大韓交通學會誌 第二卷 第一號, pp. 89~101, 1984

□論 文□

Disaggregate, Two-Stage Travel Demand Model: Choice Set Reduction and Choice

複合的 交通選擇模型 : 選擇範圍 決定과 選擇

車 東 得*

(國土開發研究院 首席研究員)

目 次

要 約

오늘날 交通計劃分野에서 個別的, 行態論的 選擇模型의 使用이 급격히 一般化 되어가고 있는 추세를 보이고 있다. 이는 종래의 集合的 模型에 비하여, 構造的으로 經濟學에서 말하는 理性的 選擇行爲를 보다 잘 說明하고 있어, 選擇行爲의 因果關係를 나타낸다고 보여지기 때문이다.

그러나 이들 模型이 주어진 選擇範圍내에서의 選擇만을 다루고 있어, 選擇範圍를 決定하는데 任意性이 內在되어 있을 뿐만 아니라, 選擇對象이 많은 경우에는 곤란하다는 것이 問題點으로 지적되고 있다. 本 論文에서는, 選擇行爲에 관한 經濟學的, 心理學的 理論에 근거하여 比較的 實用的인 選擇範圍決定過程을 開發하여 旣存의 個別的, 行態論的 模型과 複合的으로 活用할수 있는 方案을 提示하였다.

* 本學會 正會員・交通工學博士

I. INTRODUCTION

In recent years, disaggregate, behavioral choice models have been extensively applied to transportation planning. The major motivation of these new approaches is that travel demand models should be based on causality to improve the forecasting ability by establishing the behavioral link between the relevant dicision variables and the individual's travel decision.

Unlike the traditional aggregate travel demand models, the shortcomings of which have been detailed in the literature, the new approaches are based on utility maximizing rational choice behavior, with the foundation on the postulate that a group of individuals with similar observed socioeconomic characteristics and transportation opportunities will have a distribution of travel behaviors independent of the location or date of observation. Although the empirical application of these new models indicate that a great deal of progress has been made in the practical procedures of forecasting travel demand, the new models themselves involve several undue behavioral assumptions.

The primary issue addressed in this paper is that current disaggregate, behavioral models assume a common choice set, and deal with how individuals make a final choice from these given alternatives (choice process), while there is little treatment of how these alternatives are generated (choice set process). It is reasonable to hypothesize that an individual in a sample population has his own choice set depending on his perceptions of alternatives. These perceptions, in turn, are constrained by the chooser's predispositions, the characteristics of alternatives and the environmental conditions of the individual. In this study it is hypothesized that an individual screens the original global set of alternatives based on his penceptions and forms a subset from which a final choice is made.

II. PROBLEMS AND ISSUES RELATED TO INDIVIDUAL CHOICE SETS

Assuming the alternative hypothesis proposed in this study is true, the application of existing models may entail several problems. The following describes some related problems and addresses relevant issues.

Problems of the Common Choice Set

In a situation where the alternative hypothesis governs, the common choice set concept will result in the biased estimation of model parameters and associated probability values.

Policy evaluations based on this biased model may be misleading and forecasting may be biased. Major problems are summarized as follows.

Bias in model parameters

This can be easily seen by looking at the log of the likelihood function. Suppose a two alternative case, i.e., alternatives 1 and 2. Two choice sets are assumed: choice set 1 contains alternative 1 only, and set 2, both alternatives. Let p_1 (x_t) and $p_2(x_t)$ denote the probabilities of set 1 and 2 to be selected respectively, where individual t evaluates the alternatives in terms of attribute vector, x_t. Since these two events constitute the entire sample space, p_1 (x_t) + P_2 (x_t) = 1. Then, an example of a logit choice will lead to

$$P(1,t) = [1+P_1\ (x_t)\cdot \exp(\beta' Z_t]/[1+\exp(\beta' Z_t)] \quad (1)$$

$$P(2,t) = [p_2\ (X_t)\cdot \exp(\beta' Z_t)]/[1+\exp(\beta' Z_t()\,)] \quad (2)$$

where P(i,t) = probability of individual t to choose alternative i

V_i = utility function for alternative i

$\beta' Z_t = V_2 - V_2$, the difference in the attribute of alternatives (Z_t)

