Computers & Industrial Engineering 169 (2022) 108189

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Disutility Entropy in Multi-attribute Utility Analysis

Muhammed Sütçü

Industrial Engineering Department, Abdullah Gul University, Kayseri, Turkey

ARTICLE INFO

Keywords:
Multiattribute Utility Function
Cross Derivative
Shannon Entropy
Cumulative Residual Entropy
Maximum Entropy
Decision making with partial information

ABSTRACT

In this paper, we present an alternative formulation of utility entropy called disutility entropy. Previous entropy measurements in the literature use utility density function in maximum entropy formulations. Also, in most of the cases, the sign of the cross derivative of utility functions makes it impossible to apply utility entropy for more than one attribute cases. To simplify entropy measurement and relieve some burden of this task, in this paper, we present how to use multiattribute utility functions in utility entropy formulation. We show the applicability of our proposed approach and how to apply the disutility entropy approach with given constraints to single-attribute, bi-attribute, and multiattribute utility functions. Therefore, the usefulness and feasibility of the proposed method in multiattribute utility theory field is improved. We finally discuss and interpret the application of maximum disutility entropy through several examples to illustrate the new proposed approach.

1. Introduction

In decision-making practice, decision makers' preferences play a key role. The decision-maker, for example, may himself have very little knowledge of the relevant content; alternatively, he may lack the analytical capacity to evaluate the problems alone. Moreover, when there are dependencies between attributes, these preferences are difficult to reveal from the decision maker. It is unrealistic to expect him to fully list his preferences regarding the decision when decision maker has limited information about the decision to be made, Thus, it is essential to construct utility functions of decision maker when only partial preference information is available.

When the decision situation is deterministic, each decision alternative is described by a single prospect. The problem of choosing the best decision alternative is that of ordering the prospects present or assigning a value function over the attributes of each prospect. The optimal decision alternative is the one that has the largest value as determined by the value function or the highest order in the ranked list.

Under uncertainty, however, it is important to incorporate dependence among attributes into the decision situation when constructing multi-attribute utility functions. We need to assess utilities on various attribute combinations to capture dependency relationships between attributes. Unfortunately, these assessments increase exponentially due to the number of uncertainties present as the number of attributes increases. For instance, let's assume there are k attributes and all are mutually independent. In this case, only k assessments will be required and the assumption of mutual independence can simplify the assessment process. However, if those k attributes are dependent and assume they are binary, then multi-attribute utility function of k binary attributes requires the lower order assessments among attributes. So, the total number of assessments is $(2^{k}-1)$ and increasing exponentially.

Different approaches in decision analysis literature to make decisions under certainty and uncertainty. One of widely used method is to integrate information theory into decision theory and decision analysis. There are several information-theoretic measures that have been discussed in the literature of information theory since the beginning of Shannon's entropy (Shannon, 1948). Some of these are more general than others but all of them, like Shannon's entropy, were defined based on the probability density function. After Shannon's entropy explanation, Jaynes (Jaynes, 1957; Jaynes, 1957) introduces the principle of maximum entropy which maximizes Shannon's entropy subject to certain constraints representing the incomplete information. Rao defined maximum cumulative residual entropy (Rao, 2005) which is an extension of Jaynes's formulation by using cumulative distributions instead of density function. For variable X, mavwith given constraints are defined as shown in Table 1.

In the decision analysis literature, great importance is given to the solution of decision-making problems especially using entropy. Smith integrates entropy formulation into the context of decision making under uncertainty by partitioning the mathematical definition of a decision problem into two parts, a probability distribution and a convex set (Smith, 1974). Jessop shows how the concept of entropy can be used in multivariate decision problems and argues that analogies are drawn between information theory and between weights and probabilities

E-mail address: muhammed.sutcu@agu.edu.tr.

https://doi.org/10.1016/j.cie.2022.108189
Received 30 July 2019; Received in revised form 16 March 2022; Accepted 18 April 2022
Available online 22 April 2022
0360-8352/© 2022 Elsevier Ltd. All rights reserved.

Table 1
Comparison of Maximum Entropy and Maximum CRE.

