ELSEVIER

Regional Science and Urban Economics 28 (1998) 91–101

regional SCIENCE & urban ECONOMICS

The additive structure of utility in discrete choice models

M. Raouf Jaïbi[a], Thijs ten Raa[b,*]

[a]Departement de Mathématiques, Faculté des Sciences de Tunis, Campus Universitaire, 1060 Tunis, Tunisia

[b]Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Received 25 April 1994; received in revised form 7 August 1996; accepted 31 January 1997

Abstract

The logit model of discrete choice is generalized with respect to not only the shape of the error distribution, but also the additive cost structure. If utilities of alternatives are distributed according to different laws and many observations are accommodated, then a utility maximizer is shown to be governed by the logit model with an additive cost structure and an error distribution defined by the laws. The representation may break down in the case where the alternatives attracting positive choice probabilities have laws featuring either slow or fast decay, but a new result covers this situation. © 1998 Elsevier Science B.V.

Keywords: Additive structure; Logit model; Discrete choice

JEL classification: C25; R22; R10

1. Introduction

If alternatives carry utilities consisting of independent, Gumbel distributed terms (which clearly may assume different values across the alternatives) and additive costs,

$$u_i = \tilde{u}_i - c_i,\ \tilde{u}_i \sim \text{Gumbel},$$

then the probability that a utility maximizer picks alternative i decays exponentially with cost c_i. This is the so called logit model of discrete choice econometrics

*Corresponding author. Tel.: (31-13) 4662365; fax: (31-13) 4663280; e-mail: tenraa@kub.nl

PII S0166-0462(97)00014-8

(for references, see Jaïbi and ten Raa (1997). It is popular in spatial economics and other fields. The specification of the model, featuring the Gumbel law or distribution, and the additive cost structure, is sometimes considered restrictive. In a recent paper, we have generalized the error component of the logit model. More precisely, Jaïbi and ten Raa (1997) show that if

$$u_i = \tilde{u}_i - c_i, \tilde{u}_i \sim F,$$

where F is any distribution with a regular upper tail, then choice probabilities are logit, provided that many observations are accommodated. In the present paper, we attempt to rationalize the additive cost structure of the logit model. The latter is revealed by a simple rewrite of the last expression,

$$u_i \sim F_i(.) = F(. + c_i).$$

In this paper, we investigate a generalized discrete choice model with arbitrarily distributed utilities,

$$u_i \sim F_i,$$

and many observations. The importance of the paper is its demonstration that even in this completely general context, the logit model including its additive cost structure can be retrieved. Alternatives attracting positive choice probability will be shown to have similar tails, featuring either exponential, slow, or fast decay. In the first case, choice probabilities are shown to be logit with additive costs, c_i, defined by the laws, F_i. The choice probabilities can be viewed to be generated by the discrete choice model with additive costs c_i and error distribution $F = F_{i_0}$, where i_0 is any alternative attracting positive choice probability. The latter degree of freedom makes that costs are defined up to a translation. In the remaining cases (slow and fast decay), choice probabilities can be represented by limiting cases of the logit model or by new formulas presented in this paper. The generation of the choice probabilities by distribution F and additive cost c_i can be considered an additive random utility representation of the model with different laws, F_i. Note that we derive F and c_i. In other words, we show that the logit model emerges in the very general framework defined by arbitrary laws F_i. For additive random utility models, see, for example, Lindberg et al. (1990).

The generalized discrete choice model will be presented in Section 2. The only restriction will be that the distributions, F_i, have comparable tails. In Section 3, this technical condition will be defined and related to the regularity concept of Jaïbi and ten Raa (1997). Section 4 contains the determination of the limiting choice probabilities. The theorem generalizes the result of Jaïbi and ten Raa (1997). In Section 5, the additive cost structure is recovered whenever possible and new results covering the remaining cases are discussed. We relate to the applied literature in Section 6 and conclude with Section 7.

