Discrete Choice with Random Utility

Toshihiko Mukoyama

April 22, 2019, revised May 20, 2019

This note summarizes some results for the discrete choice model with random utility, developed by McFadden (1973) and Rust (1987). Some parts of the derivation follow the web appendix of Artuç et al. (2010). Similar results as (a subset of) Result 3 appear in Iskhakov et al. (2017).

Result 1 below is the standard Multinomial Logit result, although most of the current applications seem to focus on the case with $\sigma = 1$. Result 2 is an extension of the standard result that tends to focus on the $\sigma = 1$ case. Result 3 shows, intuitively, that the outcome with random utility converges to the outcome without the taste shocks as the variance of the shocks goes to zero. This result is convenient because it means that, with a small value of σ, we can have "smooth" expression of $\mathbf{V}$ (i.e. the expression that doesn't involve "kinks" with respect to state variables that influence $\{V_i\}$, associated with the max operator) which is still close to $\max\{V_1, V_2, ..., V_n\}$. The Result 3 also provides a natural tie-breaking rule when there are multiple options that provides the max.

1 Setting

Consider a the choice of n alternatives:

$$\mathbf{V} = E_\varepsilon[\max\{V_1 + \varepsilon_1, V_2 + \varepsilon_2, ..., V_n + \varepsilon_n\}].$$

Here, V_i is the utility from alternative i. The random variable ε_i is mean zero and i.i.d. across i. Assume that ε_i follows type-I extreme value (Gumbel) distribution. The expectation $E_\varepsilon[\cdot]$ is taken for the vector $\varepsilon = \{\varepsilon_1, \varepsilon_2, ..., \varepsilon_n\}$. The distribution function is given by

$$F(\varepsilon_i) = \exp\left(-\exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right)\right),$$

where $\gamma = 0.5772...$ is Euler's constant. $\sigma > 0$ is the scale parameter, and the variance of the distribution is increasing in σ. The density function is

$$f(\varepsilon_i) = \frac{1}{\sigma}\exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma - \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right)\right).$$

2 Results

The following results hold.

Result 1: The probability that the option i is chosen, p_i, is

$$p_i = \frac{\exp\left(\frac{V_i}{\sigma}\right)}{\sum_{k=1}^{n} \exp\left(\frac{V_k}{\sigma}\right)}.$$

Result 2: The resulting expected utility is

$$\mathbf{V} = \sigma \log\left(\sum_{k=1}^{n} \exp\left(\frac{V_k}{\sigma}\right)\right).$$

Result 3: As $\sigma \to 0$, the limiting outcome is that $p_i = 1$ if $V_i > V_k$ for all k. In this case, $p_k = 0$ for all other k. When there are multiple options that yield the maximum i, $p_i = 1/m$ for these options, where m is the number of ties. As $\sigma \to \infty$, $p_i \to 1/n$ for all i. As $\sigma \to 0$, $\mathbf{V} \to \max\{V_1, V_2, ..., V_n\}$.

3 Proofs

3.1 Proof of Result 1

The probability that i is chosen can be computed from

$$\begin{aligned} p_i &= \Pr\left[V_i + \varepsilon_i \geq V_k + \varepsilon_k \text{ for all } k\right] \\ &= \Pr\left[\varepsilon_k \leq V_i + \varepsilon_i - V_k \text{ for all } k\right] \\ &= \int_{-\infty}^{\infty} f(\varepsilon_i) \prod_{k \neq i} F(V_i + \varepsilon_i - V_k) d\varepsilon_i. \end{aligned}$$

In the integral, $\prod_{k\neq i} F(V_i+\varepsilon_i-V_k)$ is the probability that, when ε_i is given, ε_k is below the value of $V_i + \varepsilon_i - V_k$ for all $k \neq i$. Multiplying the density for corresponding ε_i and integrating over all possible ε_i provides the probability that the option i is chosen.