Maximum Entropy	Maximum CRE
$f^{*}(x) = \text{argmax} - \int f(x_i)\log(f(x_i))$ s.t. $\int h_i(x)f(x_i) = \mu_i$ $\int f(x_i) = 1$ $f(x_i) \geqslant 0$	$H^{*}(x) = \text{argmax} - \int H(x)\log(H(x))$ s.t. $\int r_i(x)H(x) = \alpha_i$ $\int H(x) \leqslant E(X)$
Solution: $f^{*}(x) = e^{-\alpha_0 - \alpha_1 h_1(x) - \alpha_2 h_2(x) - \ldots - \alpha_n h_n(x) - 1}$	***Solution:*** $H^{*}(x) = e^{-1-\lambda_1 r_1(x) - \lambda_2 r_2(x) - \cdots \lambda_n r_n(x)}$

(Jessop, 1999). Berhold discusses use of distribution functions to represent utility functions which allow the decision maker to take advantage of their mathematical properties to develop useful results using density functions of utility (Berhold, 1973). Then, Barron examines the extent to which ranking information about weights can be used to determine the best alternative, or falling uniqueness determines an easily applied rule for choosing a 'best' alternative (Barron, 1992).

However, while solving the decision problems, it is hard to elicit full information. Therefore, there is a wide range of studies that focus on making decisions with partial information. In these studies, principle of maximum entropy is applied to utility theory, particularly to the estimation of utility functions when only partial preference information is available. Hadfi and Ito (Hadfi and Ito, 2012) extend maximum entropy utility principle to an asymptotic maximum entropy utility principle to reveal preference in a decision situation subject to a large estimation uncertainty with a small learning sample. Herfert and La Mura (Herfert and La Mura, 2004) use a non-parametric approach utility model and minimize sequential decision-theoretical cross-entropy with the given information to predict consumer preferences in the context of a simple consumer preference model. Abbas (Abbas, 2004) presents a maximum entropy method to find an optimal question-algorithm for eliciting von Neumann and Morgenstern utility values (Von Neumann and Morgenstern, 1947) and choosing the minimum number of questions required to elicit utility. In another paper, he applies a discrete form of maximum entropy to estimate a higher order joint probability distribution when the lower order assessments are known (Abbas, 2006). Abbas also shows how to apply maximum entropy to single attribute utility functions when only partial information is available and shows how to apply maximum entropy to multiattribute utility cases using value functions (Abbas, 2006). Pires et al. explore the potentialities of the generalized maximum entropy estimation of von Neumann and Morgenstern utility functions using only partial information about the decision maker's preferences by using utility elicitation methods (Pires et al., 2013). Kim and Ahn extend the maximum entropy utility methodology and propose a distance-based solution, so-called the centralized utility increments which are obtained by minimizing the expected quadratic distance to the set of vertices that varies upon partial preferences (Kim and Ahn, 2019). Bajgiran explores maximum entropy models with minimum elaborations over quantiles of uniform and moment-based maximum entropy models (Bajgiran et al., 2021). Maeda et al. address the problem of maximizing probabilistic entropy under a functional constraint induced by the available partial information (Maeda et al., 1993). Chen and Dai propose the principle of maximum entropy for uncertain variables to choose the one with maximum entropy from all uncertainty distributions satisfying given constraints (Chen and Dai, 2011). Also, there are interesting applications of maximum entropy in the literature that focus on probability density functions or utility density functions (Harju et al., 2019; Wang and Bier, 2013; Dionisio et al., 2008; Gao et al., 2015; Al-Omari, 2016; Chaji et al., 2018; Sutcu, 2020).

Abbas uses the density of single attribute utility functions and measure the uncertainty for only single-attributes due to the constraints of utility densities (Abbas, 2006). However, in real world applications, most of the decision situations contain more than one attribute. It is impossible to apply utility entropy directly to the multi-attribute case because of the sign of the cross-derivative of utility functions. Moreover, when the decision situation has more than one attribute, Abbas (Abbas, 2006) shows a way to solve by assigning a utility function over a value function. Thus, utility function has only one value function as an attribute and represents decision maker's risk attitude towards value with respect to utility function. Instead of using a utility function over a value function, it is impossible to use attributes directly on the utility function in case of multiple attributes, due to the negativity of the cross derivative of the utility function. Assume a two-attribute decision situation; the decision maker has a utility function $U(x,y) = 1 - e^{-x-y}$ and has a cross derivative of the utility function $\frac{\partial^2 U(x,y)}{\partial x \partial y} = -e^{-x-y}$ is negative. In this case, we cannot use the cross derivative of the utility function in log part of entropy expression. Montiel and Bickel propose new methods including maximum entropy for estimating partially specified probability distributions and comparatively discuss the difficulties of making approximations with limited information (Montiel and Bickel, 2013). Montiel and Bickel also discuss the assessment and the characterization of multilinear utility functions by providing a range to assessments when partial preference information is available (Montiel and Bickel, 2014). Finally, Ebrahimi explores several hazard rate function estimation problems based on the maximum entropy principle. Potential applications include developing several classes of the maximum entropy distributions which can be used to model distributions that produce different data that meet available information constraints on the hazard rate function (Ebrahimi, 2000).

