Multivariate Extreme Value models

Nested and cross-nested logit models

Michel Bierlaire

Mathematical Modeling of Behavior

EPFL

Outline

Nested logit model

Multivariate Extreme Value distribution

Multivariate Extreme Value model

Cross-nested logit model

Case study

Appendix: derivation of the MEV model

Transportation mode choice

- Transportation mode choice between car and bus.
- The bus company happens to have blue buses and red buses.
- We use a logit model.
- Example where the travel times by bus and by car are the same.

Mode choice model

Original model

$$
\begin{aligned}
U_{\text{car}} &= \beta T + \varepsilon_{\text{car}}, \\
U_{\text{bus}} &= \beta T + \varepsilon_{\text{bus}}.
\end{aligned}
$$

Choice probability

ε_i i.i.d. $\text{EV}(0, \mu)$:

$$
P(\text{car}|\{\text{car}, \text{bus}\}) = \frac{e^{\mu\beta T}}{e^{\mu\beta T} + e^{\mu\beta T}} = \frac{1}{2}.
$$

Mode choice model

New model

$$\begin{aligned} U_{\text{car}} &= \beta T + \varepsilon_{\text{car}}, \\ U_{\text{blue bus}} &= \beta T + \varepsilon_{\text{blue bus}}, \\ U_{\text{red bus}} &= \beta T + \varepsilon_{\text{red bus}}. \end{aligned}$$

Choice probability

ε_i i.i.d. $\text{EV}(0, \mu)$:

$$P(\text{car}|\{\text{car, blue bus, red bus}\}) = \frac{e^{\mu\beta T}}{e^{\mu\beta T} + e^{\mu\beta T} + e^{\mu\beta T}} = \frac{1}{3}.$$

Policy implications

- If you paint the buses in a city red and blue, the mode share for cars drops from 50% to 33%.
- It looks like an efficient way to remove congestion from our cities.

Explaining the paradox

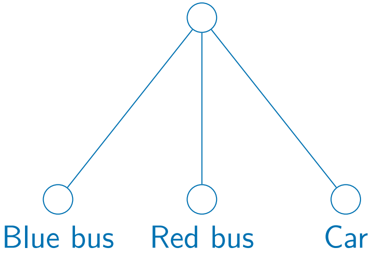
Model specification

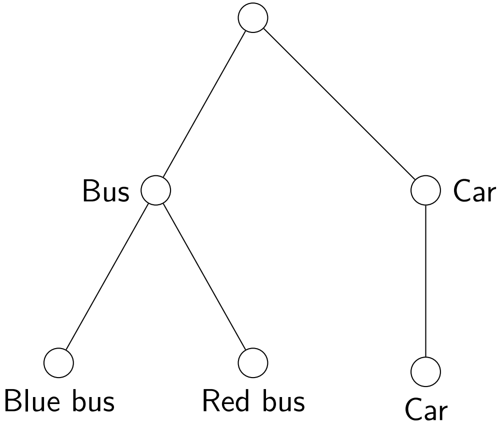
- Only travel time appears in the utility function.
- Other attributes are captured by the error term.
- Some of them are shared by $\varepsilon_{\text{blue bus}}$ and $\varepsilon_{\text{red bus}}$
  - fare
  - headway
  - comfort
  - convenience
  - etc.

Logit model

- Assumes that $\varepsilon_{\text{blue bus}}$ and $\varepsilon_{\text{red bus}}$ are independent.
- Inappropriate assumption in this case.

Transportation mode choice





The nested logit model

$$\begin{aligned} U_{\text{car}} &= \beta T + \varepsilon_{\text{car}}, \\ U_{\text{blue bus}} &= \beta T + \varepsilon_{\text{blue bus}} + \varepsilon_{\text{bus}}, \\ U_{\text{red bus}} &= \beta T + \varepsilon_{\text{red bus}} + \varepsilon_{\text{bus}}. \end{aligned}$$

ε_{bus} captures all the unobserved attributes common to both alternatives.

Within the nest

If bus is chosen then

$$P(\text{blue bus}|\{\text{blue bus}, \text{red bus}\}) = \Pr(\beta T + \varepsilon_{\text{blue bus}} + \cancel{\varepsilon_{\text{bus}}} \geqslant \beta T + \varepsilon_{\text{red bus}} + \cancel{\varepsilon_{\text{bus}}})$$

Choice probability

We assume $\varepsilon_{\text{blue bus}}$ and $\varepsilon_{\text{red bus}}$ i.i.d. $\text{EV}(0, \mu_b)$.

$$P(\text{blue bus}|\{\text{blue bus}, \text{red bus}\}) = \frac{e^{\mu_b \beta T}}{e^{\mu_b \beta T} + e^{\mu_b \beta T}} = \frac{1}{2}.$$

Across nests

What about the choice between bus and car?

$$
\begin{aligned}
U_{\text{car}} &= \beta T + \varepsilon_{\text{car}} \\
U_{\text{bus}} &= V_{\text{bus}} + \varepsilon_{\text{bus}}
\end{aligned}
$$

with

$$
\begin{aligned}
V_{\text{bus}} &= V_{\text{bus}}(V_{\text{blue bus}}, V_{\text{red bus}}) \\
\varepsilon_{\text{bus}} &= ?
\end{aligned}
$$

Idea

- Use a logit model at the higher level.
- Define V_{bus} as the expected maximum utility of red bus and blue bus.

Expected maximum utility

Definition

For a set of alternatives $\mathcal{C}$, define

$$U_{\mathcal{C}} \equiv \max_{i \in \mathcal{C}} U_i = \max_{i \in \mathcal{C}} (V_i + \varepsilon_i),$$

and

$$U_{\mathcal{C}} \equiv V_{\mathcal{C}} + \varepsilon_{\mathcal{C}}.$$

Expected maximum utility

For logit

If ε_i are i.i.d. EV(0, μ_b),

$$U_{\mathcal{C}} = V_{\mathcal{C}} + \varepsilon_{\mathcal{C}},$$
$$V_{\mathcal{C}} = \frac{1}{\mu_b} \ln \sum_{i \in \mathcal{C}} e^{\mu_b V_i},$$
$$E[\varepsilon_{\mathcal{C}}] = \frac{\gamma}{\mu_b}.$$

Expected maximum utility

Back to the blue/red buses

$$
\begin{aligned}
V_{\text{bus}} &= \frac{1}{\mu_b} \ln\left(e^{\mu_b V_{\text{blue bus}}} + e^{\mu_b V_{\text{red bus}}}\right) \\
&= \frac{1}{\mu_b} \ln\left(e^{\mu_b \beta T} + e^{\mu_b \beta T}\right) = \frac{1}{\mu_b} \ln\left(2e^{\mu_b \beta T}\right) \\
&= \beta T + \frac{1}{\mu_b} \ln 2,
\end{aligned}
$$

where μ_b is the scale parameter for the logit model associated with the choice between red bus and blue bus.

