Recent methodological developments in discrete choice models

Michel Bierlaire

Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Outline

- Introduction
- MEV models
- Mixtures of MEV
- Testing
- Route choice (no time...)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Introduction

- Discrete choice models:

$$P(i|\mathcal{C}_n) \text{ where } \mathcal{C}_n = \{1,\ldots,J\}$$

- Random utility models:

$$U_{in} = V_{in} + \varepsilon_{in}$$

and

$$P(i|\mathcal{C}_n) = P(U_{in} \geq U_{jn}, j = 1,\ldots,J)$$

- Utility is a latent concept

Multinomial Logit Model

- Assumption: ε_{in} are i.i.d. Extreme Value distributed.
- Independence is both across i and n
- Choice model:

$$P(i|\mathcal{C}_n) = \frac{e^{V_{in}}}{\sum_{j\in\mathcal{C}_n} e^{V_{jn}}}$$

Relaxing the independence assumption

...across alternatives

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

that is

$$U_n = V_n + \varepsilon_n$$

and ε_n is a vector of random variables.

Relaxing the independence assumption

- $\varepsilon_n \sim N(0, \Sigma)$: multinomial probit model
  - No closed form for the multifold integral
  - Numerical integration is computationally infeasible
- Extensions of multinomial logit model
  - Nested logit model
  - Multivariate Extreme Value (MEV) models

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

Family of models proposed by McFadden (1978)
Idea: a model is generated by a function

$$G : \mathbb{R}^J \longrightarrow \mathbb{R}$$

From G, we can build

- The cumulative distribution function (CDF) of ε_n
- The probability model
- The expected maximum utility

Called Generalized EV models in DCM community

MEV models

1. G is homogeneous of degree $\mu > 0$, that is

$$G(\alpha x) = \alpha^\mu G(x)$$

2. $\lim_{x_i \to +\infty} G(x_1, \ldots, x_i, \ldots, x_J) = +\infty$, $\forall i$,

3. the kth partial derivative with respect to k distincts x_i is non negative if k is odd and non positive if k is even, i.e., for all (disctincts) indices $i_1, \ldots, i_k \in \{1, \ldots, J\}$, we have

$$(-1)^k \frac{\partial^k G}{\partial x_{i_1} \ldots \partial x_{i_k}}(x) \leq 0, \ \forall x \in \mathbb{R}^J_+.$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

- Density function:
  $F(\varepsilon_1,\ldots,\varepsilon_J) = e^{-G(e^{-\varepsilon_1},\ldots,e^{-\varepsilon_J})}$
- Probability: $P(i|C) = \frac{e^{V_i+\ln G_i(e^{V_1},\ldots,e^{V_J})}}{\sum_{j\in C} e^{V_j+\ln G_j(e^{V_1},\ldots,e^{V_J})}}$ with $G_i = \frac{\partial G}{\partial x_i}$. This is a closed form
- Expected maximum utility: $V_C = \frac{\ln G(\ldots)+\gamma}{\mu}$ where γ is Euler's constant.
- Note: $P(i|C) = \frac{\partial V_C}{\partial V_i}$.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MEV models

Example: Multinomial logit:

