Simulating Nested Logit Draws

Max Maydanchik, University of Chicago

August 8, 2023

Overview

These notes were developed for Dingel and Tintelnot (2020). I hope it is helpful to others, given the usefulness of simulations in economic modeling.

Many economics papers work with discrete choice models and nested logit errors. In these models, each choice has a component drawn from a T1EV distribution that is correlated between choices that share nests. Nesting structures can be arbitrarily deep.

Surprisingly, until recently, it was not known how to simulate from these distributions in closed-form, at least beyond a single-level nest. (For example, the R package `mvevd` can simulate single-level nests, but grows unwieldy beyond that. Approximations using the characteristic function were possible.)

In this note, I illustrate a simple method to simulate exact, closed-form nested logit errors.

The representation result

Galichon (2022) provides a proof of the Cardell (1997) theorem for multi-level nested logit models. Stephenson (2003) provides a similar proof for a single-level nest, along with a better method to simulate draws. (For transparency: Galichon (2022) refers to Ridout (2009) for a simulation method, though I believe it is more complex than necessary; I should also note that Bunch and Rocke (2016) derive an approximate method matching moments of the characteristic function but not the exact representation.) The core representation result is this:

$$\epsilon \sim T1EV(0,1), \quad Z \sim PS(\alpha)$$
$$\implies \alpha\left(\epsilon + \log Z\right) \sim T1EV(0,1)$$

where $PS(\alpha)$ is the positive stable distribution of parameter α, given by (from Kanter (1975))

$$U \sim U[0,\pi], \quad W \sim \exp(1)$$
$$\implies \left(\frac{\sin((1-\alpha)U)}{W}\right)^{\frac{1-\alpha}{\alpha}} \frac{\sin(\alpha U)}{(\sin U)^{\frac{1}{\alpha}}} \sim PS(\alpha)$$

That is, the linear combination of a standard Gumbel and a log positive stable random variable is itself a standard Gumbel random variable.

```r
# Define stable distribution draws function
rstable <- function(n, alpha) {

    # Can avoid computation if alpha = 1
    if (alpha == 1) {
        return(rep(1, n))
    }

    # Draw from U[0,pi] distribution
    u <- runif(n) * pi

    # Draw from exp(1) distribution
    w <- rexp(n, 1)

    # Compute stable distribution
    calpha <- 1-alpha
    x <- (sin(calpha * u)/w)^(calpha / alpha) * sin(alpha * u) / sin(u)^(1/alpha)

    return(x)
}

# Define parameters
set.seed(1)
n <- 50000
alpha <- 0.5

# Draw from stable distribution
Z <- rstable(n, alpha)

# Draw from Gumbel distribution
library(evd)
e <- rgumbel(n, 0, 1)

# Represenation result
x <- alpha * (e + log(Z))

# Plot histogram with gumbel(0,1) pdf overlain
library(ggplot2)

ggplot(data.frame(x), aes(x)) +
  geom_histogram(aes(y = ..density..), bins = 50, fill = "white", color = "black") +
  geom_function(fun = function(x) dgumbel(x, 0, 1), color = "red") +
  theme_bw()
```