Then, the log likelihood (LL) and derivatives are given by

$$LL = \sum_t \left\{ f_{1t} \cdot \log P(1,t) + (1\text{-}f_{1t} \cdot \log\,[1\text{-}P(1,t)] \right\} \quad (3)$$

$$\frac{\partial l\,L}{\partial \beta} = \sum_t \{ f_{1t} - \frac{1 + [p\,(xt)\cdot e^{\beta' Z_t})}{1 + e^{\beta' Z_t}} \}\, Z_t \quad (4)$$

where f_{1t} = 1, if t chooses alternative i, otherwise ϕ

Parameters are theoretically obtained as solutions of the simultaneous equations, $\partial LL/\partial B = O$. If, as is assumed in the conventional model, the common choice set represented by both alternatives governs the situation, $(P_1\ (x_t) \rightarrow \phi)$ then the differential equation converges to that of the conventional logit model. In all other general cases, however, the equation will be distorted by the bracket term and produce different parameters (B) from those of the conventional logit model.

Bias in probability values

Ben-Akiva (1977) gives a good theoretical discussion about this issue. Consider a simple example where the sample populations comprise two groups only: captives who must choose alternative i and non-captives who consider all alternatives. That is, two levels of the choice set are assumed; A_1 = {1} and A_n = {all alternatives}. Let u_i denote choice set probabilities for the respective cases. Then the probability of choosing alternative i,

P(i), is

$$P(i/An)\cdot [1+u_i\,(\frac{1}{P\,(i\ \ An)} - 1)] \quad (5)$$

Since the term in the bracket is greater than 1, $P(i) > P(i/An)$. This has an important implication for the testing of the proposed hypothesis, since the choice probability in the choice set identifying situation must be greater than that for the common choice set case, which means the choice set identification process helps individuals to better discriminate among alternatives.

Lack of adequate sensitivity

The choice set process is basically related to identifying how a particular change in the transportation system will affect whom, depending on where the individual is located and what the personal background is. Therefore, the conventional model without the choice set process may not be sufficiently sensitive to a system change since it simply deals with the average responses. The identification of the individual choice sets will open one way to understand who will gain or lose how much from the policy implemented, and what the future equilibrium will be.

Operational problems

The application of the common choice set model to a large choice set situation can make the computational requirements of a choice model prohibitively expensive. Perhaps more important as the number of choices increases, there is an increase in the likelihood of obtaining large errors in the parameter estimates and in prediction. The identification of the choice set process may alleviate this difficulty.

In summary, the concept of the common choice set may lead to a biased model and hence forecasting errors, a degradation of model transferability, less sensitivity to system change, and operational difficulties in a large choice set case.

Issues Related to the Choice Set Identification

With the above discussed problems in mind, it is now necessary to address issues associated with choice set identification

Degree of conformity to the proposed hypothesis

The degree of conformity to the proposed hypothesis will vary depending on the situation, that is, as the number of available alternatives increases and the similarity between alternatives decreases, the conformity is expected to increase. And the conformity can not be expressed by a single criterion. Instead, it should be understood as a composite of various related indices.

Characteristics of choice set identification

Normally available transportation data usually contains only two kinds of information: chosen alternatives and characteristics of individuals and alternatives. No information is available about individual choice sets. Under this situation, one of the best ways to represent individual choice sets is by inbedding the choice set process within the choice model.

III. THEORETICAL FRAMEWORK

It is important to note how the individual's predispostitions in combination with the environmental variables determine the behavior. Individuals place a subjective value or utility on the choice alternatives. Each alternative is evaluated on the basis of anticipated satisfaction, which are conditioned by the individual's socioeconomic characteristics. That is, the utility for each alternative is expressed by a set of attributes of the alternative and the individual's socio-economic background. Then the utilities are compared to each other and the decision is made that maximizes utility. This economic utility-maximizing choice behavior postulates that each individual is able to rank alternatives in order of preference. Several worthwhile points should be noted here.