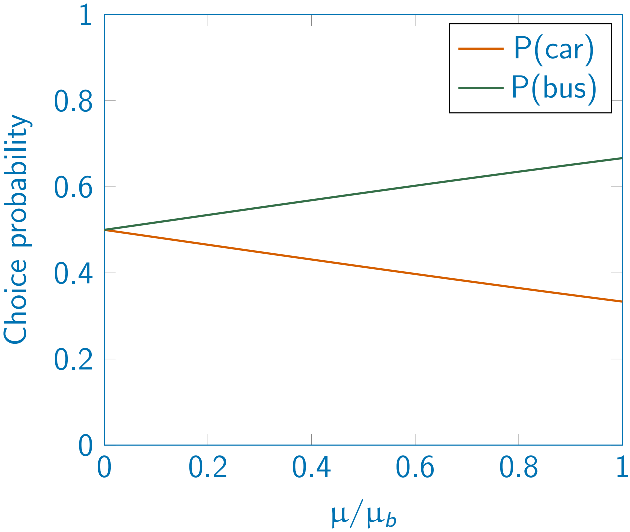
Nested Logit Model

If we assume that ε_{car} and ε_{bus} are i.i.d. EV(0,μ):

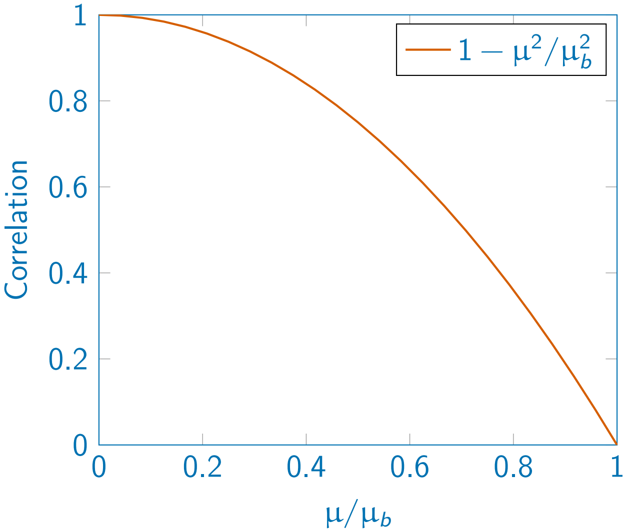
Probability model: car

$$
\begin{aligned}
P(\text{car}) &= \frac{e^{\mu V_{\text{car}}}}{e^{\mu V_{\text{car}}} + e^{\mu V_{\text{bus}}}} \\
&= \frac{e^{\mu \beta T}}{e^{\mu \beta T} + e^{\mu \beta T + \frac{\mu}{\mu_b} \ln 2}} \\
&= \frac{1}{1 + 2^{\frac{\mu}{\mu_b}}}.
\end{aligned}
$$

Nested Logit Model



Nested Logit Model



Nested logit model

Comments

- A group of similar alternatives is called a nest.
- Each alternative belongs to exactly one nest.
- The model is named Nested Logit.
- The ratio μ/μ_b must be estimated from the data.
- $0 < \mu/\mu_b \leqslant 1$.
- Going down the tree, μ's must increase, variance must decrease.

Nested logit model

- Let $\mathcal{C}$ be the choice set.
- Let $\mathcal{C}_1, \ldots, C_M$ be a partition of $\mathcal{C}$.
- The model is derived as

$$P(i|\mathcal{C}) = \sum_{m=1}^{M} \Pr(i|m, \mathcal{C}) \Pr(m|\mathcal{C}).$$

- Each i belongs to exactly one nest m.

$$P(i|\mathcal{C}) = \Pr(i|m) \Pr(m|\mathcal{C}).$$

- Each nest is associated with a scale parameter μ_m and an expected maximum utility:

$$\tilde{V}_m = \frac{1}{\mu_m} \ln \sum_{i \in \mathcal{C}_m} e^{\mu_m V_i}.$$

Nested logit model

- Within nest probability:

$$\Pr(i|m) = \frac{e^{\mu_m V_i}}{\sum_{j \in \mathcal{C}_m} e^{\mu_m V_j}}.$$

- Across nest probability:

$$\Pr(m|\mathcal{C}) = \frac{e^{\mu \tilde{V}_m}}{\sum_{p=1}^{M} e^{\mu \tilde{V}_p}}.$$

- μ is normalized to 1.

Nested logit model

Correlation matrix: block diagonal

$$\begin{pmatrix} 1 & 1-\frac{\mu^2}{\mu_1^2} & \cdots & 0 & 0 & 0 & \cdots \\ 1-\frac{\mu^2}{\mu_1^2} & 1 & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & 1 & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & \cdots & 1 & 1-\frac{\mu^2}{\mu_m^2} & 1-\frac{\mu^2}{\mu_m^2} & \cdots \\ 0 & 0 & \cdots & 1-\frac{\mu^2}{\mu_m^2} & 1 & 1-\frac{\mu^2}{\mu_m^2} & \cdots \\ 0 & 0 & \cdots & 1-\frac{\mu^2}{\mu_m^2} & 1-\frac{\mu^2}{\mu_m^2} & 1 & \cdots \\ \vdots & \vdots & & \cdots\vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Outline

Multivariate Extreme Value distribution

Multivariate distribution

$$\begin{pmatrix} U_{1n} \\ \vdots \\ U_{Jn} \end{pmatrix} = \begin{pmatrix} V_{1n} \\ \vdots \\ V_{Jn} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1n} \\ \vdots \\ \varepsilon_{Jn} \end{pmatrix}$$

that is

$$U_n = V_n + \varepsilon_n,$$

and ε_n is a vector of random variables.

Multivariate Extreme Value distribution

Definition

$$\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{Jn})$$

follows a multivariate extreme value distribution if it has the CDF

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{Jn}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{Jn}})},$$

where $G : \mathbb{R}_+^{J_n} \to \mathbb{R}_+$ is a positive function with positive arguments.

Multivariate Extreme Value distribution

Valid CDF must verify three properties

- $F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, -\infty, \ldots, \varepsilon_{J_n n}) = 0.$
- $F_{\varepsilon_n}(+\infty, \ldots, +\infty) = 1.$
- For any set of $\widehat{J_n} \leqslant J_n$ distinct indices $i_1, \ldots, i_{\widehat{J_n}}$,

$$\frac{\partial^{\widehat{J_n}} F_{\varepsilon_n}}{\partial \varepsilon_{i_1 n} \cdots \partial \varepsilon_{i_{\widehat{J_n}} n}}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) \geqslant 0.$$

The limit property

Valid CDF

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, -\infty, \ldots, \varepsilon_{J_n n}) = 0.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function

$$G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty.$$

The zero property

Valid CDF

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty) = 1.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function

$$G(0, \ldots, 0) = 0.$$

The strong alternating sign property

Valid CDF

$$\frac{\partial^{\widehat{J}_n} F_{\varepsilon_n}}{\partial \varepsilon_{i_1 n} \cdots \partial \varepsilon_{i_{\widehat{J}_n} n}}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) \geqslant 0.$$

MEV

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})}.$$

Valid G function

- The right-hand side changes sign each time it is differentiated.
- To obtain $\geqslant 0$, G must also change sign each time it is differentiated.