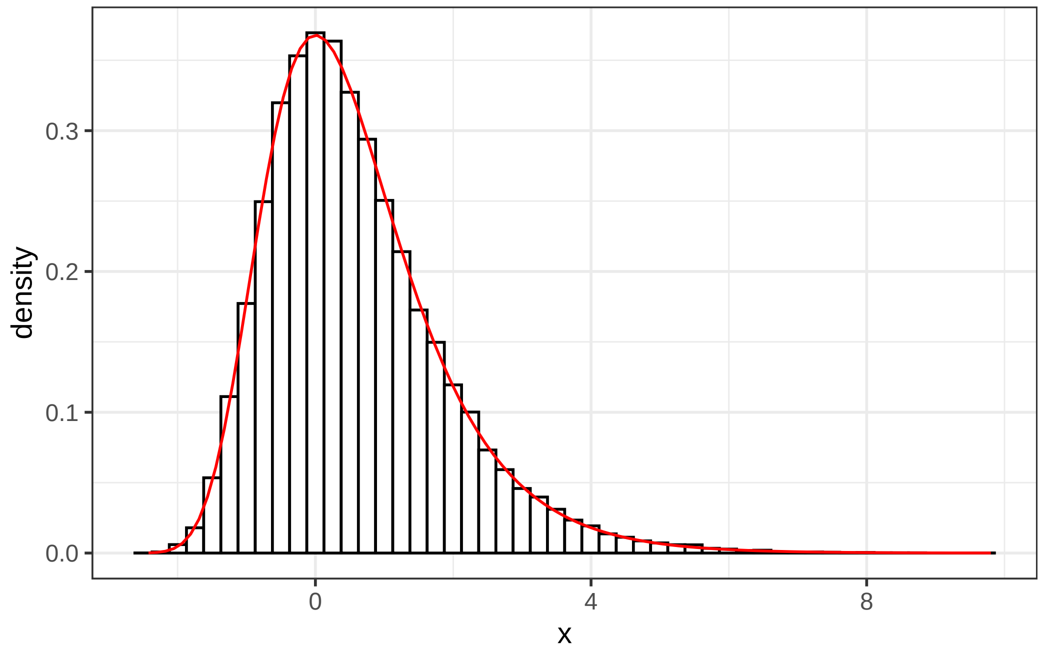

Figure 1: Illustrating the representation result. The histogram takes 50,000 draws from PS and T1EV distributions, the red line is the T1EV pdf.

Nested logit errors

The Cardell (1997) and Galichon (2022) results imply the following form for nested logit errors:

$$\varepsilon_j = \sum_{t=1}^{d_{n_j}} \Lambda_{n_j(t)} log Z_{n_j(t)} + \Lambda_{n_j} \epsilon_j$$

where the notation represents the tree structure of the nested logit.

- j: the choice index, making up the "leaves" of the tree
- n_j: the parent node of choice j, corresponding to nests
- d_n: the depth of node n from the tree top ($n = 0$), where the root node has depth $d_0 = 0$
- λ_n: the nest parameter for the given node, and $\lambda_0 = 1$ for the root node
- Λ_n is the product of nest parameters along the direct path from the root node to node n

Examples

Commonly provided forms of nested logit errors follow neatly.

1. Take a simple example of a single-level nested logit with product j in nest g. The error structure is

$$\varepsilon_j = \zeta_g + \lambda_g \epsilon_j$$

where $\zeta_g \equiv \lambda_g \log Z_g$ is a group-specific term, with $Z_g \sim PS(\lambda_g)$ and $\epsilon_j \sim T1EV(0, 1)$.

2. Take a two-level nested logit with product j in subgroup hg within group g. The error structure is
$$\varepsilon_j = \zeta_g + \lambda_g \zeta_{hg} + \lambda_{hg}\lambda_g \epsilon_j$$

Galichon (2022) uses the representation result to proof the correlation between any two choices, j and j' is $1 - \Lambda_n^2$, where n is the lowest common ancestor of j and j'. In example 1, denoting $g(j)$ as the nest of choice j, this is:

$$corr(\varepsilon_j, \varepsilon_{j'}) = \begin{cases} 1 - \lambda_g^2, & g(j) = g(j') \\ 0, & g(j) \neq g(j') \end{cases}$$

In example 2, denoting $g(j)$ and $hg(j)$ as the nest and subnest of choice j, respectively, this is:

$$corr(\varepsilon_j, \varepsilon_{j'}) = \begin{cases} 1 - (\lambda_g \lambda_{hg})^2, & hg(j) = hg(j'), \\ 1 - \lambda_g^2, & g(j) = g(j'), \quad hg(j) \neq hg(j') \\ 0, & g(j) \neq g(j') \end{cases}$$

Simulating nested logit errors

To illustrate this, we'll simulate a two-level nested logit model with 12 products, $j = 0, ..., 11$ (where 0 is the outside good), where subnests group items $\mathcal{H} = \{\{0\}, \{1\}, \{2, 3, 4\}, \{5, 6\}, \{7\}, \{8, 9, 10, 11\}\}$, and the nests groups items $\mathcal{G} = \{\{0\}, \{1, 2, 3, 4, 5, 6\}, \{7, 8, 9, 10, 11\}\}$.

