Discrete Choice Methods with Simulation

Kenneth Train
University of California, Berkeley
National Economic Research Associates

Version dated March 8, 2002

Publisher: Cambridge University Press
Scheduled publication date: Autumn 2002.

Please contact me with any corrections, comments,
and suggestions, at train@econ.berkeley.edu
or 415-291-1023.

4.4.2 Generalized nested logit

Nests of alternatives are labeled $B_1, B_2, \ldots, B_K$. Each alternative can be a member of more than one nest. Importantly, an alternative can be in a nest to varying degrees. Stated differently, an alternative is allocated among the nests, with the alternative being in some nests more than other nests. An "allocation" parameter α_{jk} reflects the extent to which alternative j is a member of nest k. This parameter must be non-negative: $\alpha_{jk} \geq 0 \ \forall j, k$. A value of zero means that the alternative is not in the nest at all. Interpretation is facilitated by having the allocation parameters sum to one over nests for any alternative: $\sum_k \alpha_{jk} = 1 \ \forall j$. Under this condition, α_{jk} reflects the portion of the alternative that is allocated to each nest.

A parameter λ_k is defined for each nest and serves the same function as in nested logit models, namely to indicate the degree of independence among alternatives within the nest: higher λ_k translates into greater independence and less correlation.

The probability that person n chooses alternative i is

$$P_{ni} = \frac{\sum_k (\alpha_{ik} e^{V_{ni}})^{1/\lambda_k} \left(\sum_{j \in B_k} (\alpha_{jk} e^{V_{nj}})^{1/\lambda_k} \right)^{\lambda_k - 1}}{\sum_{\ell=1}^K \left(\sum_{j \in B_\ell} (\alpha_{j\ell} e^{V_{nj}})^{1/\lambda_\ell} \right)^{\lambda_\ell}}. \tag{4.7}$$

This formula is similar to the nested logit probability given in equation 4.2, except that the numerator is a sum over all the nests that contains alternative i, with weights applied to these nests. If each alternative enters only one nest, with $\alpha_{jk} = 1$ for $j \in B_k$ and zero otherwise, the model becomes a nested logit model. And if, in addition, $\lambda_k = 1$ for all nests, then the model becomes standard logit. Wen and Koppelman (2001) derive various cross-nested models as special cases of the GNL.

To facilitate interpretation, the GNL probability can be decomposed as:

$$P_{ni} = \sum_k P_{ni|B_k} P_{nk},$$

where the probability of nest k is

$$P_{nk} = \frac{\sum_j (\alpha_{jk} e^{V_{nj}})^{1/\lambda_k}}{\sum_{\ell=1}^K \left(\sum_{j \in B_\ell} (\alpha_{j\ell} e^{V_{nj}})^{1/\lambda_\ell} \right)^{\lambda_\ell}}$$

and the probability of alternative i given nest k is

$$P_{ni|B_k} = \frac{(\alpha_{ik} e^{V_{ni}})^{1/\lambda_k}}{\sum_j (\alpha_{jk} e^{V_{nj}})^{1/\lambda_k}}.$$

4.5 Heteroskedastic Logit

Instead of capturing correlations among alternatives, the researcher may simply want to allow the variance of unobserved factors to differ over alternatives. Steckel and Vanhonacker (1988), Bhat (1995), and Recker (1995) describe a type of GEV model, called "heteroskedastic extreme value" (HEV), that is the same as logit except with different variance for each alternative. Utility is specified as $U_{nj} = V_{nj} + \varepsilon_{nj}$ where ε_{nj} is distributed independently extreme value with variance $(\theta_j \pi)^2/6$. There is no correlation in unobserved factors over alternatives; however, the variance of the unobserved factors is different for different alternatives. To set the overall scale of utility, the variance for one alternative is normalized to $\pi^2/6$, which is the variance of the standardized extreme value distribution. The variances for the other alternatives are then estimated relative to the normalized variance.

The choice probabilities for this heteroskedastic logit are ((Bhat, 1995)):

$$P_{ni} = \int \left[\prod_{j \neq i} e^{-e^{-(V_{ni} - V_{nj} + \theta_i w)/\theta_j}} \right] e^{-e^{-w}} e^{-w} dw$$

where $w = \varepsilon_{ni}/\theta_i$. The integral does not take a closed form; however, it can be approximated by simulation. Note that $exp(-exp(-w))exp(-w)$ is the extreme value density, given in section 3.1. P_{ni} is therefore the integral of the term in square brackets over the extreme value density. It can be simulated as follows. (1) Take a draw from the extreme value distribution, using the procedure described in section 9.2.3. (2) For this draw of w, calculate the term in brackets, namely: $\prod_{j \neq i} exp(-exp(-(v_{ni} - V_{nj} + \theta_i w)/\theta_j))$. (3) Repeat steps 1 and 2 many times and average the results. This average is an approximation to P_{ni}. Bhat (1995) shows that, since the integral is only one-dimensional, the heteroskedastic logit probabilities can be calculated effectively with quadrature rather than simulation.