The strong alternating sign property

Valid G function

- G must change sign each time it is differentiated.
- For any set of $\widehat{J}_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geqslant 0,$$

  where

$$G_i = \partial G / \partial y_i.$$

Homogeneity

We need another property: homogeneity

A function G is homogeneous of degree μ, or μ-homogeneous, if

$$G(\alpha y) = \alpha^{\mu} G(y), \; \forall \alpha > 0 \text{ and } y \in \mathbb{R}_{+}^{J_n}.$$

It will imply two results

- the marginals are univariate extreme value distributions,
- the choice model has a closed form.

Marginal distribution

ith marginal distribution

$$F_{\varepsilon_n}(+\infty, \dots, +\infty, \varepsilon_{in}, +\infty, \dots, +\infty) = e^{-G(0,\dots,0,e^{-\varepsilon_{in}},0,\dots,0)}.$$

If G is μ-homogeneous, we have

$$G(0, \dots, 0, e^{-\varepsilon_{in}}, 0, \dots, 0) = e^{-\mu\varepsilon_{in}} G(0, \dots, 0, 1, 0, \dots, 0),$$

or equivalently,

$$G(0, \dots, 0, e^{-\varepsilon_{in}}, 0, \dots, 0) = e^{-\mu\varepsilon_{in} + \ln G(0,\dots,0,1,0,\dots,0)},$$

Multivariate Extreme Value distribution

ith marginal: univariate extreme value distribution

Define $\ln G(0, \ldots, 0, 1, 0, \ldots, 0) = \mu\eta$, so that:

$$F_{\varepsilon_n}(+\infty, \ldots, +\infty, \varepsilon_{in}, +\infty, \ldots, +\infty) = \exp\left(-e^{-\mu(\varepsilon_{in}-\eta)}\right).$$

Multivariate Extreme Value distribution

Three conditions on G

- The limit property

$$G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty.$$

- The strong alternating sign property

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geqslant 0.$$

- Homogeneity (which implies the zero property)

$$G(\alpha y) = \alpha^{\mu} G(y), \ \forall \alpha > 0 \text{ and } y \in \mathbb{R}_+^{J_n}.$$

Outline

The random utility model

- Assume that $\varepsilon_n = (\varepsilon_{1n}, \ldots, \varepsilon_{J_n n})$ is a multivariate random variable,
- with CDF

$$F_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}),$$

- and pdf

$$f_{\varepsilon_n}(\varepsilon_1, \ldots, \varepsilon_{J_n}) = \frac{\partial^{J_n} F}{\partial \varepsilon_1 \cdots \partial \varepsilon_{J_n}}(\varepsilon_1, \ldots, \varepsilon_{J_n}).$$

Then $P_n(i|\mathcal{C}_n) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon.$$

Derivation in lecture on Theoretical Foundations.

Multivariate Extreme Value model

Choice model

$P_n(i|\mathcal{C}_n) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon.$$

Multivariate Extreme Value distribution

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})},$$

where G verifies the MEV conditions.

Derivation in the appendix.

MEV: choice model

The multivariate extreme value model:

$$P_n(i) = \frac{e^{V_{in} + \ln G_i\left(e^V\right)}}{\sum_j e^{V_{jn} + \ln G_j\left(e^V\right)}}.$$

where $G_i = \partial G / \partial y_i$, and G verifies

- (i) the limit property: $G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty$.
- (ii) the strong alternating sign property: for any set of $\widehat{J}_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geqslant 0.$$

- (iii) the homogeneity property:

$$G(\alpha y) = \alpha^{\mu} G(y),\ \forall \alpha > 0 \text{ and } y \in \mathbb{R}_+^{J_n}.$$

MEV: choice model

Probability generating function

A function G, which is μ homogeneous, that verifies the MEV properties is called a μ-MEV function.

Expected maximum utility

$$\mathrm{E}[\max_{j \in \mathcal{C}_n} U_{jn}] = \frac{1}{\mu}(\ln G(e^{V_{1n}}, \ldots, e^{V_{J_n n}}) + \gamma),$$

where γ is Euler's constant

Euler's constant

$$\gamma = -\int_0^{+\infty} e^{-x} \ln x \, dx \approx 0.5772.$$

Variance-covariance matrix

$$\begin{aligned}\text{Cov}(\varepsilon_{in}, \varepsilon_{jn}) &= \text{E}[\varepsilon_{in}\varepsilon_{jn}] - \text{E}[\varepsilon_{in}]\,\text{E}[\varepsilon_{jn}] \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \xi_i \xi_j \frac{\partial^2 F_{\varepsilon_n}(\xi_i, \xi_j)}{\partial \xi_i \partial \xi_j} d\xi_i d\xi_j - \gamma^2,\end{aligned}$$

where $\text{E}[\varepsilon_{in}] = \gamma$, $F_{\varepsilon_n}(\xi_i, \xi_j) = F_{\varepsilon_n}(\ldots, +\infty, \xi_i, +\infty, \ldots, +\infty, \xi_j, +\infty, \ldots)$ is the bivariate marginal cumulative distribution, and

$$\frac{\partial^2 F_{\varepsilon_{in},\varepsilon_{jn}}(\xi_i, \xi_j)}{\partial \xi_i \partial \xi_j} = F_{\varepsilon_{in},\varepsilon_{jn}}(\xi_i, \xi_j) e^{-\xi_i} e^{-\xi_j} (G_i^{ij} G_j^{ij} - G_{ij}^{ij})$$

where

$$G_i^{ij} = \frac{\partial G(\ldots, 0, e^{-\xi_i}, 0, \ldots, 0, e^{-\xi_j}, 0, \ldots)}{\partial y_i}$$

and

$$G_{ij}^{ij} = \frac{\partial^2 G(\ldots, 0, e^{-\xi_i}, 0, \ldots, 0, e^{-\xi_j}, 0, \ldots)}{\partial y_i \partial y_j}.$$

Logit model

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$, $\mu > 0$

1. $G(\alpha y) = \sum_{i=1}^{J} (\alpha y_i)^{\mu} = \alpha^{\mu} \sum_{i=1}^{J} y_i^{\mu} = \alpha^{\mu} G(y)$.
2. $\lim_{y_i \to +\infty} G(y) = +\infty$, $i = 1, \ldots, J$.
3. $\dfrac{\partial G}{\partial y_i} = \mu y_i^{\mu-1}$ and $\dfrac{\partial^2 G}{\partial y_i \partial y_j} = 0$.

G complies with the theory

Logit model

Example: $G(y) = \sum_{i=1}^{J} y_i^{\mu}$, $\mu > 0$

$$\begin{aligned} F(\varepsilon_1, \ldots, \varepsilon_J) &= e^{-G(e^{-\varepsilon_1}, \ldots, e^{-\varepsilon_J})} \\ &= e^{-\sum_{i=1}^{J} e^{-\mu \varepsilon_i}} \\ &= \prod_{i=1}^{J} e^{-e^{-\mu \varepsilon_i}}. \end{aligned}$$

Product of i.i.d EV.

