Some useful properties and formulas for random utility models with logit, nested logit, and ordered nested logit stochastic components

Victor Aguirregabiria*
University of Toronto

January 8, 2010

Abstract

Within the framework of discrete choice Random Utility Models (RUM) with additive stochastic components, this note reviews existing results on closed-form expressions for several key functions: the distribution of the maximum utility, the expected maximum utility, the choice probabilities, and the selection function. The analysis considers three different specifications for the distribution of the stochastic component: i.i.d. type I extreme value distribution, nested extreme value distribution, and ordered generalized extreme value distribution.

1 Random Utility Models

Consider a discrete choice Random Utility Model (RUM) with additive stochastic component. See McFadden (1974, 1981) for seminal descriptions of the RUM, and Anderson, De Palma, and Thisse, 1992) for a thorough analysis of these models containing some of the results in this note.

The optimal choice, a^*, is defined as:

$$a^* = \arg\max_{a \in \mathcal{A}} \ \{u_a + \varepsilon_a\} \tag{1}$$

where $\mathcal{A} = \{1, 2, ..., J\}$ is the set of feasible choice alternatives, $\mathbf{u} = (u_1, u_2, ..., u_J)$ is the vector with the deterministic component of the utility, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_J)$ is the vector

*Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

with the stochastic component. The vector $\boldsymbol{\varepsilon}$ has a joint CDF $G(.)$ that is continuous and strictly increasing with respect to the Lebesgue measure in the Euclidean space.

This note derives closed-form expressions for:

1. The probability distribution of the maximum utility, $\max_{a\in\mathcal{A}} \{u_a + \varepsilon_a\}$.

2. The expected maximum utility, $\mathbb{E}(\max_{a\in\mathcal{A}} \{u_a + \varepsilon_a\} |u)$.

3. The expected value of ε_a conditional on alternative a being optimal: $\mathbb{E}(\varepsilon_a \mid a^* = a)$.

4. The choice probabilities, $\Pr(a^* = a|\mathbf{u})$,

under three different specifications for the distribution of the vector $\boldsymbol{\varepsilon}$:

a. i.i.d. Type I Extreme Value distribution: multinomial logit RUM.

b. i.i.d. nested Extreme Value distribution: nested logit RUM.

c. i.i.d. Ordered Generalized Extreme Value distribution: OGEV RUM.

The following definitions and properties are used in the note.

Definition: *A random variable X has a Type I Extreme Value distribution (also denoted Gumbel or Double Exponential distribution) with location parameter μ and dispersion parameter σ if its CDF is:*

$$G(x) = \exp\left\{ - \exp\left(- \left[\frac{x-\mu}{\sigma} \right] \right) \right\} \tag{2}$$

for any $x \in (-\infty, +\infty)$. ■

Definition: *Maximum utility. Let v^* be the random variable that represents the maximum utility: $v^* \equiv \max_{a\in\mathcal{A}} \{u_a + \varepsilon_a\}$. This maximum utility is a random variable because it depends on the vector of random variables ε.* ■

Definition: *McFadden's Social Surplus function. The social surplus function $S(\mathbf{u})$ is the expected value of the maximum utility conditional on the vector of constants $\mathbf{u}$: $S(\mathbf{u}) \equiv \mathbb{E}(\max_{a\in\mathcal{A}} \{u_a + \varepsilon_a\} |u)$.* ■

Definition: *Conditional choice probabilities (CCPs). The conditional choice probability $P(a|\mathbf{u})$ is the probability that alternative a is the optimal choice: $P(a|\mathbf{u}) \equiv Pr(a^* = a|\mathbf{u})$.* ■

Definition: *Conditional choice expected utilities (CCEU). The conditional choice expected utility $e(a, \mathbf{u})$ is the expected value of utility $u_a + \varepsilon_a$ conditional on the vector $\mathbf{u}$ and on the event that alternative a is the optimal choice: $e(a, \mathbf{u}) \equiv \mathbb{E}(u_a + \varepsilon_a|\mathbf{u}, a^* = a)$.* ■

Definition: *Selection-bias function. The selection function $\lambda(a, \mathbf{u})$ is the expected value of the stochastic component of the utility, ε_a, conditional on the vector $\mathbf{u}$ and on the event that alternative a is the optimal choice: $\lambda(a, \mathbf{u}) \equiv \mathbb{E}(\varepsilon_a|\mathbf{u}, a^* = a)$.* ■

2 Williams-Daly-Zachary Theorem

Williams-Daly-Zachary (WDZ) Theorem is an important property of discrete choice RUM with additive stochastic component. It is the discrete-choice version of Roy's Identity in consumer theory. I use this property in several parts of this note. I include here an enunciation of the Theorem and a simple proof.