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

MEV models

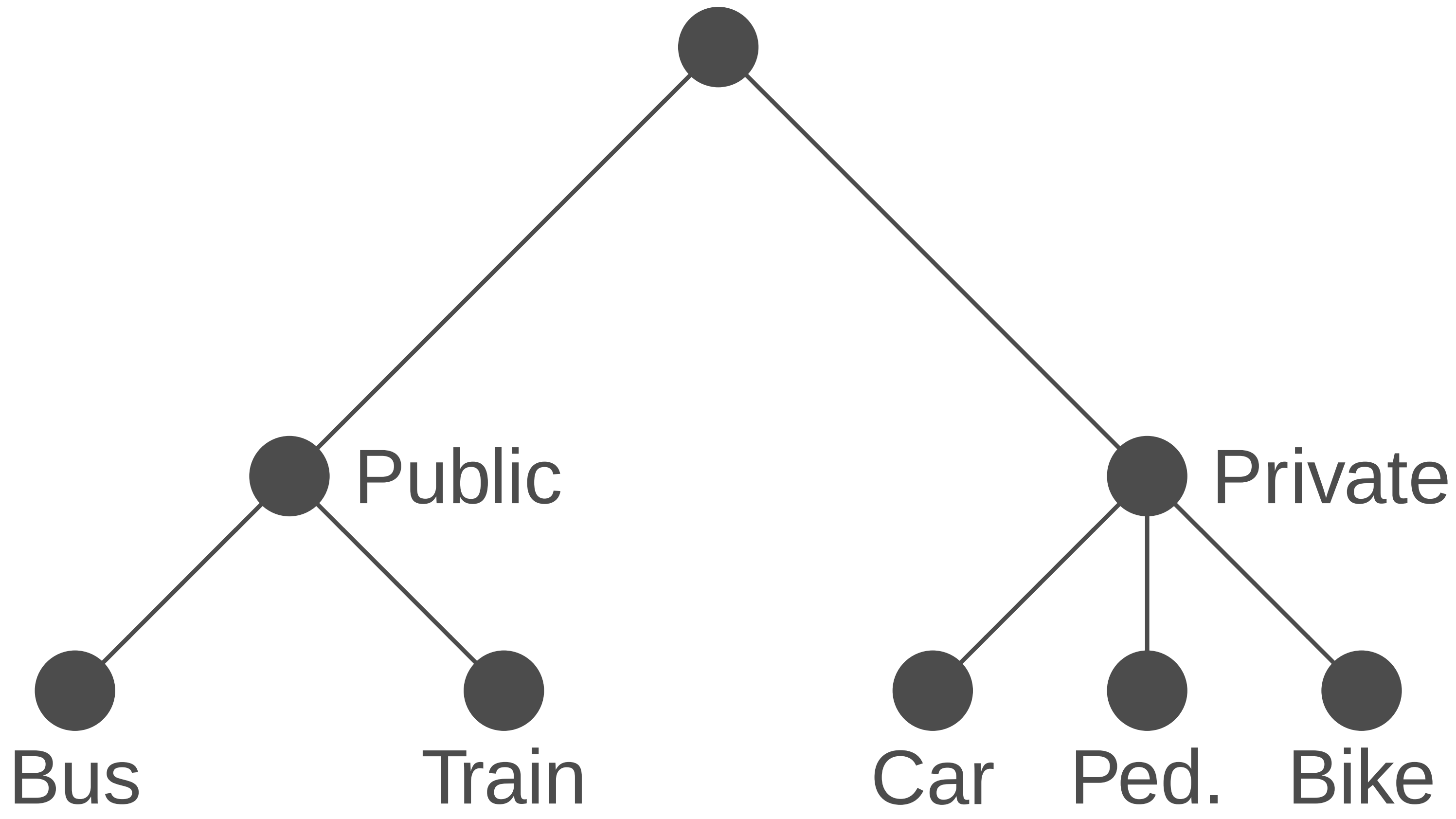
Example: Nested logit

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

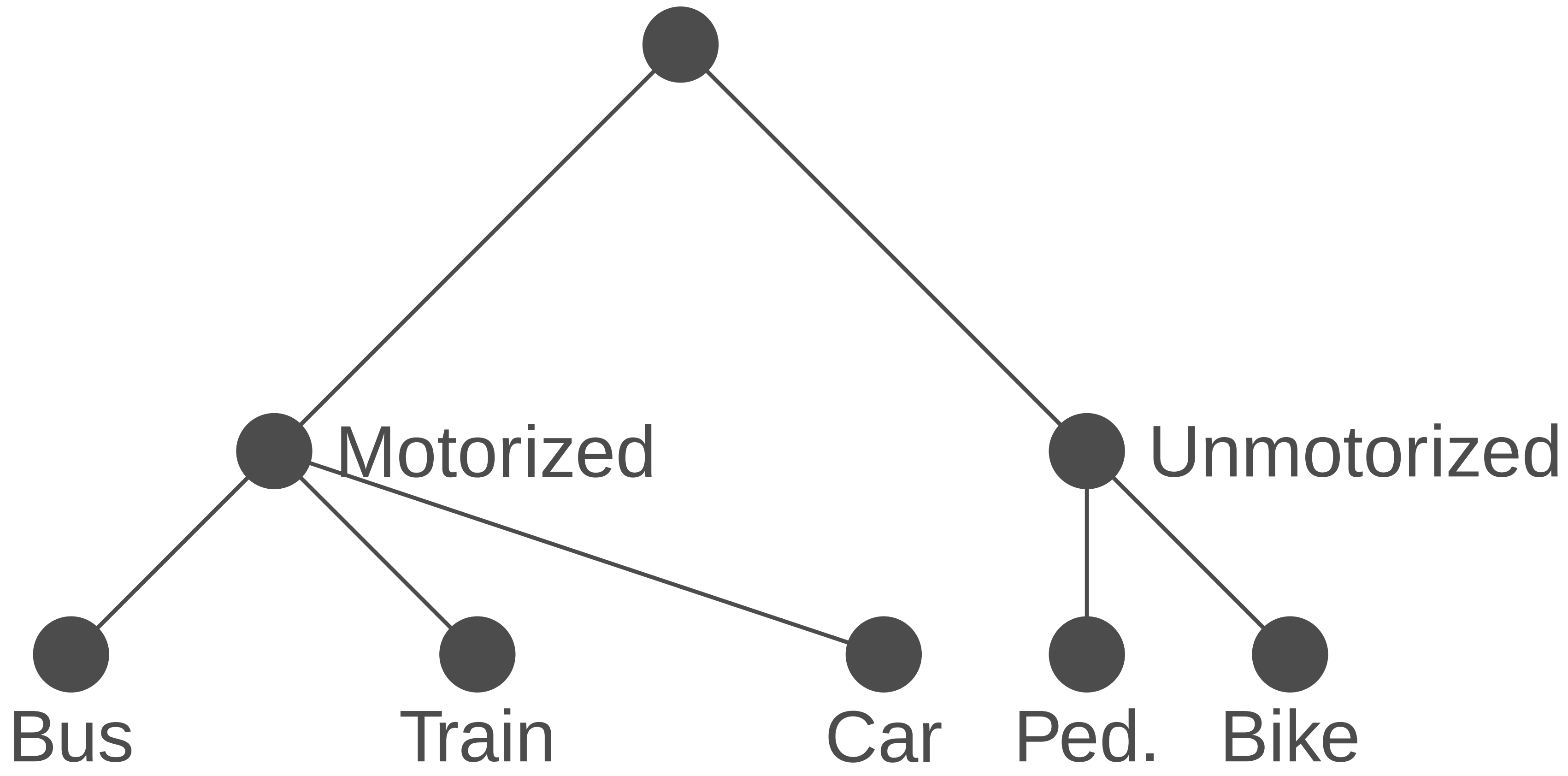
Example: Cross-Nested Logit

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

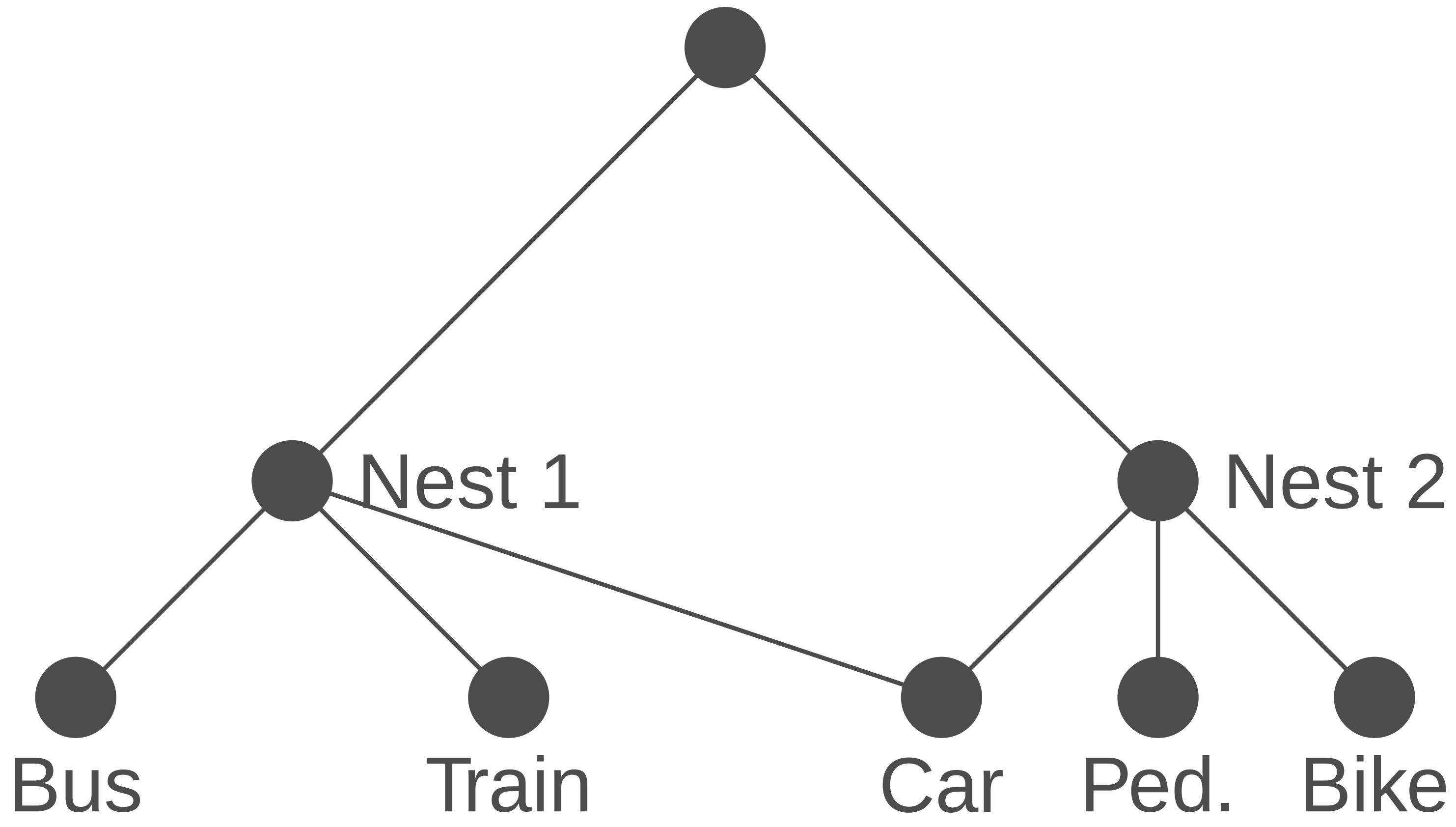
Nested Logit Model



Nested Logit Model



Cross-Nested Logit Model



MEV models

Advantages:

- Closed form probability model
- Provides a great deal of flexibility

MEV models

Issues:

- Formulation not in term of correlations

  Abbe, Bierlaire & Toledo (2005)
- Require heavy proofs

  Daly & Bierlaire (2006)
- Homoscedasticity
  - McFadden & Train (2000)
- Sampling issues
  - Bierlaire, Bolduc & McFadden (2006)

Mixture of MEV

In statistics, a mixture density is a pdf which is a convex linear combinations of other pdf's.
If $f(\varepsilon, \theta)$ is a pdf, and if $w(\theta)$ is a nonnegative function such that $\int_a w(a)da = 1$ then

$$g(\varepsilon) = \int_a w(a) f(\varepsilon, \theta) da$$

is also a pdf. We say that g is a mixture of f.

If f is the pdf of a MEV model, it is a mixture of MEV

Mixture of MEV

Discrete mixtures are also possible. If $f(\varepsilon, \theta)$ is a pdf, and if w_i, $i = 1, \ldots, n$ are nonnegative weights such that $\sum_{i=1}^{n} w_i = 1$ then

$$g(\varepsilon) = \sum_{i=1}^{n} w_i f(\varepsilon, \theta_i)$$

is also a pdf. We say that g is a discrete mixture of f.