Logit model

Example: $G(e^{V_1}, \ldots, e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$, $\mu > 0$

$$P(i) = \frac{e^{V_i + \ln G_i(e^{V_1},\ldots,e^{V_J})}}{\sum_{j \in C} e^{V_j + \ln G_j(e^{V_1},\ldots,e^{V_J})}} \text{ with } G_i(x) = \mu y_i^{\mu-1}.$$

$$\begin{aligned} e^{V_i + \ln G_i(e^{V_1},\ldots,e^{V_J})} &= e^{V_i + \ln \mu + (\mu-1)\ln e^{V_i}} \\ &= e^{\ln \mu + \mu V_i}. \end{aligned}$$

$$P(i) = \frac{e^{\ln \mu + \mu V_i}}{\sum_{j \in C} e^{\ln \mu + \mu V_j}} = \frac{e^{\mu V_i}}{\sum_{j \in C} e^{\mu V_j}}.$$

Logit model

Example: $G(e^{V_1}, \ldots, e^{V_J}) = \sum_{i=1}^{J} e^{\mu V_i}$, $\mu > 0$

$$\begin{aligned} \mathrm{E}[\max_{j \in \mathcal{C}_n} U_{jn}] &= \tfrac{1}{\mu}\left(\ln G(e^{V_1}, \ldots, e^{V_J}) + \gamma\right) \\ &= \tfrac{1}{\mu} \ln \sum_{i=1}^{J} e^{\mu V_i} + \frac{\gamma}{\mu}. \end{aligned}$$

Nested logit model

Generating function

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

with $\mu > 0$, $\mu_m > 0$.

Homogeneity

$$G(\alpha y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} (\alpha y_i)^{\mu_m} \right)^{\frac{\mu}{\mu_m}} = \alpha^{\mu} \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

Nested logit model

Generating function

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

with $\mu > 0$, $\mu_m > 0$.

Limit property

$$\lim_{y_i \to +\infty} G(y) = +\infty, i = 1, \ldots, J$$

Nested logit model

Generating function

$$G(y) = \sum_{m=1}^{M} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}}$$

with $\mu > 0$, $\mu_m > 0$.

Strong alternating sign property

$$\frac{\partial G}{\partial y_i} = \frac{\mu}{\cancel{\mu_m}} \cancel{\mu_m} y_i^{\mu_m - 1} \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m} - 1} \geqslant 0.$$

$$\frac{\partial^2 G}{\partial y_i y_j} = \mu \mu_m y_i^{\mu_m - 1} y_j^{\mu_m - 1} \left(\frac{\mu}{\mu_m} - 1 \right) \left(\sum_{i=1}^{J_m} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m} - 2} \leqslant 0, \text{if} \mu \leqslant \mu_m.$$

MEV models

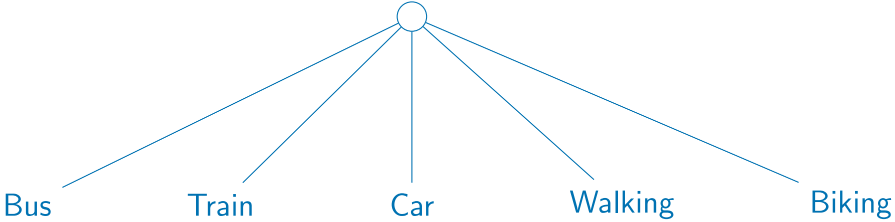
- Marginal distributions: extreme value.
- Variance-covariance matrix. (Not in closed form.)
- Normalization: $\mu = 1$.
- Expected Maximum Utility:

$$E[\max_{j \in \mathcal{C}_n} U_{jn}] = \frac{1}{\mu}(\ln G(e^{V_{1n}}, \ldots, e^{V_{J_n n}}) + \gamma)$$

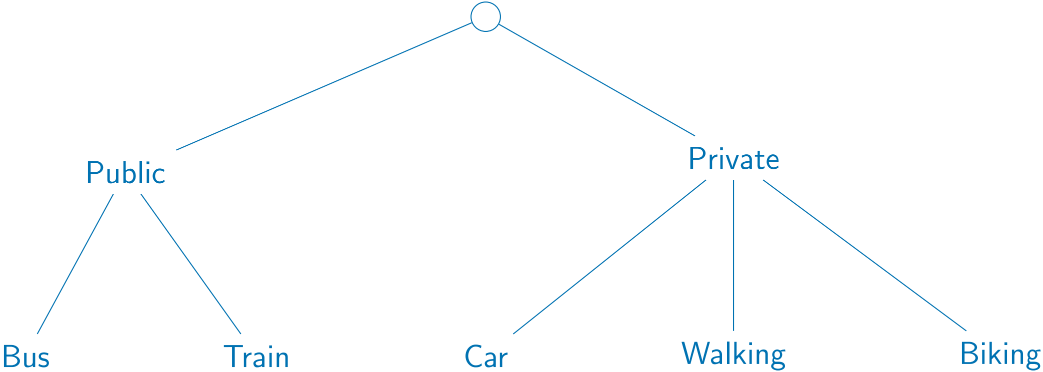
- Logit and nested logit are members of the MEV family.