Williams-Daly-Zachary (WDZ) Theorem. *For any choice alternative $a \in \mathcal{A}$, the CCP $P(a|\mathbf{u})$ can be obtained as the partial derivative of the surplus function $S(\mathbf{u})$ with respect to utility $u(a)$:*

$$P(a|\mathbf{u}) = \frac{\partial S(\mathbf{u})}{\partial u_a} \qquad \blacksquare \tag{3}$$

Proof: By definition of $S(\mathbf{u})$, we have that:

$$\frac{\partial S(\mathbf{u})}{\partial u_a} = \frac{\partial}{\partial u_a} \int \max_{j \in \mathcal{A}} \{u_j + \varepsilon_j\} \ dG(\boldsymbol{\varepsilon}) \tag{4}$$

Given the conditions on the CDF of $\boldsymbol{\varepsilon}$, we can move the partial derivative inside the integral such that:

$$\begin{aligned} \frac{\partial S(\mathbf{u})}{\partial u_a} &= \int \frac{\partial \max_{j \in \mathcal{A}} \{u_j + \varepsilon_j\}}{\partial u_a} \, dG(\boldsymbol{\varepsilon}) \\ &= \int 1\{u_a + \varepsilon_a \geq u_j + \varepsilon_j, \ \forall j \in \mathcal{A}\} \, dG(\boldsymbol{\varepsilon}) \\ &= P(a|\mathbf{u}) \end{aligned} \tag{5}$$

where $1\{.\}$ is the indicator function. ■

3 Multinomial logit (MNL)

Suppose that the random variables in the vector $\boldsymbol{\varepsilon}$ are i.i.d. with Type I Extreme Value distribution with a location parameter $\mu = 0$ and unrestricted dispersion parameter σ. That is, for every alternative $a \in \mathcal{A}$, the CDF of ε_a is $G(\varepsilon_a) = \exp\left\{-\exp\left(-\frac{\varepsilon_a}{\sigma}\right)\right\}$.

3.1 Distribution of the maximum utility

The maximum utility v^* is a random variable because it depends on the vector of random variables ε. By definition, the cumulative probability distribution of v^* is:

$$\begin{aligned} F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \prod_{a\in\mathcal{A}} \Pr(u_a + \varepsilon_a \leq v) \\ &= \prod_{a\in\mathcal{A}} \exp\left\{-\exp\left(-\frac{v-u_a}{\sigma}\right)\right\} \\ &= \exp\left\{-\exp\left(-\frac{v}{\sigma}\right)\ U\right\} \end{aligned} \tag{6}$$

where $U \equiv \sum_{a\in\mathcal{A}} \exp\left(\frac{u_a}{\sigma}\right)$. We can also write this expression as:

$$F_{v^*}(v) = \exp\left\{-\exp\left(-\frac{v-\sigma\ \ln U}{\sigma}\right)\right\} \tag{7}$$

This expression shows that the maximum utility v^* is a double exponential random variable with dispersion parameter σ and location parameter $\sigma\ \ln U$. Therefore, the maximum of a vector of i.i.d. double exponential random variables is also a double exponential random variable. This is the reason why this family of random variables is also called "extreme value". The density function of v^* is:

$$f_{v^*}(v) \equiv H'(v) = F_{v^*}(v)\ \frac{U}{\sigma}\ \exp\left(-\frac{v}{\sigma}\right) \tag{8}$$

3.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*|\mathbf{u})$. Therefore,

$$S(\mathbf{u}) = \int v^*\ h(v^*)\ dv^* = \int v^*\ \exp\left\{-\exp\left(-\frac{v^*}{\sigma}\right) U\right\}\ \frac{U}{\sigma}\ \exp\left(-\frac{v^*}{\sigma}\right)\ dv^* \tag{9}$$

Applying the change in variable $z = \exp(-v^*/\sigma)$, such that $v^* = -\sigma \ln(z)$, and $dv^* = -\sigma(dz/z)$, we have:

$$\begin{aligned} S(\mathbf{u}) &= \int_{+\infty}^{0} -\sigma \ln(z) \ \exp\{-z\ U\} \ \frac{U}{\sigma} \ z \ \left(-\sigma \frac{dz}{z}\right) \\ &= -\sigma U \int_{0}^{+\infty} \ln(z) \ \exp\{-z\ U\} \ dz \end{aligned} \tag{10}$$