It is difficult to believe that, in the perception process, the individual considers all the attributes of all the alternatives. Miller (1956) found that only a limited number of attributes are considered in any particular choice situation. Inclusion of irrelevant or seemingly not very relevant attributes in a linear utility function has no effect on the bias of the estimators, but it does affect the variances and reduces the estimation efficiency, see Horowitz (1979).

It is frequently observed that people, at the time of comparing two things, are more concerned about the difference between them than the absolute level of evaluation scores. We can assume that, with a few salient attributes, the individual evaluates alternatives along differences. This has two implication: (1) the attributes for which the individual perceives the alternatives most differently are more likely to be used in his evaluation, and (2) in this case, choice will not be influenced by the attributes that do not vary sufficiently over alternatives. However, in addition, individual specific attributes which are not a function of the alternatives, influence individual's values and so affect the way he perceives alternatives.

Then how is the difference considered by the individual in disoriminating among alternatives? A small difference in the perceived value of an attribute may not influence the choice decision. There is probably a certain 'threshold' level in the utility difference that makes the individual unable to discriminate among alternatives, even if real differences exist. Krishnan (1977), in fact, introduced to a conventional logit model a "minimum perceived difference (MPD)" to reflect the threshold by which the utility of the chosen alternative exceeds all other inferior alternatives, and reported an improved prediction.

The discussions here are relevant to both the choice set and the choice processes. However, they may be more applicable to the choice-set process which is governed by simplicity and sureness. The nature of the choice-set process as a preliminary screening of alternatives may require that the number of salient attributes be more severely limited to only those the individual considers to be very important and that the threshold value of discriminating utilities be realtively large.

Discriminating Utility

The assumption of individual-specific utility and associated discriminating threshold values for the choice-set attribute implies that discriminating utilities exist by which possible choices are eliminated to form a reasonable set of alternatives for a particular individual. This is because the threshold values are used to eliminate "clearly" undesirable alternatives by discriminating with desirable ones for further consideration. Therefore, it is possible to use a decision rule based on discriminating utilities between alternatives as a proxy of threshold values which are difficult to measure. The decision rule for choice set formation adopted in this study is that the individual eliminates clearly undersirable alternatives from a global choice set based on the utility distance between alternatives.

Now two relevant questions arise: What is the base from which utility distances are measured? This is compatible with underlying behavior for the choice-set process, but raises a question inherent in non-compensatory evaluation of attributes. That is, the result may be dependent on the selection of attributes since exclusion of important variables will make the screening process ineffective. This is an undesirable aspect of a modeling process since governing attributes for the choice set process may not be known a priori. However, transportation choice is a result of repeated experience and some of the important attributes are generally known, for example, travel time, out-of-pocket costs, and the like.

Specification of discriminating utilities

If rank orders of alternatives are similar for each attribute, it would suffice to pick up one such attribute to explain the difference in alternatives. Otherwise, if they conflict for different attributes, a selection of the two conflicting attributes is necessary.

Let Z denote the levels of each governing attribute and Y, individual or environmental characteristics. The proposition in this study is that each individual evaluates given Z's on a different scale constrained by Y's. Thus, the discriminating utilities, (the evaluation of Z's constrained by Y's), may be expressed as interaction terms.

In treating the compensation issue, the discriminating utility may be specified in two ways.

Strict non-compensation representation

In this manner, Z's are the levels of difference between alternatives for each attribute, and measured from the alternative with the highest score of that attribute. The discriminating utilities (DU) are specified independently for each attribute.

By structuring some causal link thorough P (i | z, θ*), these models can be used to predict the future demand or the demand in a new area, based on the assumption that the causal link will continue to hold even if the joint distribution of (i, z) pairs in the population is changed.

Choice Models

The random utility model usually breaks down the utility (U) of an alternative into two components: (1) representative or mean utility (v) of an alternative common to all individuals with homogenous socio-economic background, and (2) unobserved utility (n) unique to each individual.

