Transportation Research Part B 59 (2014) 81–95

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Extreme values, invariance and choice probabilities

Lars-Göran Mattsson [a,*], Jörgen W. Weibull [b,c,d], Per Olov Lindberg [a]

[a] KTH Royal Institute of Technology, Department of Transport Science, Stockholm, Sweden
[b] Stockholm School of Economics, Sweden
[c] Institute for Advanced Study in Toulouse, France
[d] KTH Royal Institute of Technology, Department of Mathematics, Stockholm, Sweden

ARTICLE INFO

Article history:
Received 5 July 2013
Received in revised form 29 October 2013
Accepted 30 October 2013

Keywords:
Extreme values
Discrete choice
Choice probabilities
Random utility
Invariance

ABSTRACT

Since the pioneering work of McFadden (1974), discrete choice random-utility models have become work horses in many areas in transportation analysis and economics. In these models, the random variables enter additively or multiplicatively and the noise distributions take a particular parametric form. We show that the same qualitative results, with closed-form choice probabilities, can be obtained for a wide class of distributions without such specifications. This class generalizes the statistically independent distributions where any two c.d.f.:s are powers of each others to a class that allows for statistical dependence, in a way analogous to how the independent distributions in the MNL models were generalized into the subclass of MEV distributions that generates the GEV choice models. We show that this generalization is sufficient, and under statistical independence also necessary, for the following invariance property: all conditional random variables, when conditioning upon a certain alternative having been chosen, are identically distributed. While some of these results have been published earlier, we place them in a general unified framework that allows us to extend several of the results and to provide proofs that are simpler, more direct and transparent. Well-known results are obtained as special cases, and we characterize the Gumbel, Fréchet and Weibull distributions.

1. Introduction

Since the pioneering work of McFadden (1974), discrete choice random-utility models, in particular logit models and their extensions (such as nested and mixed logit), have become work horses in many areas of economics, such as travel demand, labor markets, housing markets, trade and environmental economics. What these applications have in common is that there is a stochastic selection process that picks a maximal or minimal element from a finite set of alternatives.

In random utility applications, there is a population of decision-makers, each of whom faces a finite set of alternatives, and each decision-maker chooses an alternative that he or she deems to be best. The analyst usually has data about some aspects of the alternatives at hand, and may also have some about the population. However, individuals may have unobserved preference differences and the alternatives may have characteristics that are unobserved by the analyst, but that may be observed by some or all decision-makers, who may care more or less about these hidden characteristics. The analyst treats each individual's choice as probabilistic, driven by a collection of random variables, one for each alternative, representing the utility or cost of the decision alternative. The decision-maker is assumed to select an alternative that either maximizes his or her utility or minimizes his or her cost. The decision-maker may be a commuter in an urban transportation

* Corresponding author. Tel.: +46 87908563.
E-mail addresses: lars-goran.mattsson@abe.kth.se (L.-G. Mattsson), jorgen.weibull@hhs.se (J.W. Weibull), polin@kth.se (P.O. Lindberg).

0191-2615/$ - see front matter
http://dx.doi.org/10.1016/j.trb.2013.10.014

network, a consumer in a product market, a firm in international trade, or a producer who strives to minimize cost. We call the probability for a given alternative to be selected its choice probability.

In his pioneering work, McFadden (1974) assumes that the utility of each alternative is the sum of a noise term and some real-valued function of observed characteristics of the alternative. The noise terms are assumed i.i.d. Gumbel distributed, and hence the random utilities will also be independent (but typically not identically) Gumbel distributed, resulting in closed-form choice probabilities of the multinomial logit type.[1] Since then, a vast literature has emerged, generalizing this pioneering approach in several directions, such as nested logit, mixed logit, and generalized extreme-value (GEV) choice models (see e.g. Train, 2003).