Using Laplace transformation we have that $\int_0^{+\infty} \ln(z) \exp\{-z\ U\}\ dz = \frac{\ln(U)+\gamma}{U}$, where γ is Euler's constant. Therefore, the expected maximum utility is:

$$S(\mathbf{u}) \ = \ \sigma U \ \left(\frac{\ln(U)+\gamma}{U}\right) \ = \ \sigma \ (\ \ln(U)+\gamma \ ) \tag{11}$$

3.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, the optimal choice probabilities can be obtained by differentiating the surplus function. Therefore, for the MNL model,

$$\begin{aligned} P(a|\mathbf{u}) &= \sigma \ \frac{\partial \ln(U)}{\partial u_a} \qquad = \sigma \ \frac{\partial U}{\partial u_a} \ \frac{1}{U} \\ &= \exp\left(\frac{u_a}{\sigma}\right) \ \frac{1}{U} \ = \frac{\exp\left(u_a/\sigma\right)}{\sum_{j \in \mathcal{A}} \exp\left(u_j/\sigma\right)} \end{aligned} \tag{12}$$

3.4 Selection-Bias function

In this section, I derive the density function of ε_a conditional on the event $a^* = a$. I show that this conditional density has the following form:

$$f_{\varepsilon_a|a^*=a}(\varepsilon_a) \ = \ \exp\Big\{ - (\varepsilon_a - \ln P(a \mid \mathbf{u})) - \exp\big\{ - (\varepsilon_a - \ln P(a \mid \mathbf{u})) \big\}\Big\} \tag{13}$$

This is the density of a Type 1 Extreme Value random variable with location parameter $\mu = \ln P(a \mid \mathbf{u})$. By definition, the mean of this random variable is the selection-bias function and is equal to $\gamma - \ln P(a \mid \mathbf{u})$. I prove this result below.

The event $a^* = a$ is equivalent to $\varepsilon_j \leq \varepsilon_a + u_a - u_j, \quad \forall j \neq a$. Therefore, the marginal conditional density $f(\varepsilon_a \mid a^* = a)$ is

$$f_{\varepsilon_a|a^*=a}(\varepsilon_a) \ = \ \frac{f(\varepsilon_a) \cdot \prod_{j \neq a} \Pr\left(\varepsilon_j \leq \varepsilon_a + u_a - u_j\right)}{\Pr(a^* = a)} \tag{14}$$

Replacing $f(\varepsilon_a)$ with the density of the Type 1 Extreme Value random variable, and replacing $\Pr(\varepsilon_j \leq \varepsilon_a + u_a - u_j)$ with its CDF evaluated at $\varepsilon_a + u_a - u_j$. we have:

$$\begin{aligned} f_{\varepsilon_a|a^*=a}(\varepsilon_a) &= \frac{\exp(-\varepsilon_a - \exp(-\varepsilon_a)) \cdot \prod_{j\neq a} \exp\left(-\exp(-(\varepsilon_a + u_a - u_j))\right)}{P(a \mid \mathbf{u})} \\ &= \frac{\exp(-\varepsilon_a) \cdot \prod_{j=1}^{J} \exp\left(-\exp(-(\varepsilon_a + u_a - u_j))\right)}{P(a \mid \mathbf{u})} \\ &= \frac{\exp(-\varepsilon_a) \cdot \exp\left(-\exp(-\varepsilon_a) \cdot \sum_{j=1}^{J} \exp(u_j - u_a)\right)}{P(a \mid \mathbf{u})} \end{aligned} \tag{15}$$

Define: $U = \sum_{j=1}^{J} \exp(u_j)$, so $P(a \mid \mathbf{u}) = \exp(u_a)/U$. Using this definition, we can rewrite the marginal conditional density $f(\varepsilon_a \mid a^* = a)$ as:

$$\begin{aligned} f_{\varepsilon_a|a^*=a}(\varepsilon_a) &= \frac{\exp(-\varepsilon_a) \cdot \exp\left(-\exp(-\varepsilon_a) \cdot U/\exp(u_a)\right)}{\exp(u_a)/U} \\ &= \exp\left\{ -(\varepsilon_a - \ln P(a \mid \mathbf{u})) - \exp\left\{ -(\varepsilon_a - \ln P(a \mid \mathbf{u})) \right\} \right\} \end{aligned} \tag{16}$$