McFadden (1978) proves that nest parameters are consistent with utility maximizing if $1 > \lambda_g \geq \lambda_{hg} > 0$, so we'll follow that here. We will define the nesting parameters asymmetrically to illustrate the flexibility of the simulation. Define the group parameters (λ_g) as $\{1, 0.89, 0.76\}$ and the subgroup parameters (λ_{hg}) as $\{1, 1, 0.53, 0.44, 1, 0.28\}$. (The outside good, and any degenerate nests, has $\lambda = 1$.)

```
# Define the nest and subnest structure
j <- 12
g <- c(1,2,2,2,2,2,2,3,3,3,3,3)
h <- c(1,2,3,3,3,4,4,5,6,6,6,6)

# Define the parameters
set.seed(1)
draws <- 50000
lambda_g <- c(1, 0.89, 0.76)
lambda_h <- c(1, 1, 0.53, 0.44, 1, 0.28)

# Draw from the stable distribution
Z_g = matrix(0, nrow = draws, ncol = length(lambda_g))
for (i in 1:length(lambda_g)) {
    Z_g[,i] <- rstable(draws, lambda_g[i])
}
logZ_g <- log(Z_g)
```

```r

Z_h = matrix(0, nrow = draws, ncol = length(lambda_h))
for (i in 1:length(lambda_h)) {
    Z_h[,i] <- rstable(draws, lambda_h[i])
}
logZ_h <- log(Z_h)

# Draw from T1EV distribution
e_j = matrix(0, nrow = draws, ncol = j)
for (i in 1:j) {
    e_j[,i] <- rgumbel(draws, 0, 1)
}

# Compute the deviations
eps <- matrix(0, nrow = draws, ncol = length(g))
for (i in 1:j) {
    eps[,i] <- lambda_g[g[i]] * logZ_g[,g[i]] +
               lambda_h[h[i]] * lambda_g[g[i]] * logZ_h[,h[i]] +
               lambda_h[h[i]] * lambda_g[g[i]] * e_j[,i]
}

# Plot the correlation matrix
library(ggplot2)
library(reshape2)

corr_mat <- cor(eps)
corr_df <- melt(corr_mat)
corr_plot <- ggplot(data = corr_df, aes(x=Var1, y=Var2, fill=value)) +
    theme(axis.title=element_blank()) +
    geom_tile(color="white") +
    geom_text(aes(label = sprintf("%.3f", value)), color = "black", size = 2.5) +
    scale_fill_gradient2(low="red", mid="white", high="blue", limits=c(-1,1), name="Correlation") +
    scale_x_continuous(breaks = 1:j, labels = 0:(j-1)) +
    scale_y_continuous(breaks = 1:j, labels = 0:(j-1))
corr_plot
```

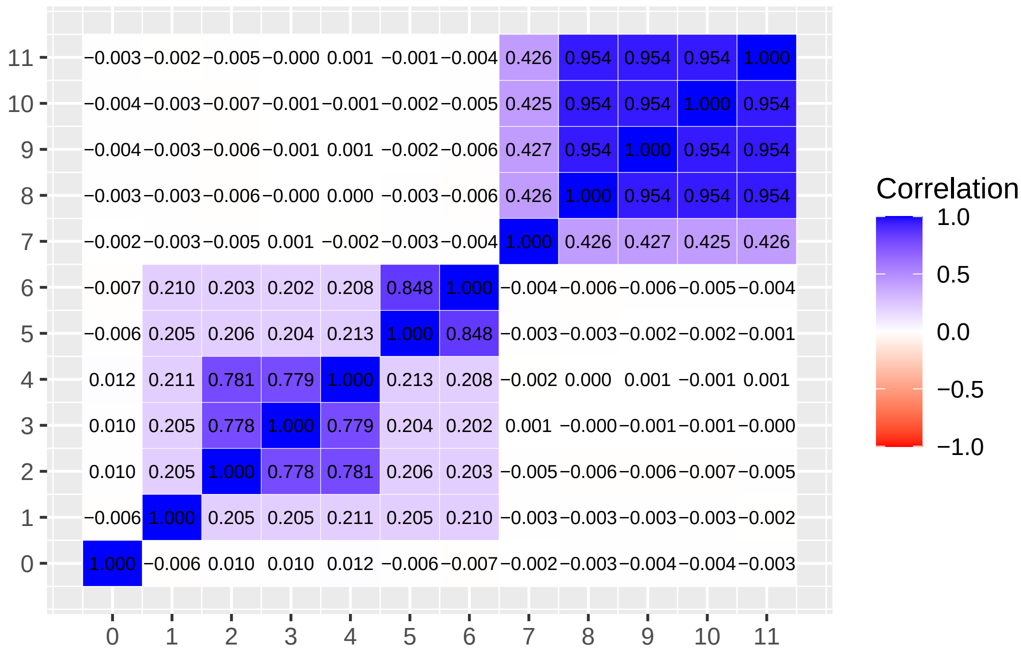

Figure 2: Correlation matrix from simulation of 50,000 draws of nested logit errors.