Outline

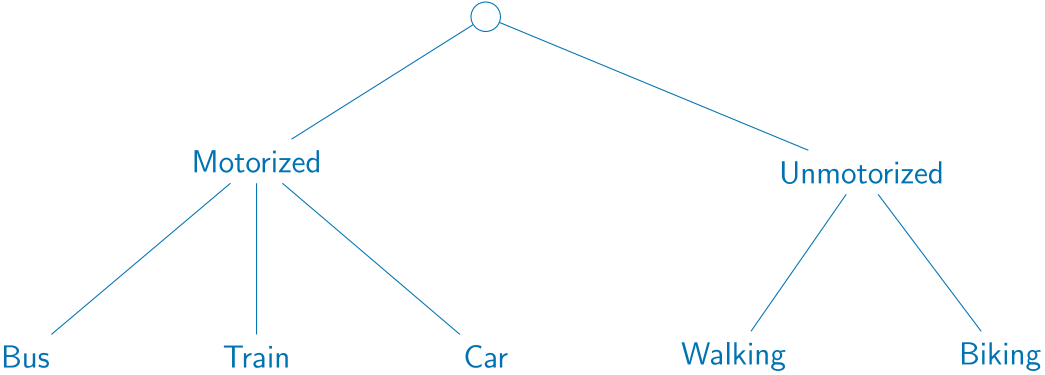
A transportation mode choice example



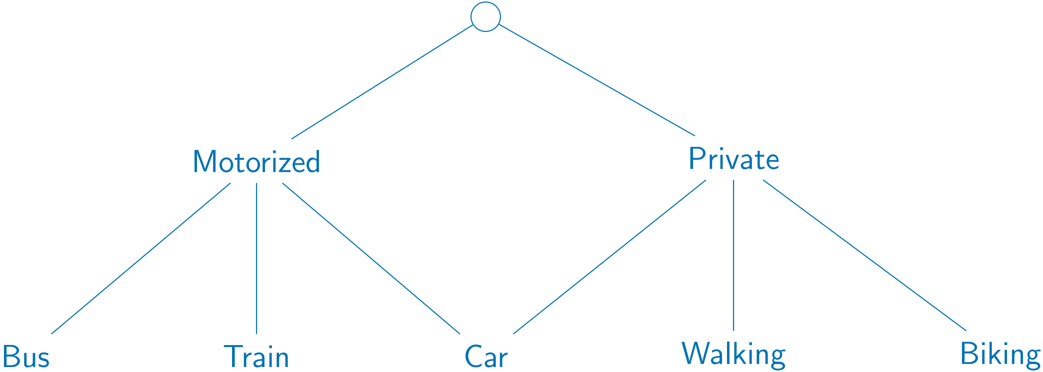
Nested logit model



Nested logit model



Cross-nested logit model



Cross-nested logit model

- Define groups of alternatives. They may overlap.
- Each of them is associated with a scale parameter μ_m.
- Like nested logit, we have $0 < \mu/\mu_m \leqslant 1$.
- For each alternative i and each nest m, there is a degree of membership $0 \leqslant \alpha_{im} \leqslant 1$.
- Each alternative must belong to at least one nest:

$$\forall j, \exists m \text{ s.t. } \alpha_{jm} > 0.$$

- Normalization:

$$\sum_m \alpha_{jm} = 1, \; \forall j.$$

MEV model

The multivariate extreme value model:

$$P_n(i) = \frac{e^{V_{in} + \ln G_i\left(e^V\right)}}{\sum_j e^{V_{jn} + \ln G_j\left(e^V\right)}}.$$

where $G_i = \partial G / \partial y_i$, and G verifies

- (i) the limit property: $G(y_{1n}, \ldots, +\infty, \ldots, y_{J_n n}) = +\infty$.
- (ii) the strong alternating sign property: for any set of $\widehat{J}_n$ distinct indices $i_1, \ldots, i_{\widehat{J}_n}$,

$$(-1)^{\widehat{J}_n - 1} G_{i_1, \ldots, i_{\widehat{J}_n}} \geqslant 0.$$

- (iii) the homogeneity property:

$$G(\alpha y) = \alpha^{\mu} G(y),\ \forall \alpha > 0 \text{ and } y \in \mathbb{R}_+^{J_n}.$$

Cross-nested logit model

Probability generating function

$$G(y_1,\ldots,y_J) = \sum_{m=1}^{M} \left(\sum_{j} (\alpha_{jm}{}^{1/\mu} y_j)^{\mu_m} \right)^{\frac{\mu}{\mu_m}},$$

with $\frac{\mu}{\mu_m} \leqslant 1$, $\alpha_{jm} \geqslant 0$, and $\forall j, \exists m$ s.t. $\alpha_{jm} > 0$.

Cross-nested logit model

Choice model

$$P(i|\mathcal{C}) = \sum_{m=1}^{M} \frac{\left(\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}\right)^{\frac{\mu}{\mu_m}}}{\sum_{\ell=1}^{M}\left(\sum_{j\in\mathcal{C}} \alpha_{j\ell}^{\mu_\ell/\mu} e^{\mu_\ell V_j}\right)^{\frac{\mu}{\mu_\ell}}} \frac{\alpha_{im}^{\mu_m/\mu} e^{\mu_m V_i}}{\sum_{j\in\mathcal{C}} \alpha_{jm}^{\mu_m/\mu} e^{\mu_m V_j}},$$

which can nicely be interpreted as

$$P(i|\mathcal{C}) = \sum_{m} P(m|\mathcal{C}) P(i|m).$$

Nested logit model

Special case

$$\alpha_{im} = \begin{cases} 1 & \text{if } i \in \mathcal{C}_m, \\ 0 & \text{otherwise.} \end{cases}$$

Outline

Case study: choice of airline itinerary

Survey

- Conducted by the Boeing Company (fall 2004).
- Sample of the customers of an Internet airline booking service.

Booking

The Internet service

- takes a specific user request for travel in a city pair,
- interrogates the web sites of airlines that provide service in that market,
- returns to the user a compiled list of available itineraries.

Case study: choice of airline itinerary

Questionnaire

- Random selection of customers for the survey.
- Three alternatives based on the origin-destination market request that the respondent entered into the itinerary search engine:
  1. a non stop flight,
  2. a flight with 1 stop on the same airline,
  3. a flight with 1 stop and a change of airline.

Case study: choice of airline itinerary

Demographic data

- Age,
- gender,
- income,
- occupation,
- education.

Context data

- Desired departure time,
- trip purpose,
- who is paying for the trip,
- the number of passengers traveling together.

Case study: choice of airline itinerary

Pick Your Preferred Flight

Three flight options are described for your trip from Chicago to San Diego . These are options that might be available on this route or might be new options actively being considered for this route as well as replacing some options that are offered now. The options differ from each other in one or more of the features described on the left.

Please evaluate these options, assuming that eveything about the options is the same except these particular features. Indicate your choices at the bottom of the appropriate column and press the Continue button.

FEATURES	Non-Stop (Option 1)	1 Stop (Option 2)	1 Stop (Option 3)
Departure time (local)	6:00 PM	4:30 PM	6:00 PM
Arrival time (local)	8:14 PM	8:44 PM	9:44 PM
Total time in air	4 hr 14 min	4 hr 44 min	4 hr 44 min
Total trip time	4 hr 14 min	6 hr 14 min	5 hr 44 min
Legroom	typical legroom	2-in more of legroom	4-in more of legroom
Airline [Airplane]	Depart Chicago Continental Airlines [B737] to San Diego	Depart Chicago Southwest Airlines [A320], connecting with Southwest Airlines [MD80] to San Diego	Depart Chicago Northwest Airlines [MD80], connecting with American Airlines [DC9] to San Diego
Fare	$565	$485	$620

1. Which is MOST attractive? ○ Option 1 ○ Option 2 ○ Option 3

2. Which is LEAST attractive? ○ Option 1 ○ Option 2 ○ Option 3

3. If these were the ONLY three options available, I would NOT make this trip by air. ○ Yes ○ No

Case study: choice of airline itinerary

Sample

- origin-destination city pairs in the USA,
- 3609 respondents,
- 1 choice each,
- we consider only data corresponding to leisure trips.