2. The model

Outcomes are 0, 1, . . . , m. Subset $I=\{1, \dots, m\}$ is the set of alternatives, such as locations. 0 is reserved for the case of no choice by absence of information. If x_i observations are made in $i\in I$, utility of alternative i is, assuming rational choice,

$$u_i = \max_{1\le j\le x_i} u_{ij}$$

where u_{ij} are random utility values at i. By convention, $u_i=-\infty$ if $x_i=0$. A referee pointed out to us that this case can be excluded by assuming $x_i\ge 1$. The random utilities at i are sampled independently from a cumulative distribution function (c.d.f.) F_i, and independently from the sample sizes (x_i).

Assumption 1: $\{u_{ij}|i\in I \text{ and } j\in\mathcal{N}\}$ is a family of independent random variables. For each $i\in I$, $\{u_{ij}|j\in\mathcal{N}\}$ are identically distributed with c.d.f. F_i.

Assumption 2: All c.d.f. F_i are continuous and have comparable tails in the sense of Definition 2 of Section 3.

For any numbers of observations, $x=(x_1, \dots, x_m)$, the probabilities of choosing alternatives i and of no choice, are defined by

$$P_i(x) = \mathbb{P}\{u_i > u_k, \text{ all } k\neq i\} \text{ and } P_0(x)=0$$

if $x\neq 0$, and by the exceptional case,

$$P_i(0)=0 \text{ and } P_0(0)=1.$$

To theorize about choice probabilities, we are interested in vectors of numbers of observations, x^n, $n\in\mathbb{N}$. The superscription randomizes the numbers of observations, driving them to infinity. The shares of alternatives will be forced to tend to proportions A_i.

Assumption 3: $\{x_i^n|i\in I$ and $n\in\mathbb{N}$ is a family of independent random integers with finite variances, independent of $\{u_{ij}|i\in I \text{ and } j\in\mathbb{N}\}$, and such that for every i,

$$\lim_{n\to\infty} \mathbb{E}(x_i^n)=\infty,$$

$$\lim_{n\to\infty} \frac{\mathbb{E}(x_i^n)}{\sum_{k=1}^{m} \mathbb{E}(x_k^n)} = A_i \in [0, 1],$$

$$\frac{x_i^n}{\mathbb{E}(x)_i^n)} \to 1 \text{ in probability as } n\to\infty.$$

Utility of alternative i is now

$$u_i^n = \max_{1 \le j \le x_i^n} u_{ij}$$

The probabilities of choosing alternatives i and of no choice, are now defined by

$$P_i^n = \mathbb{P}\{u_i^n > u_k^n, \text{ all } k \neq i\} \quad P_0^n = \mathbb{P}\{x^n = 0\}.$$

Limiting choice probabilities are defined by

Definition 1: $p_i = \lim_{n\to\infty} P_i^n$, $i = 0, 1, \ldots, m$.

3. Tails of distributions

Since the numbers of observations go to infinity in probability, only the upper tails of the utility distributions matter.

Definition 2: C.d.f.'s F and G have *comparable tails*, if

$$\alpha(F, G) = \lim_{u \uparrow \sup\{u|G(u)<1\}} \frac{1 - F(u)}{1 - G(u)}$$

is well defined. F and G are *tail similar*, if $\alpha(F, G)$ is positive and finite. F *dominates* G, if $\alpha(F, G)$ is infinite.

In our model, alternatives have distributions with comparable tails. In extreme value theory (see Jaïbi and ten Raa, 1997), a more narrow classification is relevant, namely tail equivalence (Resnick, 1971). It should be mentioned that tail similarity is implicit in the last reference. F and G are *tail equivalent* if and only if $\alpha(F, G) = 1$. We shall focus on the wider concept of similarity. If F and G are tail similar, then $\sup\{u|F(u)<1\} = \sup\{u|G(u)<1\}$ and $\alpha(G, F) = 1/\alpha(F, G)$ is also positive and finite. Therefore, similarity is a symmetric relation. Since it is also reflexive and transitive, tail similarity is an equivalence relation and any set $\mathcal{F}$ of distributions with comparable tails can be divided in equivalence classes of distributions with similar tails. The set of equivalence classes, $\hat{\mathcal{F}}$, is totally ordered by the relation $\preceq$ defined by

$$\hat{F} \preceq \hat{G} \text{ if and only if } \alpha(F, G) < \infty \text{ for some } F \in \hat{F} \text{ and } G \in \hat{G}$$