These two components, in combination, constitute the utility of an alternative. Depending on the type of distribution applied to the random portion of the utility various specific models results. The conventional logit formulation, which assumes a Weibull distribution, is not ideal; it is not truly disaggregate in that the current practice of these models is based on some level of data aggregation, and not truly behavioral in a sense that it does not have a built-in process of identifying individual choice sets.

McFadden and Reid (1975), found that the first problem may not be critical; models which are calibrated using zonal aggregates for the independent variables, but use individually observed choices as the dependent variable yield unbiased disaggregate estimates as long as models are calibrated on sufficiently large amounts of data.

This research has been designed to alleviate the second problem in the hope of improving the prediction ability and model transferability by incorporating a choice set identification process.

IV. MODEL FORMULATION

Integrated Process

The individual-specific discriminating threshold value (h) of utilities which discrimantes one alternative from another is not observable. Instead, we assume that the discriminating utilities (DU_i) defined above are specific realizations of the threshold values when a particular alternatives is included in the choice set, and that they have some distribution over population. For computational simplicity, we assume an exponential distribution for h. Consequently, indivitual t will

include alternative i if $DU_{it} \leqslant h_t$
not include alternative i; otherwise

For the assumed exponential distribution,

$$F(h) = 1 - \exp(-\lambda h), h \geqslant \phi \quad (9)$$

$$DU = Z \cdot s(\beta' Y) \quad \cdots \quad (6)$$

where

DU = discriminating utility in choice set process

β = coefficients of Y, colunm vector

Y = individual and environmental characteristics, column vector

s = scaling function

Z = attribute difference from the maximum

Z is bounded $[0 + \infty]$, but $\beta'Y$ is unbounded, therefore, to bound DU in a similar manner to Z, a scaling function such as the exponential can be used:

$$DU = Z \cdot \exp(\beta' Y)$$

Compensation-compromised representation

The above representation may become more complicated as the number of choice-set governing attributes increases. Here we allow compensation in forming choice-set utilities (DU°). Based on the choice-set utility, the discriminating utility is constructed as a difference in DU° from the maximum.

Representation of a Causal Link

In choice models the dependent variable, choice probability, is not observable, but represented by the discrete event occurrence. The structural rationale of this type of model is summarized as follows: the outcome of transportation choice is the result of the interaction through some causal mechanism between the attributes of individuals and the environmental nature of choices open to them. The causality representation in the current control of behavior models can be summarized as follows.

Based on a sample population (T) facing a dommon, finite set of alternatives (C), it is assumed that the frequency distribution of choices (i) and the attribute vetor (Z) of the population can be characterized as:

$$f(i, z) = P(i | z, \theta^*) \cdot P(z) \quad \cdots \quad (7)$$

where

$f(i, z)$ = probability density represented by i and z

$P(i \mid z, \theta^*)$ = choice probability with known parameters

$P(z)$ = attribute distribution

In this case, the population share of alternative i, Q (i), is defined as

$$Q(i) = \int P(i \mid z, \theta^*) P(z) dz \quad \cdots \quad (8)$$

where

λ = reciprocal of average h

Then the selection criterion is

$$P(i \in A_t) = \text{Probability of } DU_i \text{ being less than or equal to } h_t$$
$$= \exp(-\lambda DU_i)$$

where

A_t = choice set of individual t

The discrimination parameter, λ, is taken as a reciprocal of average DU. DU's may be averaged over both all population and all alternatives (λ), over all population for the chosen alternative only (λ chosen), or over all alternatives only for each individual (λ_t).

Assuming there are L governing attributes and evaluations are independent for each attribute,

$$P(i \in A_t) = \exp\left[-\sum_{\ell=1}^{L} \lambda'(DU_i^{\ell})\right] \quad \text{.......... (10)}$$

We assume that the two stages of decision-making are independent and the choice probability is logistic. Then, the integrated process can be expressed after normalization as:

$$P_t(i) = \frac{\exp[v_i - \lambda^{\ell}(DU_i^{\ell})]}{\sum_{k=1}^{n} \exp[v_k - \lambda^{\ell}(DU_k^{\ell})]} \quad \text{.......... (14)}$$

We end up with a compact model similar to a conventional logit form. However, the proposed model includes an explict term which accounts for choice set effects in terms of discriminating utilities.

