Choice Probability Generating Functions

DCA Workshop 2014, EPFL

Anna Fernández Antolín
Michel Bierlaire
Mogens Fosgerau
Daniel McFadden

Transport and Mobility Laboratory
School of Architecture, Civil and Environmental Engineering
Ecole Polytechnique Fédérale de Lausanne

June 20^{th}, 2014

Outline

Motivation

To fully characterize finite-mean RUM

It can be done analogously to how MEV models are introduced in McFadden (1978):

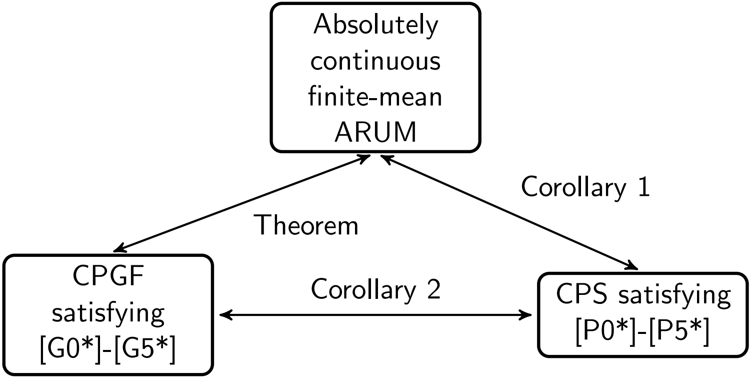
- Define a function G satisfying some conditions
- Then $U = \log G(\mathrm{e}^{V_1}, ..., \mathrm{e}^{V_J}) + \gamma$ satisfies

$$U = E[\max_{j \in \mathcal{C}} U_j]$$

where $\gamma = 0.57721566$ is Euler's constant.

by relaxing some of the conditions of G.

Conclusions



Definitions

Concepts needed:

- Random utility model (RUM)
- Additive random utility model (ARUM)
- Transformed additive random utility model (TARUM)
- Choice probability generating function (CPGF)
- Choice probability system (CPS)

Random utility model

A random utility model (RUM) is a vector $U = (U_1, ..., U_J) \in \mathbb{R}^J$ with CDF $R(u_1, ..., u_J)$. It induces an observable choice probability:

$$P_{\mathcal{C}}(j) = \Pr(U_j > U_k \text{ for } j \neq k \in \mathcal{C}) = \int_{-\infty}^{\infty} \nabla_j R(u, ...u)\, du$$

- Given a continuously differentiable increasing transformation $r : \mathbb{R} \to \mathbb{R}$, the image of a RUM is a RUM with the same associated choice probability.
- This defines a class.
- A representative can be chosen such that $E(U_j)$ is finite $\forall j$

Additive random utility model

Given:

- A finite-mean RUM with CDF $R(u_1, ..., u_J)$ and choice probability $P^0_{\mathcal{C}}(j)$
- $\mathbf{m} = (m_1, ..., m_J) \in \mathbb{R}^J$ a location vector

Then the family,

$$U = (U_1, ..., U_J) = \mathbf{m} + U^0$$

is an additive random utility model (ARUM). Each element is a finite-mean RUM with CDF $R(u_1 - m_1, ..., u_J - m_J)$ and choice probability

$$P_{\mathcal{C}}(j|\mathbf{m}) = \Pr(U_j \geq U_k \text{ for } k \in \mathcal{C}) = \Pr(U^0_j + m_j \geq U^0_k + m_k \text{ for } k \in \mathcal{C})$$

Transformed additive random utility model (TARUM)

Given

- r_j a continuously differentiable increasing transformation with a continuously differentiable inverse,
- $\varepsilon = (\varepsilon_1, ..., \varepsilon_J)$ with an absolutely continuous CDF $F(\varepsilon)$, and
- $\mathbf{m} = (m_1, ..., m_J)$ a location vector.

A finite-mean family of RUM defined implicitly by $\varepsilon_j = r_j(U_j - m_j)$ is a finite-mean transformed additive random utility model (TARUM).

RUM, ARUM, TARUM

Every RUM has a representation that can be embedded in an ARUM with finite mean, or an observationally equivalent TARUM.