This statement is easily established. Reflexivity and transitivity are trivial. To establish anti-symmetry, suppose $\hat{F} \preceq \hat{G}$ and $\hat{G} \preceq \hat{F}$. Then, for $F \in \hat{F}$ and $G \in \hat{G}$, $\alpha(F, G) < \infty$ and $\alpha(G, F) = 1/\alpha(F, G) < \infty$. Hence, $0 < \alpha(F, G) < \infty$, that is, F and G are tail similar and $\hat{F} = \hat{G}$. Lastly, the order is total, because for any pair F and G with comparable tails, either $\alpha(F, G)$ or $\alpha(G, F)$ is finite.

Consider $\mathcal{F} = \{F_1, \ldots, F_m\}$, of Assumption 2. It contains a nonempty subset, $\mathcal{F}_M$, which is the maximal element of $\hat{\mathcal{F}}$ with respect to $\preceq$. The c.d.f.'s of $\mathcal{F}_M$ are tail similar and dominate all others. $\mathscr{F}_M$ is called the *dominant subset* of $\mathscr{F}$. In the next section, it will be shown to attract all choice probabilities. The remainder of this section is devoted to the connection with tail regularity in the sense of Jaïbi and ten Raa (1997).

C.d.f. F has *regular upper tail*, if

$$\varphi_F(c) = \lim_{u \uparrow \sup\{u|F(u)<1\}} \frac{1-F(u+c)}{1-F(u)}$$

is well defined for $c \geq 0$. Jaïbi and ten Raa (1997) consider c.d.f.'s with regular upper tails which are translations of each other, as reviewed in the introduction. In this case,

$$\varphi_{F_1}(c) = \ldots = \varphi_{F_m}(c) = \exp(-\mu c).$$

Here μ is zero, positive, or infinite, which they refer to by slow, exponential, or fast decay of the upper tail, respectively. In case of slow or exponential decay, all F_i can be seen to be tail similar. In case of fast decay, however, the positions of the c.d.f.'s, that is costs c_i, govern a pattern of domination,

$$F_j \preceq F_i \text{ if and only if } c_j \geq c_i.$$

Also c.d.f.'s with either slow or exponential decay and a common φ-function may feature domination, provided we leave the framework of translations or additive costs. For example, $F(u)=1-\exp(-u)$ $(u>0)$ and $G(u)=1-\exp(-u-\log u)$ $(u>1)$ have regular upper tails with $\varphi_F(c)=\varphi_G(c)=\exp(-c)$, but $\alpha(F, G)$ is infinite. In short, regularity and its nature bear little on similarity. Conversely, similarity does not bear on regularity. For example, $F(u)=1-\exp(-2u-\sin u)$ $(u>0)$ and $G(u)=1-\exp[-2u-\sin(u+\frac{1}{u})]$ $(u>1)$ are tail similar, even equivalent in the sense of Resnick (1971), but have no regular upper tails. However, if tail similar p.d.f.'s have regular upper tails, then their natures are the same, as the following lemma shows.

Lemma: Tail similar c.d.f.'s either have regular upper tails with common φ-function, or no regular upper tails at all.

Proof: See Appendix A.

Corollary: All members of the dominant subset have regular upper tails with common φ-function, or no regular upper tails at all.

4. Limiting choice probabilities

Recall $I=\{1,\ldots,m\}$ is the set of alternatives, $\mathcal{F}=\{F_i|i\in I\}$ is the set of utility c.d.f.'s associated with the alternatives, and $\mathcal{F}_M$ is the dominant subset. Define $I_M=\{i\in I|F_i\in\mathcal{F}_M\}$. Then there exists $i_0\in I_M$. For any pair of alternatives, put $\alpha_{ij}=\alpha(F_i,\ F_j)$. By Assumption 2, all α_{ij} are well defined. By definition of domination, α_{ii_0} is zero if $i\notin I_M$ and positive and finite if $i\in I_M$. In the latter case, $\alpha_{ki}\alpha_{ii_0}=\alpha_{ki_0}(k\in I_M)$.