This paper presents a normative method for measuring utility entropy and presents a functional form of maximum utility entropy distribution among attributes without using cross-derivative of utility functions. To simplify the decision situation and overcome the disadvantage of cross-derivative of utility functions, we focus on using multiattribute utility functions to calculate the utility entropy function instead of using utility density functions. Our proposed measure is; (i) non-negative, (ii) valid for both continuous and discrete cases, (iii) no need to estimate cross-derivative of utility functions (iv) no need to cross-derivative to be positive (v) easy to implement. These features help the decision maker and expert in the simple decision-making process. We emphasize the generality of the maximum utility entropy approach for constructing representative utility functions. It also provides flexibility to estimate utility functions without any assumption about a particular form of a utility function and integrates little information into the utility function while making decisions.

The remainder of this paper is organized as follows. Section 2 presents entropy definitions and maximum entropy formulation in probability and utility. Section 3 presents the disutility entropy measure with its definitions and formulations. Section 4 proposes the maximum disutility entropy utility approach based on partial preference information or given constraints. Section 5 covers the application of maximum disutility entropy to single, two-attribute and multiattribute utility functions and moment constraints. Section 6 includes the application of maximum disutility entropy with an application. Section 7 contains the concluding remarks and future work.

2. Review of previous work

In this section, we present several definitions of entropy both in probability and utility.

2.1. Entropy in probability

Entropy is a measure of uncertainty associated with a random variable. The concept was introduced by Claude E. Shannon in his paper "A Mathematical Theory of Communication" (Shannon, 1948). Shannon entropy of a discrete random variable is defined as $H(X) = -\sum_{i=1}^{n} p_i log p_i$. The continuous case of entropy with probability density function *f(x)* is defined by analogy to the discrete case as $H(f(x)) = -\int_a^b f(x)logf(x)dx$. After several decades of Shannon's entropy definition, Rao et al. (Rao et al., 2004) proposed a measure of entropy by using the cumulative distribution of a random variables. Rao et al. introduce cumulative residual entropy of a random variable in their paper as $CRE(X) = -\int_a^b P(|X|\rangle x)logP(|X|\rangle x)dx$. The main difference in this entropy measure is the cumulative joint distributions, especially survival functions, instead of density functions.

After Shannon's entropy explanation, Jaynes (Jaynes, 1957) introduces the principle of maximum entropy which maximizes Shannon's entropy subject to certain constraints representing the incomplete information. The maximum entropy probability distribution function for a continuous variable, *X*, having n outcomes, with given constraints is.

$$\begin{aligned} f(x)_{maxent} &= \text{argmax} -\int_a^b f(x)\log(f(x))dx \\ & s.t. \\ &\int_a^b h_i(x)f(x)dx = \mu_i \; i = 1,2,\cdots,n \\ &\int_a^b f(x)dx = 1 \\ & f(x) \geqslant 0 \end{aligned} \tag{1}$$

This probability distribution function has entropy which is the maximum when the entropy distribution has the form.

$$f^*(x) = e^{-\alpha_0-\alpha_1 h_1(x)-\alpha_2 h_2(x)-\cdots-\alpha_n h_n(x)-1} \tag{2}$$

where $f^*(x)$ is the maximum entropy solution, $h_i(x)$ is indicator function or moment constraint, α_i's are the Lagrange multipliers for the given constraints.