Choice probability system (CPS)

A complete choice probability system is a familiy of non-negative functions $P_{\mathcal{C}}(j|\mathbf{m})$ for $j = 1, ..., J$ that sum up to one.

A partial choice probability system is a family of non-negative functions $P_{\mathcal{C}}(j|\mathbf{m})$ defined for $j = 1.,,, k < J$ whose sum does not exceed 1.

Choice probability generating function

A choice probability generating function (CPGF) is a function $G : [0, +\infty]^J \to [0, +\infty]$ with the following properties:

- G1 Weak alterning signs property (ASP)
- G2 Homogeneity
- G3 Boundary
- G4 Integrability

The gradient

$$P_{\mathcal{C}}(j|\mathbf{m}) = \frac{\partial \ln G(\mathrm{e}^{\mathbf{m}})}{\partial m_j}$$

is the choice probability generated by the CPGF G.

Properties [G1]-[G4] are equivalent to another set of properties [G0*]-[G5*]

Theorem

Given

- A finite-mean TARUM $\varepsilon_j = r_j(U_j - m_j)$ with $\varepsilon = (\varepsilon_1, ..., \varepsilon_J)$ with CDF $F(\varepsilon)$
- An observationally equivalent ARUM $U_j = m_j + \zeta_j$ with $\zeta = (\zeta_1, ..., \zeta_J)$ with CDF $R(\zeta) \equiv F(r(U - \mathbf{m}))$

Then, an associated CPGF exists and is defined by:

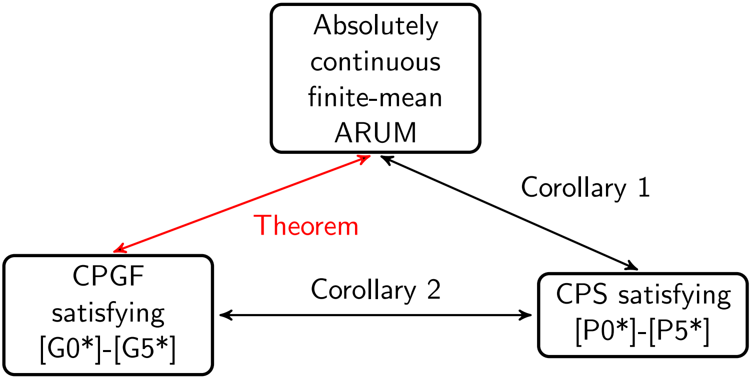
$$\ln G(\mathrm{e}^{\mathbf{m}}) \equiv E[\max_{j \in \mathcal{C}} U_j] \equiv \int_0^{+\infty} [1 - R(u - \mathbf{m})]\, du - \int_{-\infty}^{0} R(u - \mathbf{m})\, du$$

The choice probability implied by the TARUM satisfies:

$$P_{\mathcal{C}}(j|\mathbf{m}) = \Pr(U_j > U_k \text{ for } k \neq j) = \partial \ln G(\mathrm{e}^{\mathbf{m}}) / \partial m_j$$

The reciprocal is also true.

Theorem


Absolutely continuous finite-mean ARUM
Theorem
Corollary 1
CPGF satisfying [G0*]-[G5*]
Corollary 2
CPS satisfying [P0*]-[P5*]

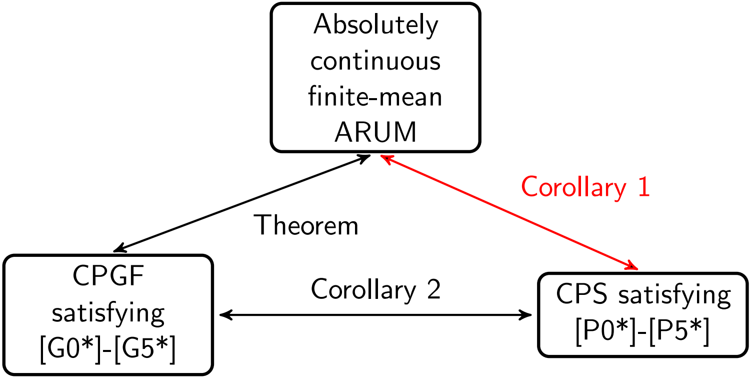
Corollary 1

A complete or partial CPS is ARUM-consistent if and only if it satisfies:

- P1 Alterning signs property (ASP)
- P2 Homogeneity
- P3 Boundary
- P4 Integrability

These can be expressed as another set of properties [P0*]-[P5*]

Corollary 1


Absolutely continuous finite-mean ARUM
Theorem
Corollary 1
CPGF satisfying [G0*]-[G5*]
Corollary 2
CPS satisfying [P0*]-[P5*]

Corollary 2

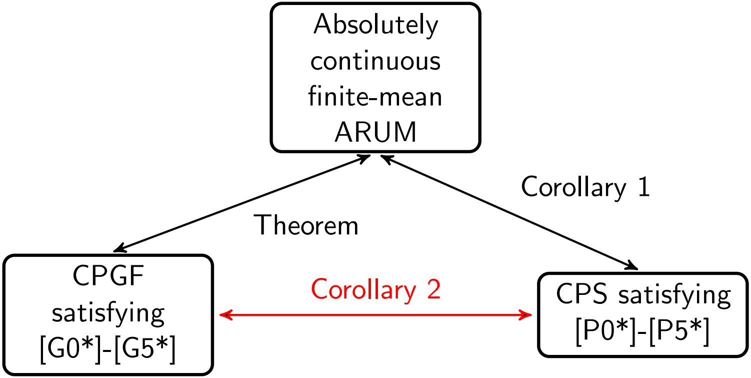
If $P_{\mathcal{C}}(1|\mathbf{m})$ is the first probability in a CPS satisfying [P0*]-[P5*] then

$$h(\mathbf{m}) = m_1 + \int_{m_1}^{+\infty} [1 - P(1|v_1, \mathbf{m}_{-1})]$$

is a CPGF satisfying [G0*]-[G5*] and whose gradient gives the partial CPS.

Notation: Given $\mathbf{m} = (m_1, ..., m_J)$, $\mathbf{m}_{-j} = (m_1, ..., m_{j-1}, m_{j+1}, ..., m_J)$ denotes a vector excluding component j.

Corollary 2


Absolutely continuous finite-mean ARUM
Corollary 1
Theorem
CPGF satisfying [G0*]-[G5*]
Corollary 2
CPS satisfying [P0*]-[P5*]

Choice probability from a single alternative is enough

A choice probability $P_{\mathcal{C}}(1|\mathbf{m})$ for a single alternative that satisfies [P0*]-[P5*] (for consistency with an ARUM) can be used to determine the CPGF from which the choice probabilities for the rest of the alternatives can be determined.

Simple example: logit model

Notation:

$$h(\mathbf{m}) = \ln G(\mathrm{e}^{\mathbf{m}})$$

Corollary 2:

$$h(\mathbf{m}) = m_1 + \int_{m_1}^{+\infty} [1 - P(1|v_1, \mathbf{m}_{-1})]$$

Multinomial logit:

$$P_{\mathcal{C}}(1|\mathbf{m}) = \frac{\mathrm{e}^{m_1}}{\sum_{j=1}^{J} \mathrm{e}^{m_j}}$$

Simple example: logit model

From one probability we can recover $h(\mathbf{m})$:

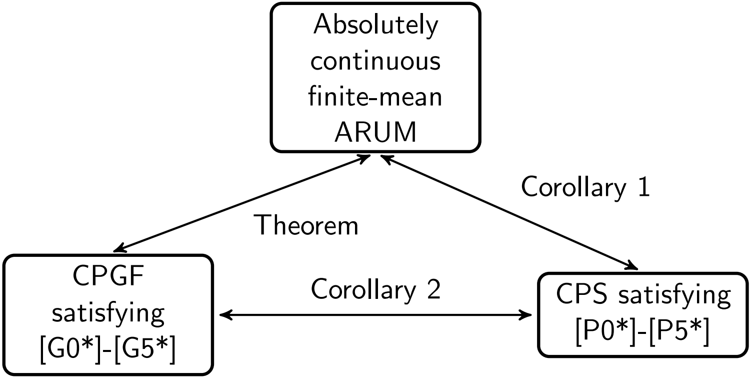
$$\begin{aligned}
h(\mathbf{m}) &= m_1 + \int_{\mathbf{m}_1}^{+\infty} 1 - P(1|v_1, m_{-1})\, dv_1 \\
&= m_1 + \int_{m_1}^{+\infty} 1 - \frac{\mathrm{e}^{v_1}}{\mathrm{e}^{v_1} + \sum_{j=2}^{J} \mathrm{e}^{m_j}}\, dv_1 \\
&= m_1 + \left[v_1 - \ln\left(\mathrm{e}^{v_1} + \sum_{j=2}^{J} \mathrm{e}^{m_j} \right) \right]_{v_1=m_1}^{v_1=+\infty} \\
&= m_1 - m_1 + \ln\left(\mathrm{e}^{m_1} + \sum_{j=2}^{J} \mathrm{e}^{m_j} \right) \\
&= \ln\left(\sum_{j=1}^{J} \mathrm{e}^{m_j} \right)
\end{aligned}$$

Simple example: logit model

And from $h(\mathbf{m})$ the rest of probabilities:

$$P_{\mathcal{C}}(i|\mathbf{m}) = \frac{\partial}{\partial m_i} h(\mathbf{m}) = \frac{\partial}{\partial m_i} \ln \left(\sum_{j=1}^{J} \mathrm{e}^{m_j} \right) = \frac{\mathrm{e}^{m_i}}{\sum_{j=1}^{J} \mathrm{e}^{m_j}}$$

Conclusions



Future work

To give an application to this theoretical work. But... how?

- To use the CPGF approach to propose new models
- To take advantage of the fact that full information on choice probabilities is contained in only one of them:
  1. Specify the choice probability for one alternative
  2. Computation of the CPGF and its gradient.
  3. This CPGF could then be used to calculate willingness to pay (WTP) for a policy change

Thanks for your attention!

Questions?

APPENDIX

Properties [G1]-[G4]

[G1] Weak ASP

In $G(y)$ satisfies the ASP, so that for any permutation σ of $(1, ..., J)$ and $k = 1, ..., J$, its mixed partial derivatives are independent of the order of differentiation and

$$\chi_{\sigma:k}(\ln y) \equiv (-1)^k \nabla_{\sigma:k} \ln G(y) \geq 0$$

[G2] Homogeneity

For each $\lambda > 0$ and $y \in [0, +\infty]^J$, $G(\lambda y) = \lambda G(y)$

Properties [G1]-[G4]

[G3] Boundary
G(0)=0, and for $j = 1, ..., J$, if $1_{(j)}$ denotes a unit vector with the jth component equal to one, then $\lim_{\lambda\to\infty} G(\lambda 1_{(j)}) \to \infty$

[G4] Integrability
Let $\mathbf{m} = (m_1, ..., m_J)$, $\mathrm{e}^{\mathbf{m}} = (\mathrm{e}^{m_1}, ..., \mathrm{e}^{m_J})$, $L = \{\mathbf{m} \in \mathbb{R}^J | \sum_{j=1}^J m_j = 0\}$.
Then the following holds:

$$\int_L \chi_{1,...,J}(\mathbf{m})\, d\mathbf{m} = J^{-1}$$

$$\int_L |m_j| \chi_{1,...,J}(\mathbf{m})\, d\mathbf{m} < +\infty$$

$$\max_{k\neq j} m_k - m_j \to \infty \implies \chi_j(\mathbf{m}) \to 0$$

Properties [G0*]-[G5*]

Let $h(\mathbf{m}) = \ln G(\mathrm{e}^{\mathbf{m}})$, then [G1]-[G4] are equivalent to [G0*]-[G5*], where:

[G0*] $h(\mathbf{m})$ is a convex function whose mixed partial derivatives exist

[G1*] $(-1)^{k-1} h_{j_1,\ldots,j_k}(\mathbf{m}) \geq 0$ for any $k = 1, \ldots, J$ and distinct indices $j_1, \ldots, j_k$

[G2*] $h(\mathbf{m} - \gamma) = h(\mathbf{m}) - \gamma$ for any scalar γ

Properties [G0*]-[G5*]