Theorem: Under Assumptions 1–3, referring to Definition 1, and if $\Sigma_{k\in I_M} A_k>0$, then

$$p_i = A_i\alpha_{ii_0}/\sum_{k\in I_M} A_k\alpha_{ki_0}$$

for $i\in I_M$ and zero otherwise.

Proof: See Appendix A.

Example: $F_i(.)=F(.+c_i)$ with F having regular upper tail. If the decay of the upper tail is fast, then $I_M=\{i\in I|c_i=\min_{k\in I}c_k\}$ and $\alpha_{ii_0}=1$ (i, $i_0\in I_M$), hence $p_i=A_i/\sum_{k\in I_M} A_k$ for $i\in I_M$ and zero otherwise. Otherwise, $I_M=I$ and $\alpha_{ii_0}=\lim_{u\to\infty}[1-F(u+c_i)]/[1-F(u+c_{i_0})]=\varphi(c_i)/\varphi(c_{i_0})$, hence $p_i=A_i\varphi(c_i)/\Sigma_{k\in I} A_k\varphi(c_k)$. If the decay of the upper tail is slow, then $\varphi=1$ and $p_i=A_i$. If it is exponential, then $\varphi(c)=e^{-\mu c}$ and p_i is given by the logit model. This example consolidates the three cases of the result of Jaïbi and ten Raa (1997).

5. Recovery of the additive structure of utility

Consider the generalized model of discrete choice, with laws $(F_1,\ldots,F_m)$. All members of the dominant subset have regular upper tails with common φ-function. (Here we apply the corollary to Lemma 3 and rule out pathological c.d.f.'s.) By Jaïbi and ten Raa (1997), $\varphi(c)=\exp(-\mu c)$ with μ zero, positive, or infinite, and all members of the dominant subset feature slow, exponential, or fast decay of the upper tail, respectively. In this section, we treat the nondegenerate case of exponential decay (μ positive and finite), relegating slow or fast decay of the dominant subset members to Appendix A.

We claim that the generalized model is equivalent to the discrete choice model with additive cost structure, $F_{i_0}(.+c_i)$, where i_0 is any element of I_M (the index set of the dominant subset of the generalized model) and $c_i=-\mu^{-1}\log\alpha_{ii_0}$. If $i\in I_M$, then α_{ii_0} is positive and finite, and c_i is a finite real number. If $i\notin I_M$, then $\alpha_{ii_0}=0$ and c_i is defined to be infinity. The discrete choice model with additive cost structure can be subjected to the theorem of Jaïbi and ten Raa (1997), yielding logit limiting choice probabilities,

$$p_i = A_i e^{-\mu c_i} / \sum_{k \in I} A_k e^{-\mu c_k}, \; i \in I$$

or, using $c_i = \infty$ for $i \notin I_M$,

$$p_i = A_i e^{-\mu c_i} / \sum_{k \in I_M} A_k e^{-\mu c_k}, \; i \in I_M \text{ and zero otherwise}$$

or, substituting the definition of c_i,

$$p_i = A_i \alpha_{ii_0} / \sum_{k \in I_M} A_k \alpha_{ki_0}, \; i \in I_M \text{ and zero otherwise}$$

which are the limiting choice probabilities of the generalized model indeed. Our claim of equivalence is completed by three arguments. First, respective c.d.f.'s of the generalized model and the model with additive cost structure are tail equivalent on the set of alternatives attracting positive choice probabilities. Secondly, the representation is independent of i_0, up to a translation of the cost structure. Thirdly, the logit representation of the generalized model generalizes the discrete choice model with additive cost structure.