Explict Two Stage Process

As before, we assume the existence of individual specific indifference range (h), specific realizations of h for alternative i as DU_i, and distributions of h over population as F(h). Both h and DU_i are bounded as to (0 to +∞). We have n alternatives and hence n realizations of h (i.e., n DU's) for a given individual. The individual can now rank n alternatives in the order of increasing DU_i. The increasing rank accounts for decreasing attractiveness for the alternative by the definition curve of h in order to represent a population share of each DU.

Suppose DU^k denotes the discriminating utilities of the alternative ranked k and there are n alternatives, and suppose that rank 1 and rank 2 alternatives only are selected through the screening process while all other alternatives are excluded. This condition requires that h for this individual should be greater than or equal to DU (2) and less than DU (3). Then the probability of selecting all alternatives ranked up to k, is expressed as

$$P(A_t^k) = F(DU_t^{k+1}) - F(DU_t^k) \quad \text{......} \quad (12)$$

where

A_t^k = reduced choice set of individual t, which includes all choices ranked up to k

F = distribution function of h

As a result, n choice-set events are generated. Conditioned on each of these events, a specific choice of a particular alternative i is made. Assuming again that the choice set generation and the choice are independent, the choice probability, $P_t(1)$ is given by

$$P_t(i) = \sum_{k=1}^{n} P(A_t^k) \cdot P(i \mid A_t^k) \quad \text{......} \quad (13)$$

For simplicity in the multi-attribute case, it is assumed that strict non-compensation between attributes is relaxed so that individuals cut off alternatives based on the total discriminating utility defined by the sum of weighted discriminating utilities for all governing attributes. That is,

$$DU_{it} = \sum_{\ell=1}^{L} \alpha_\ell \; Z^\ell{}_i \exp(B'Y_t) \quad \text{......} \quad (14)$$

where DU_{it} = total discriminating utility of individual t for alternative i

ℓ = number of choice-set governing attributes

= $1, 2, \ldots, \ell \ldots L$

$Z\ell_i$ = attribute score of alternative i for attribute ℓ

α_ℓ = weight for attribute ℓ

Y_t = individual and environmental characteristics vector of individual t

β' = parameter vector of Y_t

Assumed values of α's and β's will determine relative rankings of alternatives in terms of DU's. These rankings remain fixed for a given individual, but vary over individuals.

Consider now the case where there are too many alternatives, and individuals deal with a small subset of those while excluding all other alternatives. Since, in this case, we are interested only in those alternatives with discriminating utilities less than or equal to DU^k, of the alternative ranked k_o, the individuals do so by specifying $F(DU^o) = \ell$. Assuming an exponential distribution of h, the probability of reduced choice sets containing the k highest ranked alternatives is expressed as

$$P(A_t^k) = \exp[-\lambda (DU_t^k) - \exp[-\lambda (DU_t^{k+1})] \quad \text{......} \quad (15)$$

Let i(r) denote the index of alternative i in association with its ranking r, then the choice probability of alternative i subject to $k > 2$ is

$$P_t[i(r)] = \sum_{k=2}^{n} P(A_t^k) \cdot P[i(r) A_t^k] \quad \text{......} \quad (16)$$

k has been set to no less than 2 since a choice situation becomes meaningful when there are at least two alternatives available.

The process discussed above explains how the individual reduces the global choice set then makes a final choice from the reduced set. This may be true in a situation with a relatively small global set such as mode choice. However, in a situation of a very large global set like destination choice, it would be more realistic to assume that the individual deals with those alternatives only ranked up to k_o, we have the following choice probability.

$$P_t\,[(r<k_o)] = \sum_{k=2}^{k_o}\left\{ e^{-\lambda\,(DU_t^k)} \cdot [1 - e^{-\lambda\,(DU_t^{k+1} - DU_t^k)}]\right\} \frac{e^{v_i(r<k_o)}}{\sum_{R=1}^{k} e^{v_{J(R)}}} \quad \cdots\cdots (17)$$

where

$$DU_t = \sum_{j=1}^{J} \alpha_j Z^j \exp(\beta' Y_t)$$

Although the number of alternatives the individual usually deals with may be limited to two or three by his capacity for information processing, the value of k sub-zero is arbitary. One systematic way of determining the value of k sub-zero is to find the one which produces the maximum fit to the data.

