Random Utility Invariance Revisited

P O Lindberg*

Dept. Transp. Science, and Centre for Transp. Studies

KTH Royal Inst. Technology, Stockholm, Sweden

***Email: polin@kth.se**

1 Introduction

By *invariance* in a random utility model, we mean that the distribution of achieved utility is invariant across the alternatives chosen.

The study of Invariance in *Random Utility* (RU) *Models* originated with [1]. It gave a characterization of the joint distribution of the random terms in Additive RU (ARU) Models, necessary and sufficient for the distribution of achieved utility to be independent of which alternative attains the maximum. Later, [2] noted and corrected some errors in the proofs of [1], and gave an equivalent but somewhat different characterization. In this note, we investigate the consequences of the R&S (Robertson and Strauss) characterization.

In an RU Model, choice makers choose between alternatives in a finite *Universal Choice Set* $G =_{df} \{1,2,...,N\}$. With each alternative $i \in G$ is associated a random utility X_i, looked upon as the utility of alternative *i* for a randomly chosen choice maker. Hence, the probability of a randomly chosen choice maker to choose alternative *i* from a nonempty subset $I \subseteq G$ is postulated to be

$$p_I(i) =_{df} \Pr\{X_i > X_j, j \in I, j \neq i\}. \tag{1}$$

It usually assumed that the probability of ties is zero.

The utility distributions of the alternatives further typically depend on parameters, such as the costs of the alternatives. A typical case in question is ARU models, where the utility X_i, has "additive" structure:

$$X_i = v_i + U_i. \tag{2}$$

Here the utility X_i, is decomposed into the sum of a deterministic *population value*, v_i, assumed known to the analyst (such as the cost of alternative *i*) and an *individual value*, U_i, assumed unknown to the analyst, and hence considered as random when studying choices by the choice makers. ARU models have become work horses in many areas of probabilistic discrete choice, such as choice of mode of transport [3], choice of residential location [4], and many others

Let $\hat{X}_I$ denote the maximum achieved utility when choosing from the set *I*, i.e. $\hat{X}_I =_{df} \max_{i \in I} \{X_i\}$. Further let $\hat{X}_{i|I}$ denote the maximal utility conditioned upon alternative *i* being chosen out of *I*. When the choices are from the universal choice set *G*, we write just $\hat{X}$ and $\hat{X}_i$ for $\hat{X}_G$ and $\hat{X}_{i|G}$.

Let *F* be the joint *cdf* (cumulative distribution function) of $\mathbf{X} = (X_1, X_2, ..., X_N)$. (We use bold face to denote vectors.) We say that *F* (and also the RUM in question) has the *invariance property*, if the distribution of $\hat{X}_i$ is independent of *i*.

In an ARU model (2), the cdf of **X** depends on the vector **v** of population values, so let us denote it by $F_{\mathbf{v}}$, i.e. $F_{\mathbf{v}}(\mathbf{x}) =_{df} \Pr\{U_i + v_i \le x_i, i \in G\} = F_U(\mathbf{x} - \mathbf{v}) = F_{\mathbf{0}}(\mathbf{x} - \mathbf{v})$, where F_U is the cdf of $\mathbf{U} = (U_1, U_2, ..., U_N)$.

The invariance result proved (modulo the corrections by Lindberg et al. [2]) by [1] says that for an ARU Model, all the $F_{\mathbf{v}}$ have the invariance property, if and only if the cdf F_U has the form

$$F_U(\mathbf{x}) = \phi(H(e^{-\mathbf{x}}) =_{df} \phi(H(e^{-x_1}, e^{-x_2}, ..., e^{-x_N}), \qquad (3)$$

where *H* is a linearly homogenous function, given of course that $\phi(H(e^{-\mathbf{x}})$ is indeed a cdf. This class of distributions, which we will term *RS-distributions*, contains among others the GEV (*Generalized Extreme Value*) distributions, [4], for which $\varphi(x) = e^{-x}$, and a fortiori the *Multi Nomial Logit* (MNL) Model, which the GEV Model generalizes.

In this talk we relate results that give necessary and sufficient conditions on ϕ and *H* for F_U according to (3) to be a cdf. We also give non-GEV examples of ϕ and *H* giving invariant RU models.