Let $x_i \equiv \varepsilon_i/\sigma + \gamma$. Then

$$\begin{aligned}
p_i &= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma - \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right)\right) \prod_{k\neq i} \exp\left(-\exp\left(-\frac{V_i + \varepsilon_i - V_k}{\sigma} - \gamma\right)\right) d\varepsilon_i \\
&= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma - \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right)\right) \exp\left(-\sum_{k\neq i} \exp\left(-\frac{V_i + \varepsilon_i - V_k}{\sigma} - \gamma\right)\right) d\varepsilon_i \\
&= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right) \exp\left(-\sum_{k=1}^{n} \exp\left(-\frac{V_i + \varepsilon_i - V_k}{\sigma} - \gamma\right)\right) d\varepsilon_i \\
&= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left[-\frac{\varepsilon_i}{\sigma} - \gamma - \sum_{k=1}^{n} \exp\left(-\frac{V_i + \varepsilon_i - V_k}{\sigma} - \gamma\right)\right] d\varepsilon_i \\
&= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left[-\frac{\varepsilon_i}{\sigma} - \gamma - \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right) \sum_{k=1}^{n} \exp\left(-\frac{V_i - V_k}{\sigma}\right)\right] d\varepsilon_i \\
&= \frac{1}{\sigma}\int_{-\infty}^{\infty} \exp\left[-\frac{\varepsilon_i}{\sigma} - \gamma - \exp\left(-\frac{\varepsilon_i}{\sigma} - \gamma\right) \left(\sum_{k=1}^{n} \exp\left(\frac{V_k}{\sigma}\right)\right) \Big/ \left(\exp\left(\frac{V_i}{\sigma}\right)\right)\right] d\varepsilon_i \\
&= \int_{-\infty}^{\infty} \exp\left[-x_i - \exp(-x_i) \left(\sum_{k=1}^{n} \exp\left(\frac{V_k}{\sigma}\right)\right) \Big/ \left(\exp\left(\frac{V_i}{\sigma}\right)\right)\right] dx_i.
\end{aligned}$$

Let

$$\lambda_i \equiv \log\left[\left(\sum_{k=1}^{n} \exp\left(\frac{V_k}{\sigma}\right)\right) \Big/ \left(\exp\left(\frac{V_i}{\sigma}\right)\right)\right]. \qquad (1)$$

Then, defining $y_i \equiv x_i - \lambda_i$,

$$\begin{aligned}
p_i &= \int_{-\infty}^{\infty} \exp\left[-x_i - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\
&= \exp(-\lambda_i) \int_{-\infty}^{\infty} \exp\left[-(x_i - \lambda_i) - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\
&= \exp(-\lambda_i) \int_{-\infty}^{\infty} \exp\left[-y_i - \exp(-y_i)\right] dy_i.
\end{aligned}$$

Note that

$$\frac{d}{dy_i} \exp\left(-\exp(-y_i)\right) = \exp\left(-y_i - \exp(-y_i)\right).$$

Thus

$$\begin{aligned} p_i &= \exp(-\lambda_i) \int_{-\infty}^{\infty} \exp\left(-y_i - \exp\left(-y_i\right)\right) dy_i \\ &= \exp(-\lambda_i) \left[\exp\left(-\exp\left(-y_i\right)\right)\right]_{-\infty}^{\infty} \\ &= \exp(-\lambda_i) \\ &= \frac{\exp\left(\dfrac{V_i}{\sigma}\right)}{\displaystyle\sum_{k=1}^{n} \exp\left(\dfrac{V_k}{\sigma}\right)}. \end{aligned}$$

3.2 Proof of Result 2

From the definition of $\mathbf{V}$,

$$\mathbf{V} = \sum_{i=1}^{n} \int_{-\infty}^{\infty} (V_i + \varepsilon_i) f(\varepsilon_i) \prod_{k \neq i} F(V_i + \varepsilon_i - V_k) d\varepsilon_i.$$

With the same transformation as above (and noting that $\varepsilon_i = \sigma(x_i - \gamma)$),

$$\begin{aligned} \mathbf{V} &= \sum_{i=1}^{n} \int_{-\infty}^{\infty} (V_i + \sigma(x_i - \gamma)) \exp\left[-x_i - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\ &= \sum_{i=1}^{n} \left[(V_i - \sigma\gamma) \exp(-\lambda_i) + \sigma \int_{-\infty}^{\infty} x_i \exp\left[-x_i - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \right]. \end{aligned}$$

Note that

$$\begin{aligned} &\int_{-\infty}^{\infty} x_i \exp\left[-x_i - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\ &\quad = \exp(-\lambda_i) \textstyle\int_{-\infty}^{\infty} x_i \exp\left[-(x_i - \lambda_i) - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\ &\quad = \exp(-\lambda_i) \textstyle\int_{-\infty}^{\infty} (x_i - \lambda_i) \exp\left[-(x_i - \lambda_i) - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\ &\qquad + \exp(-\lambda_i) \textstyle\int_{-\infty}^{\infty} \lambda_i \exp\left[-(x_i - \lambda_i) - \exp\left(-(x_i - \lambda_i)\right)\right] dx_i \\ &\quad = \gamma \exp(-\lambda_i) + \lambda_i \exp(-\lambda_i) \end{aligned}$$

holds. Here, I used the fact that (Patel et al. (1976, p.35))

$$\int_{-\infty}^{\infty} y_i \exp\left[-y_i - \exp\left(-y_i\right)\right] dy_i = \gamma.$$

(Once again, here $y_i \equiv x_i - \lambda_i$).

