Sampling Of Alternatives In Random Regret Minimization Models

C. Angelo Guevara
Faculty of Engineering and Applied Science,
Universidad de los Andes, Santiago, Chile
Caspar G. Chorus
Faculty of Technology, Policy and Management
Delft University of Technology
Moshe E. Ben-Akiva
Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
C. Angelo Guevara
Faculty of Engineering and Applied Science,
Universidad de los Andes, Santiago, Chile
Email: aguevara@uandes.cl

1 Need for Sampling of Alternatives in Discrete Choice Modeling

When choice sets are very large, like is the case in many route- and destination-choice models, sampling of alternatives becomes necessary to ensure the practical feasibility of discrete choice-model formulation and estimation. In the context of the classical Random Utility Maximization-based (RUM) Logit model, a convenient method has been proposed (McFadden, 1978) to obtain a consistent estimator for model parameters. This estimator capitalizes on the fact that, due to its independently and identically distributed (or: iid) errors, the RUM-based Logit model exhibits the IIA-property.

Although very convenient from a modeler's perspective, this IIA-property is often considered to be restrictive in terms of the implied behavior of decision-makers. Over the past few decades, this observation has led to the development of a number of alternative discrete choice model forms whose errors are not iid.

2 The Random Regret Minimization Model (RRM)

Recently, a choice model has been approach that does not exhibit the IIA-property even though (when written in Logit-form) its errors are iid. This Random Regret Minimization (RRM) model (Chorus, 2010), which is the focus of this research, is based on a regret minimization-based decision rule. The model postulates that when decision makers choose between alternatives, they try to avoid the situation where a non-chosen alternative outperforms a chosen one in terms of one or more attributes. This translates into a regret function for a considered alternative that by definition features all attributes of all competing alternatives. Since its introduction a few years ago, the RRM model has been successfully estimated and applied by various authors in the context of a variety of different choice contexts, involving – to name a few examples – travelers choices between vehicle types, destinations, modes, routes, departure times, and driving maneuvers; politicians' choices between policy options; patients choices between medical treatments; and tourists' choices between leisure activity-locations. Recent studies on RRM can be found in, for example, Chorus & de Jong (2011), Kaplan & Prato (2012), Hensher et al. (2012).

One disadvantage of the RRM model which was highlighted in Chorus (2012) is that runtimes may suffer from combinatorial explosion when choice sets become very large. This issue of course is a direct result from the behavioral postulate, incorporated in the regret function, that every alternative is compared with every other alternative in the choice set in terms of every attribute. As a consequence, finding a proper way to estimate RRM models on sampled choice sets is an important condition for the model to be useful in the context of choice situations involving very large numbers of alternatives. At this point it should be noted that, because of the fact that the RRM model does not exhibit the IIA-property, McFadden's 1978-result does not apply. As mentioned, this is the case even when – such as is the case for RRM-based Logit models – errors are distributed iid.

3 A Novel method for Sampling of Alternatives in RRM

Guevara and Ben-Akiva (2012) recently proposed a method to address sampling of alternatives in MEV models. The method consists in expanding the components of the MEV model that get truncated because of the sampling procedure. In the present article it is shown that the method proposed by Guevara and Ben-Akiva (2012) can be directly extended the problem of sampling of alternatives in RRM.

The paper analyzes the conditions required for consistency, asymptotic normality and efficiency using first order asymptotics. Two cases are considered: when the researcher has full access to that data and when he or she does not. In the second case, the expansion factors required for addressing the problem of sampling of alternatives in RRM depend on the choice probabilities. To address this difficulty, two practical methods are studied.

Besides the theoretical derivation, a Monte Carlo experiment is performed to illustrate the application of the proposed method, to study the small sample properties of the estimators, and to analyze the efficacy and efficiency of the method in recovering the true coefficients of each model depending of the number of alternatives sampled.

4 Conclusion

In line with expectations, Monte Carlo experiments showed that sampling of alternatives causes a significant bias in the estimators of the model parameters and in the estimated shares when no correction is applied. In addition, the proposed method for correcting the terms that get truncated because of the sampling performed reasonably well. In cases where the researcher has full control of the data and it is possible to obtain an additional sample to expand the terms that get truncated, the method proposed is easily applicable. When it is not possible to re-sample, the method requires knowledge of the choice probabilities in order to build the expansion factors. In this final case, two practical approximation methods showed reasonably good results.

The sample size required to obtain good estimators while sampling alternatives in Random Regret models will vary on a case-by-case basis and cannot be expressed as a percentage of the cardinality of the true choice-set. In general, an appropriate strategy to determine if the size of the sample of alternatives is large enough is to test the stability of the estimators with different number of alternatives sampled.

References

Chorus, C.G., 2010. A new model of Random Regret Minimization. *European Journal of Transport and Infrastructure Research*, 10(2), 181-196

Chorus, C.G., de Jong, G.C., 2011. Modeling experienced accessibility for utility-maximizers and regret-minimizers. *Journal of Transport Geography*, 19, 1155-1162

Chorus, C.G., 2012. *Random regret-based discrete choice modeling: A tutorial*. Springer Briefs in Business, Springer, Heidelberg, Germany

Guevara C.A. and Ben-Akiva, M., 2012. Sampling of Alternatives in MEV Models. *Transportation Research Part B.* Accepted.

Hensher, D.A., Greene, W.H., Chorus, C.G., 2012. Random Regret Minimization or Random Utility Maximization: An exploratory analysis in the context of automobile fuel choice. *Journal of Advanced Transportation* (in press).

Kaplan, S., Prato, C.G., 2012. The application of the random regret minimization model to drivers' choice of crash avoidance maneuvres. *Transportation Research Part F* (in press)

McFadden D. 1978. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Residential Location*, Karlquist, Lundqvist, Snickers and Weibull (eds). North Holland, Amsterdam, 75-96.