is choosing an alternative i from a finite set $\{1, 2, ..., n\}$. Each alternative has a known (deterministic) utility u_i. The decision maker can also himself decide with what probability he will get each alternative i. However, there is a cost attached to a certain choice of the probability vector $\mathbf{p} = (p_1, p_2, ..., p_n)$, namely $v(\mathbf{p})$. The problem of the decision maker is to maximise the expected utility, i.e. to maximise $\sum_{i=1}^n p_i u_i - v(\mathbf{p})$. Calculate the optimal probability vector if the cost function is $v(\mathbf{p}) = \sum_{i=1}^n p_i \ln p_i$.

8. A structured (syn. nested, hierarchic) logit model with two levels can be written

$$P_{im} = \frac{e^{\mu_1 (V_i + \widetilde{V}_i)}}{\sum\limits_j e^{\mu_1 (V_j + \widetilde{V}_j)}} \cdot \frac{e^{\mu_2 V_{im}}}{\sum\limits_m e^{\mu_2 V_{im}}} \tag{8}$$

where $\widetilde{V}_i = \frac{1}{\mu_2} \ln \sum_m e^{\mu_2 V_{im}}$.

(a) Explain/interpret the terms V_i, $\widetilde{V}_i$ and V_{im}.

(b) Explain the model intuitively in two ways: on one hand as a model of a decision process, and on the other as a description of how groups of error terms are connected.

(c) Show that the model can be written as a *simple* (syn. simultaneous, "not nested") model if $\mu_1 = \mu_2$.

9. We want to estimate a simple logit model, where the utilities of the alternatives are given by $V_i = \alpha_i + \beta c_i$, using the maximum likelihood (ML) method.At our service we have a dataset with, on one hand, the different costs for the alternatives for each individual n, c_i^n, and on the other hand which of the alternatives this individual has chosen, δ_i^n (1 if i was chosen, 0 otherwise).

   (a) State the log-likelihood function to be maximised.

   (b) Derive the ML equations with which the parameters α_i and β are determined.

   (c) Which aggregate properties in the data material will be reproduced by the model?