To present the first argument, let $i \in I_M$ and compare F_i and $F_{i_0}(.+c_i)$.

$$\alpha(F_i, F_{i_0}(.+c_i)) = \alpha(F_i, F_{i_0})\alpha(F_{i_0}, F_{i_0}(.+c_i)) = \alpha_{ii_0} \frac{1}{\varphi(c_i)} = e^{-\mu c_i} \frac{1}{e^{-\mu c_i}}$$
$$= 1.$$

To present the second argument, let $i_1 \in I_M$ be an alternative point of reference. Then the discrete model with additive cost structure becomes $F_{i_1}(.+d_i)$ with $d_i = -\mu^{-1} \log \alpha_{ii_1}$. If $i \notin I_M$, then $\alpha_{ii_1} = \alpha_{ii_0} = 0$ and $d_i = c_i = \infty$. If $i \in I_M$, then $\alpha_{ii_1} = \alpha_{ii_0}\alpha_{i_0 i_1}$. Taking logs and dividing by $-\mu$, $d_i = c_i - \mu^{-1} \log \alpha_{i_0 i_1}$. This shows that an alternative point of reference merely translates the cost structure by addition of a constant.

To present the third argument, take $F_i = F(.+c_i)$ with F featuring exponential decay. Then the costs associated with F_i are $-\mu^{-1} \log \alpha_{ii_0} = -\mu^{-1} \log[\varphi(c_i)/\varphi(c_{i_0})] = -\mu^{-1} \log(e^{-\mu c_i}/e^{-\mu c_{i_0}}) = c_i - c_{i_0}$, that is c_i up to an additive constant.

6. Implications for applied work

Witlox (1994) states that the most well known applications of discrete choice models are in the area of spatial analysis, namely (i) travel demand analysis whereby travellers pick modes of transportation, (ii) housing markets or residential location whereby households choose the location or community in which to rent or buy housing of one or more types, (iii) college choice whereby students pick schools, (iv) shopping behavior and retailing whereby shoppers pick stores, (v) recreational behavior whereby people choose between different recreational trips, (vi) labor participation whereby workers select job offers, (vii) choice of energy whereby heating systems are selected, (viii) product differentiation whereby consumers select brands or stores, and (ix) industrial location whereby firms review alternative locations. Our theory applies to the cases where there are many individual alternatives within discrete groups of alternatives. Of the above these spring to mind: housing markets and product differentiation, according to an anonymous referee. For housing markets, see Quigley (1985); Lerman (1977); Anas (1981), (1982); Li (1977); Ellickson (1981); Anas and Chu (1984); Gabriel and Rosenthal (1989). For product differentiation, see Anderson et al. (1992) and the references given there.

In the literature it is assumed that the random utilities of each dwelling or each brand or store are extreme value or, in other words, Gumbel distributed; logit choice probabilities are derived for each dwelling or brand/store and aggregated to obtain choice probabilities for residential locations or for store locations (or commodity types). The implication of our theory for this work is that the resulting spatial choice probabilities are robust with respect to the assumed random utility distributions. In fact, as long as one is interested in aggregated choice probabilities, say by location, the distributions of the error terms associated with the underlying individual dwellings or units within locations may be any. They need not even be translations of each other. The common assumption that utility is the sum of a systematic component (typically the negative of either transportation costs or the departure from an ideal point) and a random component is not necessary at the level of individual dwellings or units, but emerges automatically when choice probabilities across groups are evaluated. In short, a direct application of the standard logit model to the choice of locational choice between groups of alternatives is justified.

The criterion for grouping is arbitrary in the light of the independence of Assumption 1. In other words, it can be driven by the issues of interest, such as locations. However, the requirement that there are many individual units per group may force some aggregation. In Jaïbi and ten Raa (1997), aggregation would be by the cost component of utility. In this paper, aggregation would have to be applied by the systematic component of utility. While from a pure statistical point of view this is an issue of distributional fit, in practice the aggregation will be by a natural measure of similarity. Theorists have been puzzled by the requirement of discrete choice theory that alternatives must be ‘distinct.’ Debreu (1960) discusses recordings of the same concerto with a live performance and McFadden (1974) considers the choice between red and blue buses. Our theory can accommodate the situation where alternatives are similar. The examples are clear candidates for aggregation and the choice probabilities across groups can be modelled by the logit model, irrespective the functional form of the underlying distribution, provided only there are many recordings of concertos or many buses.