[G3*] $\lim_{m_j \to -\infty} \frac{h(\mathbf{m})}{m_j} = 0$,
$\lim_{m_j \to -\infty} m_j h_j(\mathbf{m}) = 0$,
$\lim_{m_j \to +\infty} \frac{h(\mathbf{m})}{m_j} = 1$,
$\lim_{m_j \to +\infty} m_j [1 - h_j(\mathbf{m})] = 0$,
$\int_{v=-\infty}^{0} h_j(m_j, \mathbf{m_{-j}})\, dm_j < +\infty$, and
$\int_{m_j=0}^{+\infty} [1 - h_j(m_j, \mathbf{m}_{-j})]\, dm_j < +\infty$.

Properties [G0*]-[G5*]

[G4*] $(-1)^{J-1} \int_{\mathbf{m_{-j}}=-\infty}^{+\infty} \frac{\partial^J h(\mathbf{m})}{\partial m_1 \ldots \partial m_J} \, d\mathbf{m_{-j}} = 1$ for each m_j

[G5*]
$- \int_{m_k=-\infty}^{+\infty} |m_k| h_{jk}(-\infty, \ldots, -\infty, m_j, -\infty, \ldots, -\infty, m_k, -\infty, \ldots, -\infty) \, dm_k < +\infty$ for each m_j and $k \neq j$

Properties [P0*]-[P5*]

[P0*]

- Mixed partial derivatives of $P_{\mathcal{C}}(j|\mathbf{m})$ with respect to $\mathbf{m}$ exist
- $P_{\mathcal{C}}(j|\mathbf{m})$ is monotone, non-decreasing in m_j and monotone non-increasing in $\mathbf{m}_{-j}$
- $\partial P_{\mathcal{C}}(j|\mathbf{m})/\partial m_i = \partial P_{\mathcal{C}}(i|\mathbf{m})/\partial m_j$ for $1 \leq i\,,\, j \leq k$

[P1*] $(-1)^n \partial^n P_{\mathcal{C}}(j_0|\mathbf{m})/\partial m_{j_1} ... \partial m_{j_n} \geq 0$ for any n and distinct indices $j_0, ..., j_k$, with $j_0 \leq k$

[P2*] $P_{\mathcal{C}}(j|\mathbf{m} - \gamma) \equiv P_{\mathcal{C}}(j|\mathbf{m})$ for any scalar γ

Properties [P0*]-[P5*]

[P3*] $\lim_{m_j \to -\infty} m_j P_{\mathcal{C}}(j|\mathbf{m}) = 0$,
$\lim_{m_j \to +\infty} m_j [1 - P_{\mathcal{C}}(j|\mathbf{m})] = 0$,
$\int_{-\infty}^{0} P_{\mathcal{C}}(j|m_j, \mathbf{m}_{-j})\, dm_j < +\infty$, and
$\int_{0}^{+\infty} [1 - P_{\mathcal{C}}(j|m_j, \mathbf{m}_{-j})]\, dm_j < +\infty$

[P4*] $(-1)^{J-1} \int_{\mathbf{m_{-j}}=-\infty}^{+\infty} \frac{\partial^{J-1} P_{\mathcal{C}}(j|\mathbf{m})}{\partial m_1 \ldots \partial m_{j-1} \partial m_{j+1} \ldots \partial m_J}\, d\mathbf{m_{-j}} = 1$ for each m_j

[P5*] $(-1)^{J-1} \int_{m_i=-\infty}^{+\infty} |m_i| \frac{\partial P_{\mathcal{C}}(j|\mathbf{m})}{\partial m_k}\, dm_i < +\infty$ for each m_j and $k \neq j$

Definition: ARUM-consistent

A complete or partial CPS is ARUM-consistent if there exists an ARUM with CDF F such that for $1 \leq j \leq k \leq J$,

$$\begin{aligned} P_{\mathcal{C}}(j|\mathbf{m}) &= \text{Prob}(\mathbf{U}|U_j + m_j \geq U_k + m_k \text{ for } k \neq j) \\ &\equiv \int_{-\infty}^{+\infty} F_j(v - m_1, \ldots, v - m_j)\, dv \end{aligned}$$