Logit model

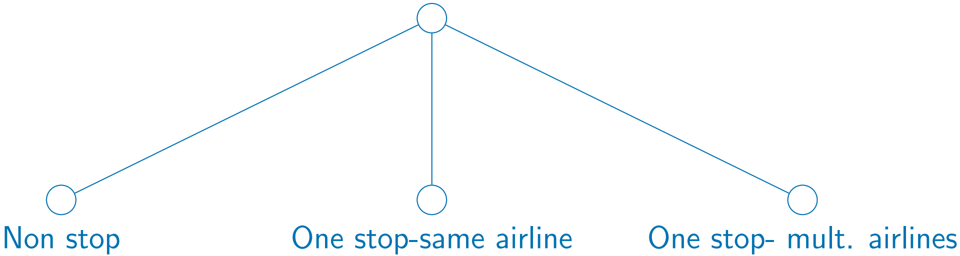
Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-1.05	0.203	-5.16	2.49e-07
2	One stop–multiple airlines dummy	-1.23	0.205	-6.0	1.93e-09
3	Round trip fare ($)	-0.0178	0.00151	-11.8	0.0
4	Elapsed time (hours) (0–2 hours)	-0.867	0.228	-3.8	0.000145
5	Elapsed time (hours) (2–8 hours)	-0.262	0.0792	-3.3	0.000953
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.83	0.785	-2.33	0.0197
7	Leg room (inches), if male (non stop)	0.234	0.0588	3.97	7.09e-05
8	Leg room (inches), if female (non stop)	0.344	0.0576	5.97	2.35e-09
9	Leg room (inches), if male (one stop)	0.205	0.0528	3.89	9.96e-05
10	Leg room (inches), if female (one stop)	0.213	0.05	4.26	2.06e-05
11	Being early (hours)	-0.149	0.019	-7.85	4.22e-15
12	Being late (hours)	-0.0985	0.0168	-5.87	4.26e-09
13	Round trip fare / income ($/$1000)	-0.244	0.082	-2.97	0.00293

Summary statistics

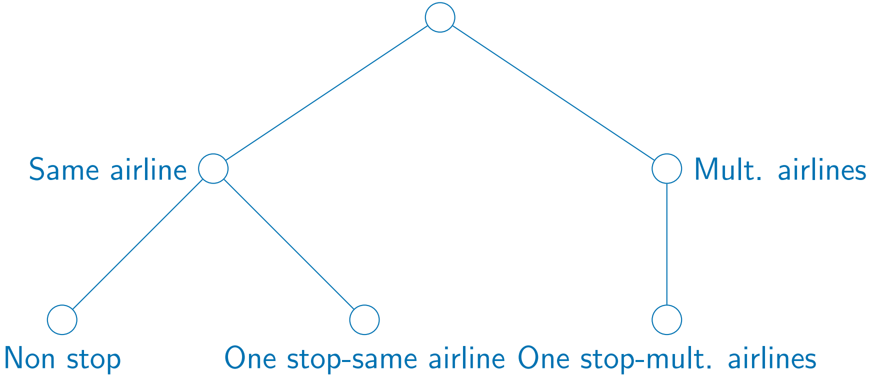
Number of observations = 2544

$$\mathcal{L}(0) = -2794.870$$
$$\mathcal{L}(\hat{\beta}) = -1637.058$$
$$-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 2315.624$$
$$\rho^2 = 0.414$$
$$\bar{\rho}^2 = 0.41$$
$$\text{AIC} = 3300.115$$
$$\text{BIC} = 3376.055$$

Logit model



Nested logit model



Nested logit model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-1.11	0.222	-5.01	5.47e-07
2	One stop–multiple airlines dummy	-1.22	0.215	-5.67	1.44e-08
3	Round trip fare (\$)	-0.0185	0.00173	-10.7	0.0
4	Elapsed time (hours) (0–2 hours)	-0.892	0.24	-3.71	0.000205
5	Elapsed time (hours) (2–8 hours)	-0.273	0.083	-3.29	0.000999
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.94	0.821	-2.36	0.0183
7	Leg room (inches), if male (non stop)	0.243	0.0624	3.89	0.000101
8	Leg room (inches), if female (non stop)	0.36	0.062	5.81	6.35e-09
9	Leg room (inches), if male (one stop)	0.212	0.055	3.85	0.000117
10	Leg room (inches), if female (one stop)	0.222	0.0519	4.27	1.92e-05
11	Being early (hours)	-0.156	0.0212	-7.36	1.9e-13
12	Being late (hours)	-0.103	0.018	-5.71	1.12e-08
13	Round trip fare / income (\$/\$1000)	-0.253	0.0857	-2.95	0.00316
14	$\mu_{\text{same airline}}$	0.929	0.0763	12.2	0.0

Summary statistics

Number of observations = 2544

$$\mathcal{L}(0) = -2794.870$$

$$\mathcal{L}(\hat{\beta}) = -1636.573$$

$$-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 2316.593$$

$$\rho^2 = 0.414$$

$$\bar{\rho}^2 = 0.409$$

$$\text{AIC} = 3301.146$$

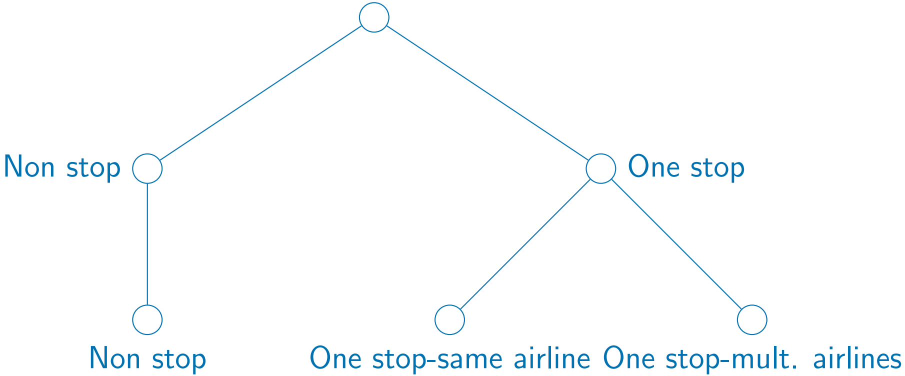
$$\text{BIC} = 3382.927$$

Nested logit model

Model is rejected

- $\mu_{\text{same airline}}$ cannot be lesser than 1.
- AIC, BIC and $\bar{\rho}^2$ show that the fit has not significantly improved.

Nested logit model



Nested logit model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-0.724	0.168	-4.32	1.58e-05
2	One stop–multiple airlines dummy	-0.865	0.169	-5.1	3.33e-07
3	Round trip fare ($)	-0.0151	0.00148	-10.2	0.0
4	Elapsed time (hours) (0–2 hours)	-0.838	0.217	-3.86	0.000112
5	Elapsed time (hours) (2–8 hours)	-0.216	0.059	-3.66	0.000249
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.71	0.722	-2.38	0.0175
7	Leg room (inches), if male (non stop)	0.227	0.0534	4.25	2.18e-05
8	Leg room (inches), if female (non stop)	0.315	0.0527	5.96	2.45e-09
9	Leg room (inches), if male (one stop)	0.159	0.0402	3.96	7.64e-05
10	Leg room (inches), if female (one stop)	0.15	0.0394	3.8	0.000143
11	Being early (hours)	-0.127	0.0161	-7.92	2.22e-15
12	Being late (hours)	-0.0734	0.0142	-5.17	2.32e-07
13	Round trip fare / income ($/$1000)	-0.212	0.0753	-2.82	0.00485
14	$\mu_{\text{one stop}}$	1.82	0.199	9.15	0.0