We expect correlations of:

$$
\begin{aligned}
(1-(0.89)^2) &= 0.208 && \text{between items 1 and 6, items 1 and 4, and items 2 and 6} \\
(1-(0.89\times 0.53)^2) &= 0.777 && \text{between items 2 and 3} \\
(1-(0.89\times 0.44)^2) &= 0.847 && \text{between items 5 and 6} \\
(1-(0.76)^2) &= 0.422 && \text{between items 7 and 11} \\
(1-(0.76\times 0.28)^2) &= 0.955 && \text{between items 8 and 11} \\
(1-(1)^2) &= 0.000 && \text{elsewhere off-diagonal}
\end{aligned}
$$

As we can see the distribution of all the errors is $T1EV(0,1)$.

```r
# Plot the distribution of errors for each product with gumbel pdf overlain
eps_df <- melt(eps)
eps_df$Var2 <- eps_df$Var2 - 1
eps_plot <- ggplot(data = eps_df, aes(x=value)) +
    geom_histogram(aes(y=..density..), bins=50) +
    facet_wrap(~Var2, ncol = 6) +
    theme_bw() +
    theme(axis.title=element_blank()) +
    geom_function(fun = function(x) dgumbel(x, 0, 1), color = "red")
eps_plot
```

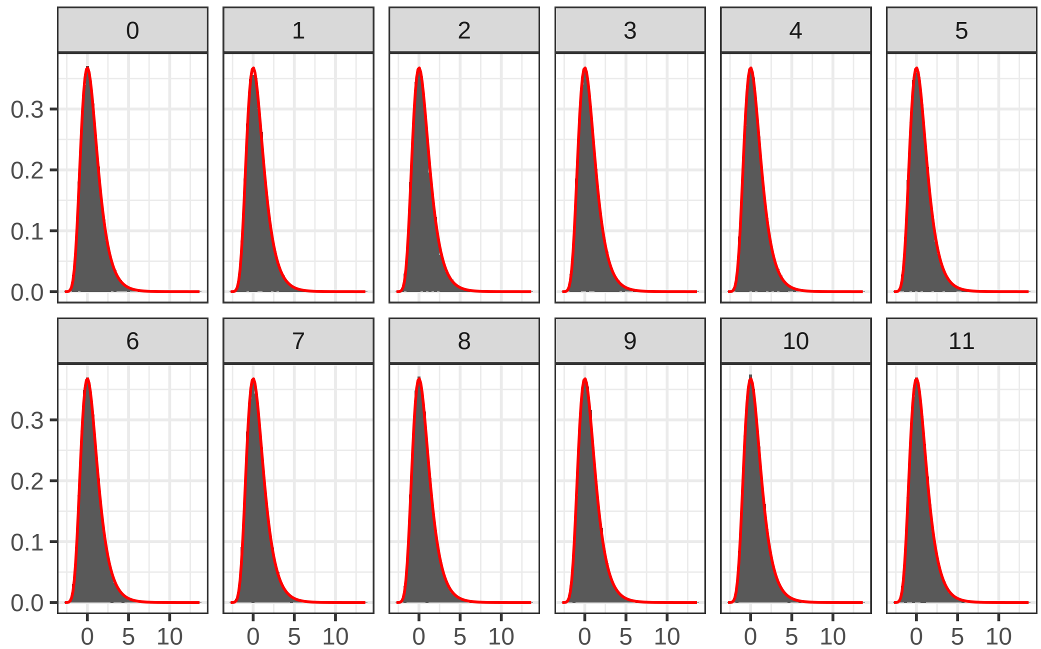

Figure 3: Distribution of errors for each product.

References

Bunch, David S., and David M. Rocke. 2016. "Variance-Component-Based Nested Logit Specifications: Improved Formulation, and Practical Microsimulation of Random Disturbance Terms." *Journal of Choice Modelling* 21: 30–35. https://doi.org/https://doi.org/10.1016/j.jocm.2016.04.001.

Cardell, N. Scott. 1997. "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity." *Econometric Theory* 13 (2): 185–213. https://www.jstor.org/stable/3532724.

Dingel, Jonathan I, and Felix Tintelnot. 2020. "Spatial Economics for Granular Settings." Working Paper 27287. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w27287.

Galichon, Alfred. 2022. "On the Representation of the Nested Logit Model." *Econometric Theory* 38 (2): 370–80. https://doi.org/10.1017/S026646662000047X.

Kanter, Marek. 1975. "Stable Densities Under Change of Scale and Total Variation Inequalities." *The Annals of Probability* 3 (4): 697–707. http://www.jstor.org/stable/2959332.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." Cowles Foundation Discussion Papers 477. Cowles Foundation for Research in Economics, Yale University.

Ridout, M. S. 2009. "Generating Random Numbers from a Distribution Specified by Its Laplace Transform." *Statistics and Computing* 19 (4): 439. https://doi.org/10.1007/s11222-008-9103-x.

Stephenson, Alec. 2003. "Simulating Multivariate Extreme Value Distributions of Logistic Type." *Extremes* 6 (1): 49–59. https://doi.org/10.1023/A:1026277229992.