As mentioned above, this is the density of a Type 1 Extreme Value random variable with location parameter $\mu = \ln P(a \mid \mathbf{u})$. Therefore,

$$\lambda(a, \mathbf{u}) = \mathbb{E}(\varepsilon_a \mid \mathbf{u}, a^* = a) = \gamma - \log P(a \mid \mathbf{u}) \tag{17}$$

4 Nested logit (NL)

Suppose that the random variables in the vector ε have the following joint CDF:

$$G(\boldsymbol{\varepsilon}) = \exp\left\{ -\sum_{r=1}^{R} \left[\sum_{a\in\mathcal{A}_r} \exp\left(-\frac{\varepsilon_a}{\sigma_r}\right)\right]^{\frac{\sigma_r}{\delta}} \right\} \tag{18}$$

where $\{\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_R\}$ is a partition of $\mathcal{A}$, and δ, σ_1, σ_2, ..., σ_R are positive parameters, with $\delta \leq 1$.

4.1 Distribution of the Maximum Utility

Using the same approach as for the MNL model, we have:

$$\begin{aligned}
F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \prod_{a\in\mathcal{A}} \Pr(u_a + \varepsilon_a \leq v, \forall\, a \in \mathcal{A}) \\
&= \prod_{a\in\mathcal{A}} \exp\left\{ -\sum_{r=1}^{R} \left[\sum_{a\in A_r} \exp\left(-\frac{v-u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ -\exp\left(-\frac{v}{\delta}\right) \sum_{r=1}^{R} \left[\sum_{a\in\mathcal{A}_r} \exp\left(\frac{u_a}{\sigma_r}\right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ -\exp\left(-\frac{v}{\delta}\right) U \right\}
\end{aligned} \tag{19}$$

where:

$$U \equiv \sum_{r=1}^{R} \left[\sum_{a\in\mathcal{A}_r} \exp\left(\frac{u_a}{\sigma_r}\right) \right]^{\frac{\sigma_r}{\delta}} = \sum_{r=1}^{R} U_r^{1/\delta} \tag{20}$$

and

$$U_r \equiv \left[\sum_{a\in\mathcal{A}_r} \exp\left(\frac{u_a}{\sigma_r}\right) \right]^{\sigma_r} \tag{21}$$

The density function of v^* is:

$$f_{v^*}(v) \equiv H'(v) \;=\; F_{v^*}(v)\, \frac{U}{\delta}\, \exp\left(-\frac{v}{\delta}\right) \tag{22}$$

4.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*)$. Therefore,

$$S(\mathbf{u}) \;=\; \int_{-\infty}^{+\infty} v^*\, h(v^*)\, dv^* \;=\; \int_{-\infty}^{+\infty} v^*\, \exp\left\{ -\exp\left(-\frac{v^*}{\delta}\right) U \right\} \frac{U}{\delta}\, \exp\left(-\frac{v^*}{\delta}\right) dv^* \tag{23}$$

Let's apply the following change in variable: $z = \exp(-v^*/\delta)$, such that $v^* = -\delta \ln(z)$, and $dv^* = -\delta(dz/z)$. Then,

$$S(\mathbf{u}) \;=\; \int_{+\infty}^{0} -\delta \ln(z)\, \exp\{-z\, U\}\, \frac{U}{\delta}\, z \left(-\delta \frac{dz}{z}\right) \;=\; -\delta U \int_{+\infty}^{0} \ln(z)\, \exp\{-z\, U\}\, dz \tag{24}$$

And using Laplace transformation:

$$S(\mathbf{u}) \;=\; \delta U \left(\frac{\ln(U) + \gamma}{U} \right) \;=\; \delta \;(\; \ln(U) + \gamma \;) \tag{25}$$

where γ is the Euler's constant.

