Fast algorithms to generate individualized designs for the mixed logit choice model

Marjolein Crabbe   Martina Vandebroek

KU Leuven, Belgium

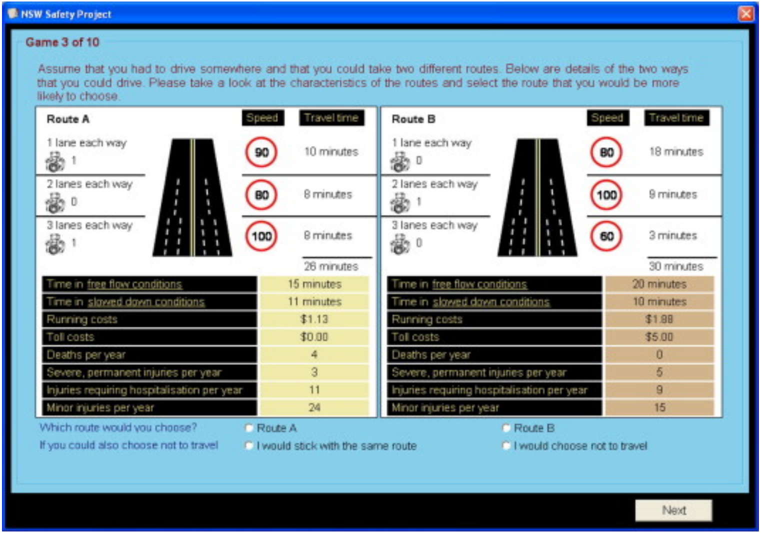
mODa 10, June 10 2013

Discrete choice experiments

- Survey methodology to study the preferences of consumers
- In a discrete choice experiment respondents must choose their preferred product in a series of choice sets contrasting multiple alternatives
- Each alternative or profile in a set is characterized by a number of attributes
- The attributes take on specific values or levels

$\Rightarrow$ The choices reveal the relative value that consumers attach to the different attributes of the product

Discrete choice experiments



Discrete choice analysis

- Choice models are based on utility maximization
- The utility that individual n receives from alternative k in choice set s

$$U_{ksn} = \mathbf{x}'_{ksn}\boldsymbol{\beta} + \varepsilon_{ksn}$$

  - $\mathbf{x}_{ksn}$ the attribute levels of the alternative
  - $\boldsymbol{\beta}$ the relative importance of the attributes
- The conditional logit choice model
  - Probability that individual n chooses alternative k in choice set s

$$p_{ksn}(\boldsymbol{\beta}) = \frac{e^{\mathbf{x}'_{ksn}\boldsymbol{\beta}}}{\sum_{t=1}^{K} e^{\mathbf{x}'_{tsn}\boldsymbol{\beta}}}$$

  - Assumes a homogeneous population: all people equally value the product attributes

Discrete choice analysis

- The mixed logit choice model
  - Accounts for heterogeneity in the preferences
  - Individual-specific coefficients $\boldsymbol{\beta}_n$
  - Aggregate choice behavior in the population modeled with a heterogeneity distribution

$$\boldsymbol{\beta}_n \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

  - Unconditional mixed logit choice probability

$$p_{ksn} = \int \frac{e^{\mathbf{x}'_{ksn}\boldsymbol{\beta}_n}}{\sum_{t=1}^{K} e^{\mathbf{x}'_{tsn}\boldsymbol{\beta}_n}} \ \phi(\boldsymbol{\beta}_n|\boldsymbol{\mu}, \boldsymbol{\Sigma}) \ d\boldsymbol{\beta}_n$$

Individualized design for the mixed logit choice model

- Bliemer and Rose (2010) constructed aggregate **locally** $\mathcal{D}$-efficient designs for the mixed logit choice model
- Generating aggregate Bayesian $\mathcal{D}$-efficient designs, taking the uncertainty about the model parameters into account, appeared infeasible in a reasonable amount of time

$$\Downarrow$$

- Individualized design
  - $\boldsymbol{\beta}_n$ assumed constant over all choice sets
  - The preferences of a specific individual are thus in essence modeled by a conditional logit choice model
  - Individual efficient designs with respect to the underlying conditional logit choice models

Individualized design for the mixed logit choice model

- The choice experiments are sequentially generated for each person separately, based on choice information from the previously administered choice sets
- Tailored to the specific preferences of an individual
- Online, interactive choice experiments
    1. Assume a prior distribution $f(\boldsymbol{\beta}_n) \equiv \phi(\boldsymbol{\beta}_n | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$
    2. After respondent n has completed $s-1$ choice sets: Bayesian update of the prior information on $\boldsymbol{\beta}_n$

$$f(\boldsymbol{\beta}_n | \mathbf{y}_n^{s-1}) = \frac{L(\boldsymbol{\beta}_n | \mathbf{y}_n^{s-1}, \mathbf{X}_n^{s-1})\ \phi(\boldsymbol{\beta}_n | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)}{\int L(\boldsymbol{\beta}_n | \mathbf{y}_n^{s-1}, \mathbf{X}_n^{s-1})\ \phi(\boldsymbol{\beta}_n | \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)\ d\boldsymbol{\beta}_n}$$

