Biases in Discrete Choice Models

Hong K LO, Quentin K. WAN
Department of Civil Engineering
Hong Kong University of Science and Technology
cehklo@ust.hk

In discrete choice models, the hypothesis is that individuals' preferences for each alternative can be depicted by the alternative's utility measure and that people choose the alternative with the highest utility. However, utilities cannot be entirely observed or measured. Traditionally, utility is modeled as a random variable. Therefore, one does not have a deterministic solution depicting the choices; instead, choices are expressed in probability terms.

The utility measure associated with each alternative is divided into two parts: a systematic (deterministic) component, $V_k(\mathbf{a})$, and a random "error term", ξ_k, that is,

$$U_k(\mathbf{a}) = V_k(\mathbf{a}) + \xi_k \, \forall k \in \kappa \,, \tag{1}$$

where $\mathbf{a}$ denotes the vector of variables, including characteristics and attributes associated with the alternative and those of the decision maker; k represents an alternative in the choice set; the random variable ξ_k captures the unobserved attributes associated with alternative k. $U_k(\mathbf{a})$ is referred to as the "perceived utility", and $V_k(\mathbf{a})$ as the "measured utility".

The probability that alternative k is chosen, P_k, is related to $\mathbf{a}$, which can be expressed as the probability that $U_k(\mathbf{a})$ is higher than the utility of any other alternatives (given $\mathbf{a}$), that is:

$$\begin{aligned} P_k(\mathbf{a}) &= \Pr\left[U_k(\mathbf{a}) \geq U_l(\mathbf{a}), \forall l \in \kappa\right] \\ &= \Pr\left[\xi_k - \xi_l \geq V_l(\mathbf{a}) - V_k(\mathbf{a}), \forall l \in \kappa\right] \end{aligned} \tag{2}$$

By specifying the distribution of the error term, ξ_k, one can determine (2). Note that the term $V_l(\mathbf{a}) - V_k(\mathbf{a})$ in (2) is deterministic. $P_k(\mathbf{a})$ depends on the error-difference distribution $\xi_k - \xi_l$.

Associated with each set of alternatives is a satisfaction function, which captures the expected utility for the set of alternatives. Since each individual selects the alternative with the maximum utility based on the realization of the error, the satisfaction function associated with the choice set is defined as the expectation of the maximum utility alternative whose expectation is taken over the error distribution. That is:

$$\tilde{S} = E\left[\max_{\forall k}\{U_k\}\right] \tag{3}$$

with the expectation taken with respect to the distribution of $\mathbf{U}$ and hence to that of ξ, expressed as:

$$\tilde{S}(\mathbf{V}) = E\left[\max_{\forall k}\{V_k + \xi_k\}\right] \tag{4}$$

Again, the error distribution has a key role in determining the satisfaction function. Much of the above is taken from Ben-Akiva and Lerman (1985), and is intended to form a background for the following discussion.

Sheffi (1985) depicted the satisfaction function associated with the choice set for the logit model, often referred to as the log-sum term:

$$\tilde{S}(\mathbf{V}) = \ln \sum_k e^{V_k} \tag{5}$$

Let us apply this satisfaction function for the choice of two modes (M1 and M2) between an origin-destination (OD) pair. For simplicity, let us assume their utility functions as:

$$\begin{matrix} V_{\mathrm{M1}} = -c_1 \\ V_{\mathrm{M2}} = -c_2 \end{matrix}, \; c_1 = \$0.2;\, c_2 = \$0.1. \tag{6}$$

c_1, c_2 are the travel costs on M1 and M2, respectively. The measured utilities of both mode choices are negative as higher costs are more undesirable. Putting (6) to (5), the satisfaction of this choice set of two modes becomes:

$$\tilde{S}(\mathbf{V}) = \ln\left(e^{-0.2} + e^{-0.1}\right) = 0.54$$

Note that the sign of the satisfaction function associated with the choice set becomes positive, opposite to the measured utilities of both choices. This cannot be a correct representation of actual behavior as cost is not a positive utility. This paradoxical illustration is not limited to the logit model. One can find similar results with the probit model and other error distributions. The root cause is due to the way the error distributions are used in determining the satisfaction function of the choice set, as will be discussed in this paper.

Typically the error term distribution has a mean of zero and has tails in both the positive and negative domains. In other words, they may either add to or subtract from the measured utility $V_k(\mathbf{a})$ to form the perceived utility $U_k(\mathbf{a})$. Therefore, depending on its realization, $U_k(\mathbf{a})$ is sometimes greater than, sometimes less than $V_k(\mathbf{a})$. Nevertheless, in determining the satisfaction function as in (3) and (4), the error term is always used in a positive way to boost up $U_k(\mathbf{a})$ in the maximization operation. In effect, the term $\left[\max_{\forall k}\{V_k + \xi_k\}\right]$ will take on the value of the alternative that has the largest $U_k(\mathbf{a})$ (which could be due to the boosting effect by a large positive error). In this sense, the larger is the spread of the error distribution toward the positive domain, the higher is the resultant satisfaction function. We refer to this effect in boosting $U_k(\mathbf{a})$ as the bias. In the paradox shown above, the error term over-boosts $U_k(\mathbf{a})$, shifting it to the positive domain. One can verify that for relatively small $V_k(\mathbf{a})$ as compared with the magnitude of the error term distribution, this reversal of sign occurs frequently. In the above example, the sign is reversed if $\left(e^{-c_1} + e^{-c_2}\right) > 1$.

In this study, we will provide a detailed discussion on this bias for a number of error distributions and provide numerical examples for this illustration. We will also propose a set of behavioral biases to explain or avoid this phenomenon. Through this discussion, we hope that new ways will be opened up to reconsider fundamental characteristics of discrete choice models.

References

Ben-Akiva, M. and S. Lerman. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. The MIT Press, Cambridge, Massachusetts.

Sheffi, Y. 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall, Inc. NJ.