Thus

$$\begin{aligned}\mathbf{V} &= \sum_{i=1}^{n}\left[(V_i - \sigma\gamma)\exp(-\lambda_i) + \sigma(\gamma+\lambda_i)\exp(-\lambda_i)\right] \\ &= \sum_{i=1}^{n}(V_i + \sigma\lambda_i)\exp(-\lambda_i).\end{aligned}$$

From the expression of λ_i in (1),

$$\begin{aligned}\sigma\lambda_i &= \sigma\log\left[\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right)\Big/\left(\exp\left(\frac{V_i}{\sigma}\right)\right)\right] \\ &= \sigma\left[\log\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right) - \frac{V_i}{\sigma}\right] \\ &= \sigma\log\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right) - V_i\end{aligned}$$

Therefore

$$\begin{aligned}\mathbf{V} &= \sum_{i=1}^{n}\sigma\log\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right)\exp(-\lambda_i) \\ &= \sigma\log\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right)\sum_{i=1}^{n}p_i \\ &= \sigma\log\left(\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)\right).\end{aligned}$$

3.3 Proof of Result 3

From the expression of p_i,

$$\begin{aligned}p_i &= \frac{\exp\left(\frac{V_i}{\sigma}\right)}{\sum_{k=1}^{n}\exp\left(\frac{V_k}{\sigma}\right)} \\ &= \frac{1}{1+\sum_{k\neq i}\exp\left(\frac{V_k - V_i}{\sigma}\right)}.\end{aligned}$$

As $\sigma \to 0$, (i) if $V_k < V_i$ for all $k \neq i$, $p_i \to 1$; (ii) if $V_k > V_i$ for some $k \neq i$, $p_i \to \infty$; and (iii) if $V_k = V_i$ for some $k \neq i$ (and there are $m-1$ such k) and $V_k < V_i$ for the rest of $k \neq i$, $p_i = 1/m$. As $\sigma \to \infty$, $\exp\left(\frac{V_k - V_i}{\sigma}\right) \to 1$ for all k and thus $p_i \to 1/n$.

From the expression of $\mathbf{V}$,

$$
\begin{aligned}
\mathbf{V} &= \sigma \log \left(\sum_{k=1}^{n} \exp \left(\frac{V_k}{\sigma} \right) \right) \\
&= \sigma \log \left(\sum_{k=1}^{n} \exp \left(\frac{V_k}{\sigma} \right) \right) - V_i + V_i \\
&= \sigma \log \left(\sum_{k=1}^{n} \exp \left(\frac{V_k}{\sigma} \right) \right) - \sigma \log \left(\exp \left(\frac{V_i}{\sigma} \right) \right) + V_i \\
&= \sigma \log \left(\sum_{k=1}^{n} \exp \left(\frac{V_k - V_i}{\sigma} \right) \right) + V_i \\
&= \sigma \log \left(1 + \sum_{k \neq i} \exp \left(\frac{V_k - V_i}{\sigma} \right) \right) + V_i,
\end{aligned}
$$

where V_i is chosen as $\max\{V_1, V_2, ..., V_n\}$. The first term converges to zero as $\sigma \to 0$. Thus $\mathbf{V} \to V_i$.

References

[1] Artuç, E., S. Chaudhuri, and J. McLaren (2010). "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review* 100, 1008–1045.

[2] Iskhakov, F., T. H. Jørgensen, J. Rust, and B. Schjerning (2017). "The Endogenous Grid Method for Discrete-Continuous Dynamic Choice Models with (or without) Taste Shocks," *Quantitative Economics* 8, 317–365.

[3] McFadden, D. (1973). "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka (ed.) *Frontiers of Econometrics*, New York: Academic Press.

[4] Patel, J. K., C. H. Kapadia, and D.B. Owen (1976). *Handbook of Statistical Distributions*, New York: Marcel Dekker.

[5] Rust, J. (1987). "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55, 999-1033.