    3. The next choice set is efficiently selected with the updated information
    4. Repetition of steps 2 and 3 until a specific amount of choice sets is obtained

Individualized design for the mixed logit choice model

- Comparison of four design criteria
  - Minimum posterior weighted $\mathcal{D}$-error ($\mathcal{D}_B$)
  - Novel criteria from optimal test design based on Kullback-Leibler divergence
    - Maximum expected Kullback-Leibler divergence between subsequent posteriors (KLP)
    - Maximum mutual information (MUI)
    - Minimum expected posterior entropy (ENT)

Minimum posterior weighted $\mathcal{D}$-error

- The (Bayesian) Fisher information matrix for design $\mathbf{X}_n^S$ with S choice sets for individual n

$$\mathbf{I}_{BFIM}(\boldsymbol{\beta}_n, \mathbf{X}_n^S) = -\mathsf{E}\left[\frac{\partial^2 \log[L(\boldsymbol{\beta}_n|\mathbf{y}_n^S, \mathbf{X}_n^S)\ f(\boldsymbol{\beta}_n)]}{\partial\boldsymbol{\beta}_n \partial\boldsymbol{\beta}_n'}\right]$$

with $f(\boldsymbol{\beta}_n)$ a prior for $\boldsymbol{\beta}_n$

- Assuming a normal prior with covariance matrix $\boldsymbol{\Sigma}_0$

$$\begin{aligned}\mathbf{I}_{BFIM}(\boldsymbol{\beta}_n, \mathbf{X}_n^S) &= \mathbf{I}_{FIM}(\boldsymbol{\beta}_n, \mathbf{X}_n^S) + \boldsymbol{\Sigma}_0^{-1} \\ &= \sum_{s=1}^{S} \mathbf{X}_{sn}'(\mathbf{P}_{sn} - \mathbf{p}_{sn}\mathbf{p}_{sn}')\mathbf{X}_{sn} + \boldsymbol{\Sigma}_0^{-1}\end{aligned}$$

Minimum posterior weighted $\mathcal{D}$-error

- Bayesian $\mathcal{D}$-efficient designs: minimize $\mathcal{D}$-error averaged over a weighting distribution for $\boldsymbol{\beta}_n$
- To select the sth choice set, minimize

$$\int \text{det}[\mathbf{I}_{BFIM}(\boldsymbol{\beta}_n, \mathbf{X}_n^s)]^{-1/p} \; f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}) \; d\boldsymbol{\beta}_n$$

Maximum expected Kullback-Leibler divergence between subsequent posteriors

- Kullback-Leibler divergence between two densities f and g for a continuous variable X

$$KL(f,g) \;=\; \int f(x) \log \frac{f(x)}{g(x)} \, dx$$

  - For any f and g, KL is non-negative and zero in case of equal densities
  - $KL(f,g)$ increases as the two densities become more divergent
  - "Distance between two densities"
  - Not a real distance measure (for instance non-symmetric: $KL(f,g) \neq KL(g,f)$)

Maximum expected Kullback-Leibler divergence between subsequent posteriors

- To select the sth set in a choice experiment, maximize the expected Kullback-Leibler distance between the current posterior distribution of $\boldsymbol{\beta}_n$ and the updated posterior one obtains with the answer to the sth choice set

$$\sum_{k=1}^{K} \pi(y_{ksn}|\mathbf{y}_n^{s-1})\ KL[f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}), f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}, y_{ksn})]$$

with

$$\pi(y_{ksn}|\mathbf{y}_n^{s-1}) = \int p_{ksn}(\boldsymbol{\beta}_n)\ f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1})\ d\boldsymbol{\beta}_n$$

Maximum mutual information

- Mutual information between two variables X and Y

$$I_M(X,Y) = \int_Y \int_X f(x,y) \log \frac{f(x,y)}{f(x)f(y)} \; dxdy$$

  - Kullback-Leibler distance between the joint distribution of X and Y and their distribution in case of independence
  - Expresses how much information one variable holds with respect to the other