7. Conclusion

The logit model of discrete choice can be generalized not only with respect to the shape of the error distribution, but also the additive cost structure. If utilities of alternatives are distributed according to different laws and many observations are accommodated, then a utility maximizer is shown to be governed by the logit model with c.d.f. represented by anyone of the alternatives attracting positive choice probability and additive costs represented by the logs of the relative tail thicknesses. The representation breaks down only if utility distributions have upper tails which feature either slow or fast decay and are similar, but not equivalent. (Equivalent tails admit a trivial representation of the choice model.)

Acknowledgements

Suggestions by Tony Smith are gratefully acknowledged. The research was performed at Tilburg University with support provided by Fellowships of the Netherlands Organisation for the Advancement of Research and the Royal Academy of Arts and Sciences, respectively.

Appendix A

Proof of Lemma: Let F and G be tail similar. Then $\sup\{u|F(u)<1\}=\sup\{u|G(u)<1\}=b$, say. If b is finite, then F and G have regular upper tails with $\mu=\infty$. Otherwise, consider for any $c\geq 0$,

$$\frac{1-G(u+c)}{1-G(u)}=\frac{1-G(u+c)}{1-F(u+c)}\,\frac{1-F(u+c)}{1-F(u)}\,\frac{1-F(u)}{1-G(u)}$$

and take limits $u\to\infty$. Either they exist and

$$\varphi_G(c)=\alpha(G,F)\varphi_F(c)\alpha(F,G)=\varphi_F(c)$$

or

$$\lim_{u\to\infty}\frac{1-G(u+c)}{1-G(u)}\text{ and }\lim_{u\to\infty}\frac{1-F(u+c)}{1-F(u)}\text{ do not exist.}\quad \textit{Q.E.D.}$$

Proof of Theorem: The proof is a modification of the proof of the theorem of Jaïbi and ten Raa (1997), to which we refer. The strategy is to minorate

$\liminf_{n\to\infty} P_i^n$ by $P_i^\infty \geq 0$ and to show that P_i^∞ add up to unity over I_M. Here, under Assumptions 1–3, for $i \in I$, $P_i^n = \mathbb{E}[P_i(x^n)]$ with

$$P_i(x) = x_i \int_u F_i(u)^{x_i - 1} \prod_{\substack{k \neq i \\ x_k > 0}} F_k(u)^{x_k} \mathrm{d}F_i(u)$$

This is proved by trivial modification of the proofs of Lemmas 1 and 2 of Jaïbi and ten Raa (1997). We proceed with the proof of the theorem. Fact 1 now reads as follows. For $i \in I_\mathrm{M}$, $x \neq 0$ and $\varepsilon \in (0,1)$, there is a B_ε^i with $F_i(B_\varepsilon^i) < 1$ and

$$P_i(x) \geq \frac{C_\varepsilon^{-1} x_i}{\sum_{k=1}^{m} x_k \varphi_k} \{1 - \exp\{-C_\varepsilon[1 - F_i(B_\varepsilon^i)]x_i\}\}$$

where $C_\varepsilon = -\varepsilon^{-1} \log(1-\varepsilon)$ and $\varphi_k = \alpha_{ki} + \varepsilon$. (In the proof of this fact, define $\bar{\varphi} = \max_{k \in I} \alpha_{ki}$. Then $1 \leq \bar{\varphi} < \infty$, since $\alpha_{ii} = 1$ and $i \in I_\mathrm{M}$. The first inequalities are now based on the definition of tail comparability,

$$0 \leq 1 - F_k(u) \leq v_k = (\alpha_{ki} + \varepsilon)[1 - F_i(u)] \leq (\bar{\varphi} + \varepsilon)[1 - F_i(B_\varepsilon^i)] \leq \varepsilon$$