V. VALIDATION AND TESTING

As implied by the above discussion, it is difficult to validate the model based only on the choice set process, because the choice set process itself is depicted as a part of the entire process, and since no observed information is available about individual choice sets. So the validation should be made through the evaluation of the entire model. The five primary evaluation criteria that are most applicable to test the proposed models are as follows.

Model transferability

The basic principle of behavioral disaggregate models is that the demand function determined in a homogenious market is stable over time and space. There always exist some discrepancies between the underlying assumptions of a model and reality. That is, a homogenious market is difficult to attain, the real choice behavior may deviate from the assumed one, the functional forms of the model may not be sufficiently adequate, and the model may have important variables missing. These reasons presumably explain why there has been only limited success in model transferability.

The proposed hypothesis appears closer to the real world choice behavior than the conventional one since it incorporates a process to explain differences in environmental factors and in the socio-economic differences among individuals. Therefore, the propose model, when properly specified, is expected to improve model transferability.

Parameter estimates

The quality of parameter estimates is usually explained by sign correctness, statistical significance, and reasonableness of realtive values. All other things being equal, however, more realistic behavioral hypothesis can upgrade the quality of parameter estimates.

Choice probability

As discussed earlier, the consideration of the individual choice set is expected to identify better discrimination between alternatives through the :screening; process, and hence to reveal a higher probability of an alternative being chosen in the estimation process and a higher probability value for the most probable alternative during the prediction process.

Prediction ability

One method of assessing the fit of a model is to examine the successfulness of prediction by alternative and overall. There are two ways to compute the index of the success prediction, unit-weighted percentage correctly predicted, and probabiliy-weighted index. Both of these indices are expected to be better in the proposed hypothesis.

Log-likelihood of convergence

This represents an overall criterion for goodness of fit and has been commonly used to test the importance of particular sets of variables and the significance of the model as a whole.

A low level of goodness of fit is explained by missing variables, functional specification errors, and errors in the variables. Although a more realistic behavioral assumption is expected to produce a higher level of goodness of fit, the improvement would not be guaranteed unless these factors are improved in a given situation. Nonetheless, this index in conjunction with prediction ability will improve the goodness of fit, relative to the conventional model, when the complexity of the decision space of the individuals increases, since as the choice set reduction process reduces erroneous decision prediction by alleviating the complicated choice situation.

VI. CONCLUSION

Equations (11) and (17) represent wo methods of integrating the choice set process with the choice process. Although they appear complex, both models can be calibrated from typical travel demand origin-destination data. Although it is beyond the scope of this paper to report specific results, we have tested these models with typical modal split data and found that they were superior to the conventional model according to all five of the evaluation criteria.

The significant contribution is not only the improvement in prediction for the "few choice" case such as modal split, but more importantly, these methodologies open the way for the application of choice models to the "large choice" case which heretofore has been nearly impossible.

REFERENCES

1. Ben-Akiva, M. (1979). Issues in Transferring and Updating Travel Behavior Models. Fourth International Conference on Behavioral Travel Modeling, Germany.
2. Horowitz, L. (1979). Sources of Error and Uncertainty in Behavioral Travel Demand Models. Fourth Interntional Conference on Behavioral Travel Modeling, Germany.
3. Krishnan, K. (July 1977). Incorporating Thresholds of Indifference in Probabilistic Choice Models. Management Science, Vol. 23, No. 11.
4. McFadden, D. and Reid, F. (1975). Aggregate Travel Demand Forecasting from Disaggregated Behavioral Models. *T. R. Record 534,* TRB, Washington, D.C.
5. Miller, G. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review,* No. 63.