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mixture of MEV

Common terminology:

- Mixed logit: incorrect
- Logit kernel: correct
- Hybrid model: inaccurate

Most appropriate terminology:

mixture of logit models
mixture of MEV models

If $w(a)$ is a normal pdf, we have

normal mixture of MEV models

Mixture of MEV

$$U_n = V_n + \varepsilon_n$$

- ε_n compliant with MEV theory
- V_n contains random parameters.

$$V_n = \beta^T X_n \text{ where } \beta \sim N(\widehat{\beta}, \Sigma)$$

- Using the Cholesky factorization, we have

$$\beta = \widehat{\beta} + P\zeta \text{ where } \Sigma = PP^T$$

and ζ are i.i.d. standard normal variates.

Heteroscedastic model

- Random parameter = alternative specific constant
- Error term becomes:

$$\varepsilon_{in} = \xi_{in} + \nu_{in}$$

- $\xi_{in} \sim N(c_i, \sigma_i^2)$
- $\mu_{in} \sim$ MEV

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Panel data

- Same individual observed several times
- Utility:

$$U_{int} = V_{int} + \xi_{in} + \nu_{int}$$

- Probability

$$P_n(i|\mathcal{C}_n) = \prod_t P_{nt}(i|\mathcal{C}_{nt})$$

where $\mathcal{C}_n = \cup_{t \in T_n} \mathcal{C}_{nt}$

- ξ_{in} is not distributed across observations, only across individuals

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Mixture of MEV

- McFadden & Train(2000)
  "Under mild regularity conditions, any discrete choice model derived from random utility maximization has choice probabilities that can be approximated as closely as one pleases by a Mixed MNL model."
- Why bother with Mixture of MEV?

Mixture of MEV

- MEV has closed form formulation
- Mixture models require simulated maximum likelihood estimation
- Capture as much as possible of the correlation using MEV
- Use the mixing distribution for the rest
- Issue: estimation

BIOGEME

Motivations

- MEV family must be explored
- Complicated implementation
- No appropriate software package
- Most researchers use commercial packages: LIMDEP, ALOGIT, HieLoW or Gauss, Matlab, SAS
- Freeware: Kenneth Train (but based on Gauss)

BIOGEME

Objectives

- Maximum likelihood estimation of a wide variety of MEV models
- Use various nonlinear optimization algorithms
- Open source
- Designed for researchers
- Flexible and easily extensible

BIOGEME

BIerlaire's Optimization toolbox for GEV Models Estimation

`biogeme.epfl.ch`

Testing

- Mixtures of MEV is very flexible (too flexible?)
- Choice of the distribution for the random parameter is important
- Need for a test to check if it is appropriate
- Fosgerau & Bierlaire (2006).

Testing: main ideas

- Random parameter: ω
- Base (postulated) distribution: f, F
- True distribution: g, G
- Unknown transformation Q, monotonic, such that

$$G(\omega) = Q(F(\omega)),$$

- Densities:

$$g(\omega) = q(F(\omega))f(\omega).$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Testing: main ideas

- Approximate q using polynomials.