4.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, choice probabilities can be obtained differentiating the surplus function. For the NL model:

$$\begin{aligned} P(a|\mathbf{u}) &= \delta \frac{\partial \ln(U)}{\partial u_a} = \delta \frac{\partial U}{\partial u_a} \frac{1}{U} = \\ &= \delta \frac{\sigma_{ra}}{\delta} \left[\sum_{j \in A_{ra}} \exp\left(\frac{u_j}{\sigma_{ra}} \right) \right]^{\frac{\sigma_{ra}}{\delta} - 1} \frac{1}{\sigma_{ra}} \exp\left(\frac{u_a}{\sigma_{ra}} \right) \frac{1}{U} \\ &= \frac{\exp\left(u_a/\sigma_{ra}\right)}{\sum_{j \in \mathcal{A}_{ra}} \exp\left(u_j/\sigma_{ra}\right)} \; \frac{\left[\sum_{j \in A_{ra}} \exp\left(u_j/\sigma_{ra}\right) \right]^{\frac{\sigma_{ra}}{\delta}}}{\sum_{r=1}^{R} \left[\sum_{j \in \mathcal{A}_r} \exp\left(u_j/\sigma_r\right) \right]^{\frac{\sigma_r}{\delta}}} \end{aligned} \tag{26}$$

The first term is $q(a|r_a)$ (i.e., probability of choosing a given that we are in group A_{ra}), and the second term is $Q(r_a)$ (i.e., probability of selecting the group A_{ra}).

4.4 Conditional choice expected utilities

As shown in general, $e(a, \mathbf{u}) = S(\mathbf{u})$. This implies that $\mathbb{E}(\varepsilon_a \mid u,\, a^* = a) = S(\mathbf{u}) - u_a$. Given that for the NL model $S(\mathbf{u}) = \delta \left(\ln U + \gamma\right)$ we have that:

$$\mathbb{E}(\varepsilon_a | u, a^* = a) = \delta\gamma + \delta \ln U - u_a \tag{27}$$

4.5 Relationship between selection function and CCPs

To write $\mathbb{E}(\varepsilon_a | u, a^* = a)$ in terms of choice probabilities, note that from the definition of $q(a|r_a)$ and $Q(r_a)$, we have that:

$$\ln q(a|r_a) \;=\; \frac{u_a - \ln U_{ra}}{\sigma_{ra}} \;\Rightarrow\; \ln U_{ra} \;=\; u_a - \sigma_{ra} \ln q(a|r_a) \tag{28}$$

and

$$\ln Q(r_a) \;=\; \frac{\ln U_{ra}}{\delta} - \ln U \;\;\Rightarrow\;\; \ln U \;=\; \frac{\ln U_{ra}}{\delta} - \ln Q(r_a) \tag{29}$$

Combining these expressions, we have that:

$$\ln U = \frac{u_a - \sigma_{ra} \ln q(a|r_a)}{\delta} - \ln Q(r_a) \tag{30}$$

Therefore,

$$\begin{aligned} e_a &= \delta\gamma + \delta \left(\frac{u_a - \sigma_{ra} \ln q(a|r_a)}{\delta} - \ln Q(r_a) \right) - u_a \\ &= \delta\gamma - \sigma_{ra} \ \ln q(a|r_a) - \delta \ \ln Q(r_a) \end{aligned}$$

5 Ordered GEV (OGEV)

Suppose that the random variables in the vector $\boldsymbol{\varepsilon}$ have the following joint CDF:

$$G(\boldsymbol{\varepsilon}) = \exp\left\{ -\sum_{r=1}^{J+M} \left[\sum\nolimits_{a \in B_r} W_{r-a} \exp\left(-\frac{\varepsilon_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \tag{31}$$

where:

- M is a positive integer;
- $\{B_1, B_2, ..., B_{J+M}\}$ are $J + M$ subsets of A, with the following definition:

  $$B_r = \{a \in \mathcal{A} : r - M \leq a \leq r\} \tag{32}$$

  For instance, if $A = \{1, 2, 3, 4, 5\}$ and $M = 2$, then $B_1 = \{1\}$, $B_2 = \{1, 2\}$, $B_3 = \{1, 2, 3\}$, $B_4 = \{2, 3, 4\}$, $B_5 = \{3, 4, 5\}$, $B_6 = \{4, 5\}$, and $B_7 = \{5\}$.
- δ, and σ_1, σ_2, ..., σ_{J+M} are positive parameters, with $\delta \leq 1$;
- W_0, W_1, ..., W_M are constants (weights) such that: $W_m \geq 0$, and $\sum_{m=0}^{M} W_m = 1$.