Recently, there has appeared a literature that goes beyond the logit formulation, such as Castillo et al. (2008) and Fosgerau and Bierlaire (2009), who instead assume that decision-makers are cost minimizers with independent multiplicative Weibull distributed noise terms.[2] Castillo et al. (2008) derive the choice probabilities for the Weibull choice model and Fosgerau and Bierlaire (2009) develop a number of its properties and demonstrate its better empirical fit in some travel applications. This line of research is further developed in Li (2011), who shows that in the independent case there exist closed-form choice probabilities under cost-minimization for a much wider class of distributions. More exactly, the survival functions for any two random cost variables need only be a positive power of each other. The class of survival functions that are positive integer powers of a given survival function is the smallest class of independent survival functions that is closed under minimization and includes the given survival function.[3] Li (2011) further shows how these theoretical results can be explored to achieve better data fit in certain applications. Dalal and Klein (1988) explore the same basic insight, but for the utility maximizing case, when they propose a flexible class of discrete choice models termed "generalized logit models," a class closely related to mixed logit models. As compared to Li (2011) and Dalal and Klein (1988), we study a more general class of distributions that allows for statistical dependence in a way analogous to how statistical dependence was introduced among Gumbel distributions when multinomial logit models were generalized into GEV choice models. Thus, some of their results are special cases of ours for statistically independent distributions. A common feature of most models suggested in this literature is that they each assume some functional form, such as additivity or multiplicativity, of some real-valued functions of the observed characteristics of the alternatives and the decision-makers, and a certain parametric representation of the noise terms.

In this paper, we formulate a class of random utility models that allows us to collect and extend several results in the literature in a unified way. It also allows us to provide proofs of earlier or extended results that are simpler, more direct and transparent without use of heavier mathematics than standard calculus.[4] Within this general framework, we also characterize the Gumbel, Fréchet and Weibull distributions. Our framework is a generalization of the subclass of multivariate extreme value (MEV) distributions introduced by McFadden (1978), see also Smith (1984), and thus allows for statistical dependence among (unobserved) noise terms.[5] We show that this generalization is sufficient, and in the special case of statistical independence also necessary, for a certain invariance property, namely that conditioning upon a certain alternative having been chosen (being best) provides no information about the utility achieved by this choice. In other words, all conditional random variables, one for each alternative when being chosen, have the same probability distribution. Such invariance has been observed before. Strauss (1979) introduced a class of additive random utility (ARU) models, containing the GEV choice models, by allowing the additive noise vector to have a more general c.d.f. Robertson and Strauss (1981) sketch a proof of Strauss' (1979) claim that the invariance property characterizes this class of ARU models. Lindberg et al. (1995) suggest an alternative representation, extend the analysis of the invariance property, and correct the characterization proof. More recently, de Palma and Kilani (2007) show that under independence, the only distribution for which the invariance property holds, within an ARU model, is the Gumbel distribution – a claim that was stated already in Strauss (1979). By way of applying and extending a result in Resnick and Roy (1990a), we give a direct and simple proof that our class of distributions is characterized by the invariance property in the independent case, without assuming additive noise. Although not immediately apparent, some of our arguments parallel those of Marley (1989), who analyzed a stochastic mechanism that generates a large class of random utility models closely related to GEV models. The mechanism is in the form of a horse race, and Marley's model gives predictions both of which horse will win and its recorded time.

Fosgerau et al. (2013) develop another general framework for additive random utility models. They obtain a characterization of choice probabilities in terms of partial derivatives of a value function that plays the same role as indirect utility in consumer theory (Roy's identity). This is a generalization of the Williams–Daly–Zachary theorem for additive random utilities (see McFadden, 1981). By contrast, we do not presume additivity and we obtain another result, namely that the choice probabilities are of a generalized "Luce form", that is, are generalizations of the ratio between a function of the

[1] Excellent textbooks on discrete choice random utility models are Ben-Akiva and Lerman (1985), Anderson et al. (1992) and Train (2003).

[2] Anderson and de Palma (1999) also analyze random cost mimimization but with additive i.i.d. reverse Gumbel noise terms. This leads to what they term the "reverse multinomial logit model" with the same choice probabilites as for random utility maximization where Gumbel distributed noise terms are subtracted rather than added to the negative of the observed costs.

[3] This concept of distributions that are stable under minimization (or maximization) should not be mixed up with the concept of min-stable (or max-stable) distributions in asymptotic extreme value theory, see, e.g., Coles (2001, Definition 3.1, p. 50).

[4] Already Weibull (1951) pointed out the general qualitative properties needed for survival functions to represent failure of the "weakest link of a chain", before suggesting the parametric form carrying his name as "the most simple function satisfying this condition" (op. cit. p. 293). One of the authors happens to be grandson of Waloddi Weibull (1887–1979).

[5] See Joe (1997, Chapter 6) for a general exposition of MEV distributions.