Summary statistics

Number of observations = 2544

$$\mathcal{L}(0) = -2794.870$$

$$\mathcal{L}(\hat{\beta}) = -1616.221$$

$$-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 2357.298$$

$$\rho^2 = 0.422$$

$$\bar{\rho}^2 = 0.417$$

$$\text{AIC} = 3260.441$$

$$\text{BIC} = 3342.222$$

Nested logit model

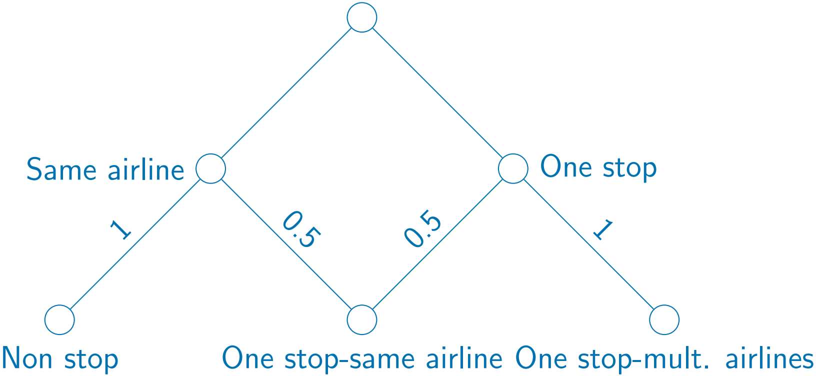
Model is accepted

- $\mu_{\text{one stop}}$ is larger than 1.
- Testing the null hypothesis that the true value is 1, the *t*-test is

$$\frac{1.82 - 1}{0.199} = 4.07.$$

  Therefore, the null hypothesis can be rejected.
- AIC, BIC and $\bar{\rho}^2$ show that the fit has significantly improved.

Cross-nested logit model



Cross-nested logit model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	*t*-stat	*p*-value
1	One stop–same airline dummy	-0.715	0.189	-3.79	0.00015
2	One stop–multiple airlines dummy	-1.03	0.162	-6.33	2.47e-10
3	Round trip fare ($)	-0.0153	0.00169	-9.03	0.0
4	Elapsed time (hours) (0–2 hours)	-0.808	0.213	-3.79	0.00015
5	Elapsed time (hours) (2–8 hours)	-0.207	0.0615	-3.36	0.00079
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.51	0.698	-2.16	0.031
7	Leg room (inches), if male (non stop)	0.222	0.0529	4.19	2.77e-05
8	Leg room (inches), if female (non stop)	0.306	0.0543	5.64	1.65e-08
9	Leg room (inches), if male (one stop)	0.168	0.0416	4.04	5.43e-05
10	Leg room (inches), if female (one stop)	0.159	0.0411	3.86	0.000112
11	Being early (hours)	-0.128	0.0176	-7.3	2.94e-13
12	Being late (hours)	-0.075	0.0154	-4.88	1.08e-06
13	Round trip fare / income ($/$1000)	-0.194	0.0772	-2.51	0.012
14	$\mu_{\text{same airline}}$	1.1	0.118	9.32	0.0
15	$\mu_{\text{one stop}}$	2.32	0.386	6.02	1.75e-09

Summary statistics

Number of observations = 2544

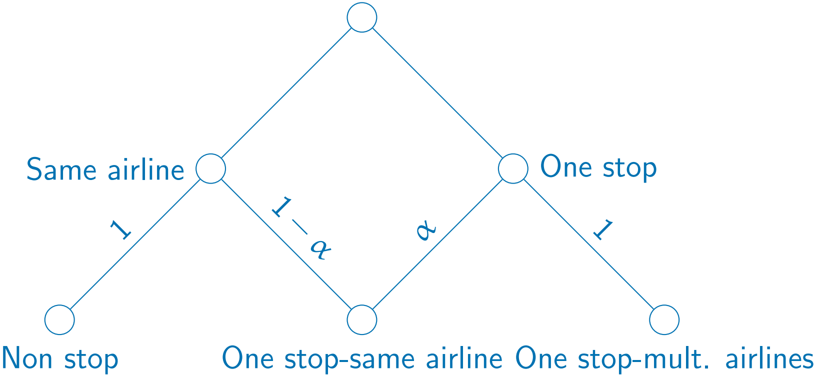
$$\mathcal{L}(0) = -2794.870$$
$$\mathcal{L}(\hat{\beta}) = -1618.303$$
$$-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 2353.134$$
$$\rho^2 = 0.421$$
$$\bar{\rho}^2 = 0.416$$
$$\text{AIC} = 3266.606$$
$$\text{BIC} = 3354.228$$

Cross-nested logit model

Model is rejected

- Fit is not better than nested logit, with one more parameter.

Cross-nested logit model



Cross-nested logit model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-0.782	0.175	-4.48	7.53e-06
2	One stop–multiple airlines dummy	-0.867	0.171	-5.06	4.17e-07
3	Round trip fare ($)	-0.0156	0.00159	-9.79	0.0
4	Elapsed time (hours) (0–2 hours)	-0.847	0.224	-3.78	0.000156
5	Elapsed time (hours) (2–8 hours)	-0.22	0.0601	-3.66	0.000248
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.73	0.778	-2.22	0.0264
7	Leg room (inches), if male (non stop)	0.232	0.0548	4.23	2.34e-05
8	Leg room (inches), if female (non stop)	0.324	0.055	5.9	3.74e-09
9	Leg room (inches), if male (one stop)	0.165	0.0404	4.08	4.49e-05
10	Leg room (inches), if female (one stop)	0.156	0.0403	3.87	0.000109
11	Being early (hours)	-0.132	0.0168	-7.83	4.88e-15
12	Being late (hours)	-0.0733	0.0148	-4.94	7.65e-07
13	Round trip fare / income ($/$1000)	-0.213	0.08	-2.66	0.00788
14	$\mu_{\text{same airline}}$	0.788	0.207	3.8	0.000144
15	$\mu_{\text{one stop}}$	2.02	0.291	6.95	3.72e-12
16	$\alpha_{\text{one stop}}$	0.932	0.103	9.01	0.0

Summary statistics

Number of observations = 2544

$$\begin{aligned}
\mathcal{L}(0) &= -2794.870 \\
\mathcal{L}(\hat{\beta}) &= -1613.924 \\
-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] &= 2361.892 \\
\rho^2 &= 0.423 \\
\bar{\rho}^2 &= 0.417 \\
\text{AIC} &= 3259.848 \\
\text{BIC} &= 3353.312
\end{aligned}$$

Cross-nested logit model

Model is rejected

- $\mu_{\text{same airline}}$ cannot be lesser than 1.
- We constrain $\mu_{\text{same airline}}$ to 1.