5.1 Distribution of the Maximum Utility

$$
\begin{aligned}
F_{v^*}(v) \equiv \Pr(v^* \leq v) &= \Pr(\varepsilon_a \leq v - u_a : for\ any\ a \in \mathcal{A}) \\
&= \exp\left\{ -\sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(-\frac{v - u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ -\exp\left(-\frac{v}{\delta} \right) \sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right\} \\
&= \exp\left\{ -\exp\left(-\frac{v}{\delta} \right) U \right\}
\end{aligned}
\tag{33}
$$

where:

$$
U \equiv \sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} = \sum_{r=1}^{J+M} U_r^{1/\delta} \tag{34}
$$

where $U_r \equiv \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\sigma_r}$. The density function of v^* is:

$$
f_{v^*}(v) \equiv H'(v) = F_{v^*}(v) \frac{U}{\delta} \exp\left(-\frac{v}{\delta} \right) \tag{35}
$$

5.2 Expected maximum utility

By definition, $S(\mathbf{u}) = \mathbb{E}(v^*|u)$. Therefore,

$$
S(\mathbf{u}) = \int_{-\infty}^{+\infty} v^* \ h(v^*) \ dv^* = \int_{-\infty}^{+\infty} v^* \ \exp\left\{ -\exp\left(-\frac{v^*}{\delta} \right) U \right\} \frac{U}{\delta} \ \exp\left(-\frac{v^*}{\delta} \right) \ dv^* \tag{36}
$$

Let's apply the following change in variable: $z = \exp(-v^*/\delta)$, such that $v^* = -\delta \ln(z)$, and $dv^* = -\delta(dz/z)$. Then,

$$
S = \int_{+\infty}^{0} -\delta \ln(z) \ \exp\{-z \ U\} \ \frac{U}{\delta} \ z \ \left(-\delta \frac{dz}{z} \right) = -\delta U \int_{0}^{+\infty} \ln(z) \ \exp\{-z \ U\} \ dz \tag{37}
$$

And using Laplace transformation:

$$
S = \delta U \left(\frac{\ln U + \gamma}{U} \right) = \delta \left(\ln U + \gamma \right) = \delta\gamma + \delta \ \ln \left[\sum_{r=1}^{J+M} \left[\sum_{a \in B_r} W_{r-a} \exp\left(\frac{u_a}{\sigma_r} \right) \right]^{\frac{\sigma_r}{\delta}} \right] \tag{38}
$$

where γ is the Euler's constant.

5.3 Choice probabilities

By Williams-Daly-Zachary (WDZ) theorem, choice probabilities can be obtained differentiating the surplus function.

$$P(a|u) = \frac{1}{U} \sum_{r=a}^{a+M} \left[\sum_{j \in B_r} W_{r-j} \exp\left(\frac{u_j}{\sigma_r}\right) \right]^{\frac{\sigma_r}{\delta}-1} W_{r-a} \exp\left(\frac{u_a}{\sigma_r}\right) = \sum_{r=a}^{a+M} q(a|r)\ Q(r) \tag{39}$$

where:

$$\begin{aligned} q(a|r) &= \frac{W_{r-a}\ \exp\left(u_a/\sigma_r\right)}{\sum_{j \in B_r} W_{r-j} \exp\left(u_j/\sigma_r\right)} = \frac{\exp\left(u_a/\sigma_r\right)}{\exp(\ln U_r/\sigma_r)} \\ Q(r) &= \frac{\exp\left(\ln U_r/\delta\right)}{\sum_{j=1}^{J+M} \exp\left(\ln U_j/\delta\right)} = \frac{\exp\left(\ln U_r/\delta\right)}{U} \end{aligned} \tag{40}$$

5.4 Conditional choice expected utilities

As shown in general, $e(a, \mathbf{u}) = S(\mathbf{u})$. This implies that $\mathbb{E}(\varepsilon_a \mid u,\, a^* = a) = S(\mathbf{u}) - u_a$. Given that for the OGEV model $S(\mathbf{u}) = \delta\left(\ln U + \gamma\right)$ we have that:

$$\mathbb{E}(\varepsilon_a | u, a^* = a) = \delta\gamma + \delta \ln U - u_a \tag{41}$$

References

Anderson, S. P., A. De Palma, and J.-F. Thisse (1992): *Discrete choice theory of product differentiation.* MIT press.

McFadden, D. (1974): “Conditional logit analysis of qualitative choice behavior,” *in Frontiers in Econometrics, P. Zarembka (ed.) Academic Press: New York*, pp. 105–142.

——— (1981): “Econometric Models of Probabilistic Choice,” *in C. Manski and D. McFadden (eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA.*