A DCE is an elicitation technique to obtain individual preferences using a series of hypothetical choices. DCEs are implemented using surveys that require a series of choices to be made between alternatives, where each alternative, although described using a set of common attributes (i.e. characteristics or traits) has different values or levels for these attributes. Alternatives are usually presented in pairs known as choice sets. For each choice set individuals are asked to select which profile they prefer based on the information presented. The response data from a DCE is analysed within the random utility model (RUM). In this model it is assumed that the different profiles presented to individuals generate a level of utility (or satisfaction) and that the individual selects the profile that yields the highest level of utility. The overall utility (U_i) for the *i*th alternative is divided into components that the analyst observes (V_i) and contributions that unfortunately we do not observe (ε_i). The relationship between the explainable and unexplainable components is assumed to be independent and additive so the level of utility for the *i*th alternative is described as:

$$U_i = V_i + \varepsilon_i$$

V_i is also known as the "representative component of utility" because it can be explained through the attributes that are observed in the DCE. The relative contribution of each attribute to the overall utility can be represented by a weight (i.e. a coefficient or parameter) that in its simplest form can take the form of a linear expression:

$$V_i = \beta_{oi} + \beta_{1i}X_{1i} + \beta_{2i}X_{2i} + \cdots + \beta_{ki}X_{ki}$$

where

β_{1i} is the parameter associated with the attribute X_{1i} and alternative i

β_{oi} is a parameter that is not associated to any of the observed attributes and represents an *alternative-specific constant* indicating on average the role of all the unobserved sources of utility

Different assumptions can be made about the error term ε_i but often for simplicity and a good starting point for the selection of choice model, it is assumed to be independent and with the exact same distribution (identically distributed) among alternatives. These sets of assumptions are known as IID (independent and identically distributed) [1].

Under RUM the individual evaluates each alternative represented as $U_j; j = 1, \ldots J$ alternatives, and compares $U_1, U_2, U_3, \ldots, U_j$ selecting the alternative with the highest utility, i.e. max (U_j). Therefore, the probability of selecting a specific alternative i compared to an alternative j can be expressed as:

$$Prob_i = Prob\left(U_i \geq U_j\right) \forall\, j \ \in j = 1, \ldots, J: i \neq j$$

In words, the probability of an individual choosing alternative i is equal to the probability that the utility of alternative i is greater than (or equal to) the utility associated with alternative j after evaluating each and every alternative in the choice set of $j = 1 \ldots i \ldots J$ alternatives.

This is equivalent to:

$$Prob_i = Prob\left[(V_i + \varepsilon_i) \ \geq \left(V_j + \varepsilon_j\right) \forall\, j \ \in j = 1, \ldots, J: i \neq j\right]$$

Also equivalent to:

$$Prob_i = Prob\left[\left(\varepsilon_j + \varepsilon_i\right) \leq \left(V_i + V_j\right) \forall j \in j = 1, \ldots, J : i \neq j\right]$$

In words, the probability of an individual choosing alternative i is equal to the probability that the difference in the unobserved sources of utility of alternative i compared to j is less than (or equal to) the difference in the observed sources of utility associated with alternative i compared to alternative j after evaluating each and every alternative in the choice set of $j = 1 \ldots i \ldots J$ alternatives.

This final expression indicates that to estimate the probability of an alternative i being selected compared to an alternative j, we need information on V_i and V_j (we can directly observe information on attributes and levels) and information on ε_i and ε_j (that we do not observe and in fact we have no idea what this looks like). Therefore, to estimate the probability of i being selected, we need to impose some structure for ε that helps us in identifying a practical choice model. The structure of the random component takes the form of a statistical distribution and a common distribution used in discrete choice analysis is the extreme value type 1 (EV1). The final selected choice model creates a relationship between the observed attributes, the unobserved attributes and the stated choice outcome. Under EV1 and IID assumptions we can derive the most widely used choice model known as multinomial logit (MNL). For a full derivation of the MNL under EV1 and IID, the reader is referred to Louviere and colleagues [2]. The predicted probabilities of an alternative i being selected from the complete set of alternatives $j = 1, \ldots, J$ in a MNL are given by:

$$Prob_i = \frac{expV_i}{\sum_{j=1}^{J} expV_i}; j = 1, \ldots, i, \ldots, J : i \neq j$$

1. Hensher, D.A., J.M. Rose, and W.H. Greene. Applied Choice Analysis. 2nd edition ed. 2015, Cambridge: Cambridge University Press.

2. Louviere, J.J., D.A. Hensher, J.D. Swait, and W. Adamowicz. Stated Choice Methods: Analysis and Application. 2000, Cambridge: Cambridge University Press.