and the remainder is modified the same way.) In Fact 2, $F(B_\varepsilon)$ is replaced by $F_i(B_\varepsilon^i)$. If follows that

$$\liminf_{n\to\infty} P_i^n \geq A_i \sum_{k=1}^{m} A_k \alpha_{ki} = A_i \sum_{k \in I_M} A_k \alpha_{ki} = P_i^\infty$$

since $\alpha_{ki} = 0$ for $k \in I_\mathrm{M}$. Fix $i_0 \in I_\mathrm{M}$. Then, for $i \in \mathrm{M}$,

$$P_i^\infty = A_i \alpha_{ii_0} \sum_{k \in I_\mathrm{M}} A_k \alpha_{ki} \alpha_{ii_0} = A_i \alpha_{ii_0} \sum_{k \in I_\mathrm{M}} A_k \alpha_{ki_0}$$

add up to unity over I_M. Q.E.D.

The special cases of Section 5: It remains to consider the cases where the members of the dominant subset feature slow or fast decay. If their tails are equivalent, then the theorem reduces to

$$p_i = A_i \sum_{k \in I_\mathrm{M}} A_k \text{ for } i \in I_\mathrm{M} \text{ and zero otherwise}$$

In this case the generalized model can be seen to be equivalent to the discrete choice model with degenerate additive cost structure, $F_{i_0}(.+c_i)$, where $i_0 \in I_\mathrm{M}$ and $c_i = 0$. This is done in the same way as in Section 5. If the tails are not equivalent, but merely similar, then the generalized model admits no representation by a discrete choice model with additive cost structure, but the theorem must be applied directly to calculate limiting choice probabilities. They are proportional to the product of the relative sample size, A_i, and the relative tail thickness, α_{ii_0}. (An example is given by $F_1 = 1 - \frac{1/2}{1+u}$ and $F_2 = 1 - \frac{1}{1+u}$.)

References

Anas, A., 1981. The estimation of multinomial logit models of joint location and travel mode choice from aggregated data. Journal of Regional Science 21 (2), 223–242.

Anas, A., 1982. Residential location markets and urban transportation. In: Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models. Academic Press, New York.

Anas, A., Chu, C., 1984. Discrete choice models and the housing price and travel to work elasticities of location demand. Journal of Urban Economics 15 (2), 107–123.

Anderson, S.P., de Palma, A., Thisse, J.-F., 1992. Discrete Choice Theory of Product Differentiation. MIT Press, Cambridge, MA and London.

Debreu, G., 1960. Review of R.D. Luce, Individual Choice Behaviour: A Theoretical Analysis. The American Economic Review L (2), 186–188.

Ellickson, B., 1981. An alternative test of the hedonic theory of housing markets. Journal of Urban Economics 9 (1), 56–80.

Gabriel, S.A., Rosenthal, S.S., 1989. Household location and race: Estimates of a multinomial logit model. The Review of Economics and Statistics LXXI (2), 240–249.

Jaïbi, M.R., ten Raa, T., 1997. An asymptotic foundation for logit models. Regional Science and Urban Economics 28, 75–90.

Lerman, S.R., 1977. Location, housing, automobile ownership, and model to work: A joint choice model. Transportation Research Record 610, 6–11.

Li, M.M., 1977. A logit model of homeownership. Econometrica 45 (5), 1081–1097.

Lindberg, P.O., Eriksson, E.A., Mattson, L.-G., 1990. Generalized Extreme Value Choice Models and the Invariance Property, Working Paper, The Royal Institute of Technology, Stockholm.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In: Zarembk, P.C. (Ed.), Frontiers in Econometrics. Academic Press, New York and London, pp. 105–142.

Quigley, J.M., 1985. Consumer choice of dwelling, neighborhood and public services. Regional Science and Urban Economics 15 (1), 41–63.

Resnick, S.I., 1971. Tail equivalence and its applications. Journal of Applied Probability 8, 135–156.

Witlox, F., 1994. Discrete Choice Modelling in a Context of Spatial Choice Behavior: A Survey, SESO Report 311, Universiteit Antwerpen.