- To select the sth set in a choice experiment, maximize the mutual information between the individual coefficients $\boldsymbol{\beta}_n$ and the choice for the next set, given the choice data of the previously administered sets

$$\sum_{k=1}^{K} \int f(\boldsymbol{\beta}_n, y_{ksn}|\mathbf{y}_n^{s-1}) \log \frac{f(\boldsymbol{\beta}_n, y_{ksn}|\mathbf{y}_n^{s-1})}{f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1})\pi(y_{ksn}|\mathbf{y}_n^{s-1})} \; d\boldsymbol{\beta}_n$$

Minimum expected posterior entropy

- Entropy for a continuous variable X and density $f(x)$

$$H(X) = -\int f(x) \log f(x) \, dx$$

- Measure of uncertainty
- To select the sth set in a choice experiment, minimize the expected posterior entropy, or equivalently maximize

$$\sum_{k=1}^{K} \pi(y_{ksn}|\mathbf{y}_n^{s-1}) \int f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}, y_{ksn}) \log f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}, y_{ksn}) \, d\boldsymbol{\beta}_n$$

Connections among the Kullback-Leibler design criteria

- KLP is the expected Kullback-Leibler distance between the current and the updated posterior

$$\sum_{k=1}^{K} \pi(y_{ksn}|\mathbf{y}_n^{s-1}) \; KL[f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}), f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}, y_{ksn})]$$

- MUI is the expected Kullback-Leibler distance between the updated and the current posterior

$$\sum_{k=1}^{K} \pi(y_{ksn}|\mathbf{y}_n^{s-1}) \; KL[f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1}, y_{ksn}), f(\boldsymbol{\beta}_n|\mathbf{y}_n^{s-1})]$$

- ENT is the expected Kullback-Leibler distance between the updated posterior and a uniform distribution

Comparison study

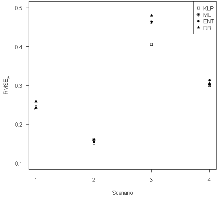
- Multiple experimental setups

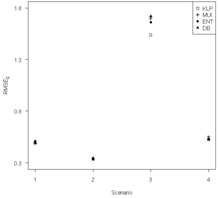
Scenario 1	$3^3/2/15$
Scenario 2	$2 \times 3 \times 2 \times 3/3/15$
Scenario 3	$3 \times 2^4/2/15$
Scenario 4	$3 \times 2 \times 3/3/15$

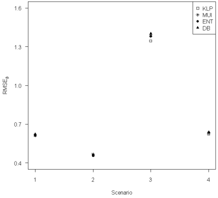
- 50 respondents
- Response simulation repeated 100 times

Estimation accuracy

Mean RMSE$_{\boldsymbol{\mu}}$, RMSE$_{\boldsymbol{\Sigma}}$ and RMSE$_{\boldsymbol{\beta}}$ values obtained with KLP, MUI, ENT and $\mathcal{D}_B$ for the different scenarios







$\Rightarrow$ No significant differences in estimation accuracy for scenario 1, 2 and 4

$\Rightarrow$ In scenario 3, KLP outperforms the other methods

Computation time

Average computation time (seconds) to select an additional choice set with KLP, MUI, ENT and $\mathcal{D}_B$ using various numbers of draws

	Scenario 1			Scenario 2		
	512	1024	2048	512	1024	2048
KLP	0.074	0.152	0.285	1.726	3.381	6.729
MUI	0.082	0.152	0.293	1.773	3.484	6.924
ENT	0.090	0.168	0.328	1.972	3.866	7.674
$\mathcal{D}_B$	1.789	3.269	6.523	35.689	71.277	142.296

Computation time

	Scenario 3			Scenario 4		
	512	1024	2048	512	1024	2048
KLP	0.207	0.402	0.805	0.219	0.418	0.809
MUI	0.215	0.414	0.816	0.223	0.426	0.836
ENT	0.242	0.468	0.914	0.246	0.473	0.934
$\mathcal{D}_B$	5.207	10.375	20.671	3.855	7.702	15.436

$\Rightarrow$ Impressive decrease in computation time from using the Kullback-Leibler design criteria instead of $\mathcal{D}_B$

$\Rightarrow$ The $\mathcal{D}_B$ computation times are approximately 20 times the KLP times

Conclusions

- This research focusses on improving the practicability of individualized choice design for the mixed logit choice model using criteria from optimal test design
- Comparison of four design algorithms
  - The efficiency to estimate the mixed logit choice model of the designs obtained with the four criteria is equivalent
  - The Kullback-Leibler criteria are preferred over $\mathcal{D}$-efficiency due to their low complexity, yielding a huge decrease in computation time
  - The Kullback-Leibler criteria warrant the feasibility of individualized choice design

Fast algorithms to generate individualized designs for the mixed logit choice model

Marjolein Crabbe Martina Vandebroek

KU Leuven, Belgium

mODa 10, June 10 2013