Cross-nested logit model

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	One stop–same airline dummy	-0.716	0.164	-4.35	1.34e-05
2	One stop–multiple airlines dummy	-0.884	0.167	-5.3	1.14e-07
3	Round trip fare ($)	-0.0151	0.00149	-10.1	0.0
4	Elapsed time (hours) (0–2 hours)	-0.831	0.217	-3.83	0.000127
5	Elapsed time (hours) (2–8 hours)	-0.211	0.0576	-3.65	0.000259
6	Elapsed time (hours) ($\geqslant$ 8 hours)	-1.62	0.738	-2.2	0.028
7	Leg room (inches), if male (non stop)	0.227	0.053	4.28	1.86e-05
8	Leg room (inches), if female (non stop)	0.312	0.0524	5.95	2.61e-09
9	Leg room (inches), if male (one stop)	0.16	0.0389	4.11	3.93e-05
10	Leg room (inches), if female (one stop)	0.148	0.0385	3.86	0.000114
11	Being early (hours)	-0.127	0.0158	-8.03	1.11e-15
12	Being late (hours)	-0.0711	0.0142	-5.0	5.73e-07
13	Round trip fare / income ($/$1000)	-0.202	0.0758	-2.66	0.00775
14	$\mu_{\text{one stop}}$	2.14	0.316	6.77	1.28e-11
15	$\alpha_{\text{one stop}}$	0.808	0.0919	8.8	0.0

Summary statistics

Number of observations = 2544

$$\mathcal{L}(0) = -2794.870$$
$$\mathcal{L}(\hat{\beta}) = -1614.425$$
$$-2[\mathcal{L}(0) - \mathcal{L}(\hat{\beta})] = 2360.89$$
$$\rho^2 = 0.422$$
$$\bar{\rho}^2 = 0.417$$
$$\text{AIC} = 3258.849$$
$$\text{BIC} = 3346.471$$

Estimated models

Model	K	$\mathcal{L}(\widehat{\beta})$	AIC
Logit	13	-1637.058	3300.115
Nested logit (same airline)		Rejected	
Nested logit (one stop)	14	-1616.221	3260.441
Cross-nested (fixed α)		Rejected	
Cross-nested		Rejected	
Cross-nested (fixed $\mu_{\text{same airline}}$)	15	-1614.425	3258.849

Correlation matrix

Cross-nested logit model

	Non stop	One stop-same airline	One stop-mult. airlines
Non stop	1	0	0
One stop-same airline	0	1	0.692
One stop-mult. airlines	0	0.692	1

Nested logit model

	Non stop	One stop-same airline	One stop-mult. airlines
Non stop	1	0	0
One stop-same airline	0	1	0.699
One stop-mult. airlines	0	0.699	1

Correlation structure

Notes

- In this case, the cross-nested logit model has a block diagonal structure, as the nested logit model.
- But it does not mean it is a nested logit model.
- Contrarily to probit models, MEV models are not characterized by the structure of their correlation matrix.

Summary

- Limitations of logit.
- Nested logit model.
- MEV models.
- Cross-nested logit model.

Relax the i.i.d. assumption

i.i.d. assumption

- ✓ Same η for all alternatives i: relaxed.
- ✓ Same η for all observations n: relaxed.
- ▶ Same μ for all alternatives i: relaxed in the lecture on mixtures.
- ✓ Same μ for all observations n: relaxed.
- ▶ Independence across alternatives i: relaxed in this lecture.
- ▶ Independence across observations n: relaxed in the lecture on panel data.

Outline

MEV: the choice model

CDF of the error terms

$$F_{\varepsilon_n}(\varepsilon_{1n}, \ldots, \varepsilon_{J_n n}) = e^{-G(e^{-\varepsilon_{1n}}, \ldots, e^{-\varepsilon_{J_n n}})},$$

Choice model: $P_n(i) =$

$$\int_{\varepsilon=-\infty}^{+\infty} \frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) d\varepsilon.$$

$$\begin{aligned}
&\frac{\partial F_{\varepsilon_{1n}, \varepsilon_{2n}, \ldots, \varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots) \\
&= e^{-\varepsilon} G_i(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots) \\
&\exp\left(-G(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots)\right)
\end{aligned}$$

MEV: the choice model

G is μ-homogeneous

so that $G_i = \partial G / \partial y_i$ is $(\mu - 1)$-homogeneous.

$$
\begin{aligned}
& e^{-\varepsilon} G_i(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots) \\
& \exp\left(-G(\ldots, e^{-V_{in}+V_{(i-1)n}-\varepsilon}, e^{-\varepsilon}, e^{-V_{in}+V_{(i+1)n}-\varepsilon}, \ldots)\right) \\
& = e^{-\varepsilon} e^{-(\mu-1)\varepsilon} e^{-(\mu-1)V_{in}} G_i(\ldots, e^{V_{(i-1)n}}, e^{V_i n}, e^{V_{(i+1)n}}, \ldots) \\
& \qquad \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(\ldots, e^{V_{(i-1)n}}, e^{V_{in}}, e^{V_{(i+1)n}}, \ldots)\right) \, .
\end{aligned}
$$

MEV: choice model

We now denote

$$e^{V} = \left(\ldots, e^{V_{(i-1)n}}, e^{V_{in}}, e^{V_{(i+1)n}}, \ldots\right),$$

and simplify the terms to obtain

$$\frac{\partial F_{\varepsilon_{1n},\varepsilon_{2n},\ldots,\varepsilon_{J_n n}}}{\partial \varepsilon_i}(\ldots, V_{in} - V_{(i-1)n} + \varepsilon, \varepsilon, V_{in} - V_{(i+1)n} + \varepsilon, \ldots)$$

$$= e^{-\mu\varepsilon} e^{-\mu V_{in}} e^{V_{in}} G_i(e^{V}) \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^{V})\right).$$

Therefore

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^{V}) \int_{\varepsilon=-\infty}^{+\infty} e^{-\mu\varepsilon} \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^{V})\right) d\varepsilon.$$

MEV: choice model

Choice probability

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \int_{\varepsilon=-\infty}^{+\infty} e^{-\mu\varepsilon} \exp\left(-e^{-\mu\varepsilon} e^{-\mu V_{in}} G(e^V)\right) d\varepsilon.$$

Define $t = -\exp(-\mu\varepsilon)$, so that $dt = \mu \exp(-\mu\varepsilon) d\varepsilon$:

$$P_n(i) = e^{-\mu V_{in}} e^{V_{in}} G_i(e^V) \frac{1}{\mu} \int_{t=-\infty}^{0} \exp\left(t e^{-\mu V_{in}} G(e^V)\right) dt,$$

which simplifies to

$$P_n(i) = \frac{e^{V_{in}} G_i(e^V)}{\mu G(e^V)}.$$

MEV: choice model

Choice probability

$$P_n(i) = \frac{e^{V_{in}} G_i(e^V)}{\mu G(e^V)}.$$

From Euler's theorem:

$$P_n(i) = \frac{e^{V_{in}} G_i\left(e^V\right)}{\sum_j e^{V_{jn}} G_j\left(e^V\right)}.$$

Logit-like form:

$$P_n(i) = \frac{e^{V_{in} + \ln G_i\left(e^V\right)}}{\sum_j e^{V_{jn} + \ln G_j\left(e^V\right)}}.$$