2. Consequences of the RS formula for ϕ and *H*

2.1. *Normalization of setup.*

We will study what conditions ϕ and H in the RS formula (3) must fulfill for $F_{RS}(\mathbf{x}) =_{df} \phi(H(e^{-\mathbf{x}}))$ to be a cdf. The analysis to some extent parallels that [5]. First of all we will normalize F_{RS}, and are able to state:

Assumption. We will WLOG assume that F_{RS} has the representation

$$F_{RS}(\mathbf{x}) =_{df} \phi(-H(e^{-\mathbf{x}})), \tag{4}$$

where $\phi : (-\infty,0] \to [0,1]$ is non-decreasing with $\phi(-\infty) = 0$ and $\phi(0) = 1$, and H is positive and linearly homogenous, with $H(\mathbf{u}_i) = 1$. (Here $\mathbf{u}_i$ is the *i*-th coordinate unit vector.) We will say that such ϕ and H are RS-*appropriate* □

Remark. Note that the Assumption implies that all univariate marginals have the same cdf $\phi(-e^{-x})$. When we use the distribution (4) for the random term in an ARU model (2), the marginals gets shifted by the v_i and the marginal cdf's become $\phi(-e^{-(x-v_i)})$. Thus, in particular, all marginals have the same variance. □

Eq. (4) still is a bit complex to analyze. To this end we show**:**

Proposition. F_{RS}, according to (4), is a cdf if and only if the function $\overline{F}_{RS}$ defined by

$$\overline{F}_{RS}(\mathbf{x}) =_{df} \phi(-H(-\mathbf{x})) \tag{5}$$

is a cdf on $(-\infty,0)^N$. □

The derivations of conditions on ϕ and H for F_{RS} to be a cdf are rather intertwined. But for the sake of presentation, we will give result in an order not following the derivations.

As compared to the GEV setting, with $\phi(x) = e^{-x}$, a general ϕ of course gives a more complex situation. We therefore need to restrict he behavior of $\phi(x)$ when $x \uparrow 0$. We will say that ϕ is *RS-well-behaved* when its derivatives fulfill

$$\phi^{(n)}(s) \text{ is } o(\phi'(s)(-s)^{1-n}) \text{ ("little o") as } s \uparrow 0. \tag{6}$$

2.2. *The support of* ϕ.

The conditions on ϕ implies that its support must be either $(-\infty,0]$ or $[-a,0]$ for some $a>0$. Hence, the marginal distributions, by the remark to the Assumption, must have supports $(-\infty,\infty)$ or (up to translation) $[0,\infty)$.

2.3 *Necessary conditions on H.*

For *H*, we recover the classical *alternating sign* conditions, [4], [5]:

$$(-1)^{n-1} H_{i_1 i_2 \ldots i_m}(\mathbf{x}) \geq 0, \qquad (7)$$

(where $H_{i_1 i_2 \ldots i_n}(\mathbf{x})$ is the partial derivative of H w.r.t. $i_1 < i_2 < \ldots < i_m \leq n$).

2.4. *Sufficient conditions for ϕ and H.*

Concerning sufficient conditions it is natural to demand

$$\phi^{(n)}(x) \geq 0 \qquad (8)$$

Conditions (7) and (8) turn out to be sufficient for F_{RS} to be a cdf.

3. Special cases of ϕ.

Here we derive, using the sufficient conditions, some special cases of ϕ giving non-GEV RS-distributions

3.1. $\phi(x) = 1/(1-x)$ on $(-\infty,0]$.

This ϕ gives marginals that are *logistic*.

3.2. $\phi(x) = 1 + x$ on $[-1,0]$.

This ϕ gives marginals that are standard *exponential*.

3.2. $\phi(y) = 1 + y(1 - \ln(-y))$ on $[-1,0]$.

This ϕ gives marginals that are standard *Erlang-2*.

These examples are to our knowledge the first examples of non-GEV RS-distributions. Note in particular that the last two distributions give nonnegative error terms in ARU models. This might be of interest for the interpretation of ARU-models, where the possibility of arbitrarily large negative random utilities may seem counterintuitive.

References

[1] Robertson, C. A. and D. J. Strauss (1981), A Characterization Theorem for Random Utility Variables, *J. Math. Psychology* **23**, 259-263.

[2] Lindberg, P.O., E.A. Eriksson and L.-G. Mattsson (1995), Invariance of Achieved Utility in Random Utility Models, *Env. Plan.* **27** 121-142.

[3] Ben-Akiva, M. and S.R. Lerman (1985), *Discrete Choice Analysis*, The MIT Press, Cambridge, MA.

[4] McFadden, D. (1978), Modeling the Choice of Residential Location, in Karlqvist et al. (eds.), *Spatial Interaction Theory and Planning Models*, 75-96, North Holland, Amsterdam.

[5] Smith, T. E. (1988) A Choice Probability Characterization of Generalized Extreme Value Models, *Applied Mathematics and Computation* **14**, 35-62.