$$q_N(x) = 1 + \sum_{k=1}^{N} \delta_k L_k(x),$$

- L_k are transformed Legendre polynomials
- Define

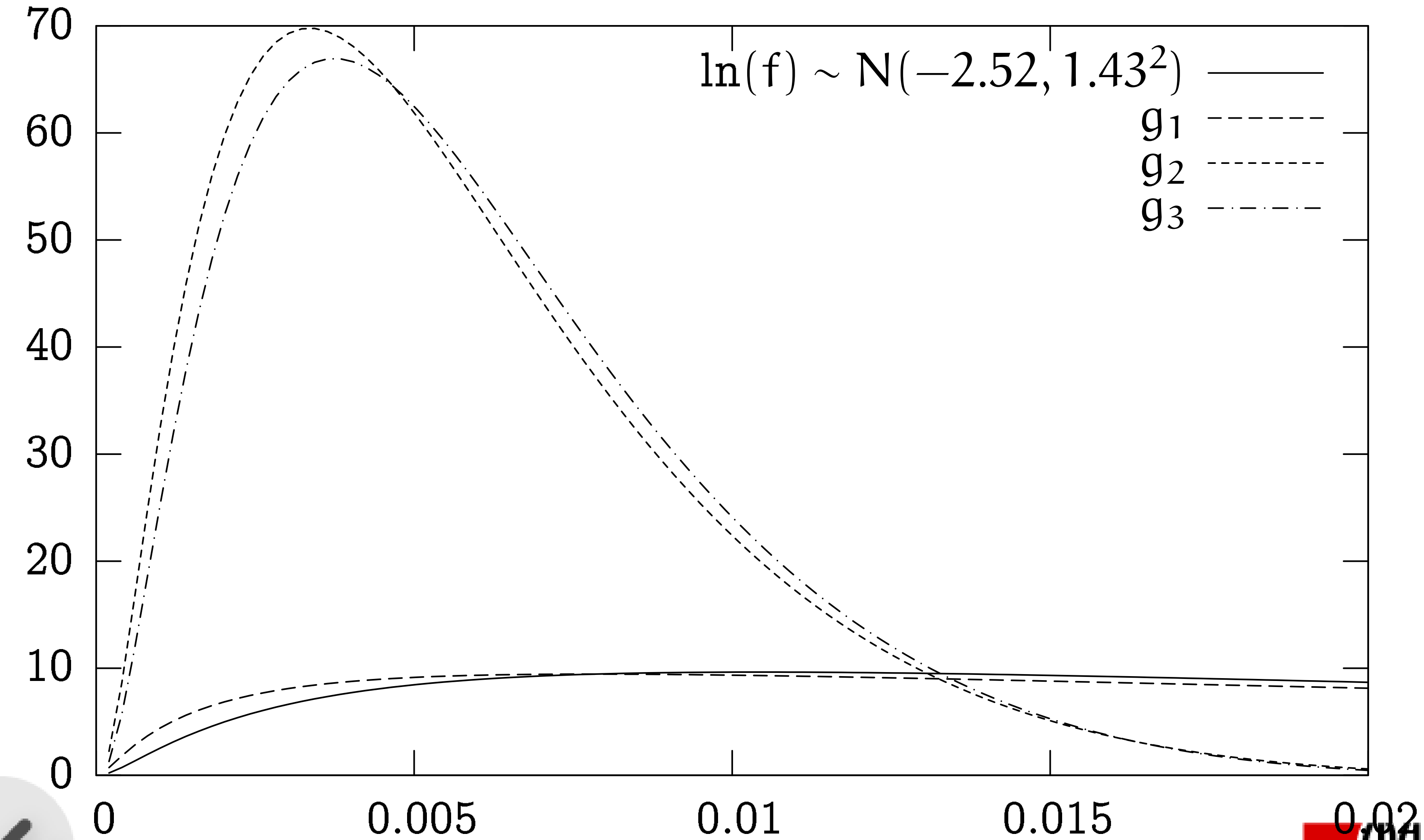
$$q(x) \approx \frac{1}{K} q_N^2(x),$$

where $K = \int_{-\infty}^{+\infty} q_N^2(F(\omega)) f(\omega) d\omega$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Testing: main ideas



Testing: main ideas

- Under the null hypothesis that $f = g$,

$$P_n(i|\mathcal{C}_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) g(\beta) d\beta,$$

is equivalent to the model

$$P_n(i|\mathcal{C}_n) = \int_{-\infty}^{+\infty} P_n(i|\beta, \mathcal{C}_n) f(\beta) d\beta.$$

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Testing: main ideas

- The two models are nested
- Likelihood ratio test can be used to test if the models are indeed equivalent
- Test implemented in Biogeme

Short course

Lausanne, March 25-29, 2007

Ben-Akiva, McFadden, Bierlaire, Bolduc

http://transp-or.epfl.ch/dca