Rao shows (Rao, 2005) how maximum cumulative residual entropy is also calculated in a similar way to Maximum Entropy. Let *X* be a non-negative random variable, then maximum cumulative residual entropy is defined as.

$$\begin{aligned} H^* &= \text{argmax} -\int F(x)\log(F(x))dx \\ & s.t. \\ &\int r_i(x)F(x)dx = \alpha_i \; i = 1,2,\cdots,n \\ & x, F(x) \geqslant 0 \end{aligned} \tag{3}$$

and cumulative residual entropy distribution that maximizes the entropy with given constraints has the form as.

$$H^*(x) = e^{\sum_{i=1}^n \lambda_i r_i(x)} \tag{4}$$

where $H^*(x)$ is the maximum cumulative residual entropy solution, $r_i(x)$ is indicator function or moment constraint, λ_i's are the Lagrange multipliers for the given constraints.

2.2. Entropy in utility

Abbas introduced the utility entropy for the first time for single attribute utility functions (Abbas, 2002). He also introduced utility increment vector to calculate entropy of a discrete utility function. The entropy of a utility increment vector *(V)* defined as $V(\Delta u_1, \Delta u_2, \cdots, \Delta u_{n-1}) = -\sum_{i=1}^{n-1}\Delta u_i log\Delta u_i$ and the entropy of a utility function (*UE*) on a domain, [*a,b*] is $UE(x) = -\int_a^b u(x)ln(u(x))dx$. He then presents the maximum entropy utility (Abbas, 2006) by using the analogy between utility and probability and defines the formulation as.

$$\begin{aligned} &\text{argmax} -\int_a^b u(x)\log(u(x))dx \\ & s.t. \\ &\int_a^b h_i(x)u(x)dx = \mu_i \; i = 1,2,\cdots,n \\ &\int_a^b u(x)dx = 1 \\ & u(x) \geqslant 0 \end{aligned} \tag{5}$$

This formulation yields a utility density form $u_{maxent(x)} = e^{-\alpha_0-\alpha_1 h_1(x)-\cdots-\alpha_n h_n(x)-1}$ where $h_i(x)$ is a given preference constraint, μ_i s are given utility values from decision maker, and α_i s are the langrange multipliers for the given constraints. This form is a result of one attribute case. He presents the multiattribute case in the same paper by a utility function over a value function. So, a maximum entropy multiattribute utility function can be constructed with partial preference information using a utility assessment over the value function in the maximum entropy formulation.

3. Disutility entropy

In this section, we present new definitions and formulations that highlight the analogy between probability and utility. When we consider the analogous interpretation of entropy when applied to utility theory, we can easily see the similarities between them easily.

We propose an alternative measure of uncertainty in an attribute *X*, using the analogy between utility and probability, to incorporate the multiattribute utility function into the Rao et al.'s entropy definition and call it Disutility Entropy of attribute *X*. We begin our approach with the observation if a utility function is normalized from zero to one. It behaves mathematically as a cumulative probability distribution (both are non-decreasing and range from zero to one). The main objective of our study is to construct a new entropy measurement by using utility functions instead of cross derivatives or the density of the utility functions. We first introduce the discrete case of disutility entropy and in the following the continuous case.

3.1. Discrete Case: Disutility vector

A utility vector is a vector which ranks the utility values of the prospects from lowest to highest utility value (Abbas, 2002). Without loss of generality to von Neumann and Morgenstern axioms (Von Neumann and Morgenstern, 1947), we can define the disutility vector (*1-U*) like the complementary cumulative distribution function in probability. In the case of probability, the complementary cumulative distribution function explains the remaining lifetime of variable at a time. In the case of utility, the disutility vector shows the residual value of utility for a prospect if perfect information does not exist or we can say the difference between the best prospect and the rest of the prospects by assigning zero and one to least preferred and most preferred prospects, respectively.

Definition 1. *Discrete disutility vector* (*DUV*) *with n prospects is defined as.*

$$\begin{aligned} DUV &\cong (1-U_1, 1-U_2, \cdots, 1-U_{n-1}, 1-U_n) \\ DUV &\cong (\overline{U}_1, \overline{U}_2, \cdots, \overline{U}_{n-1}, \overline{U}_n) \\ &= (1, \overline{U}_2, \overline{U}_3, \cdots, \overline{U}_{n-1}, 0) \end{aligned} \tag{6}$$

We can easily see the analogy between cumulative discrete complementary cumulative distribution function and disutility vector, both are monotonically decreasing (non-increasing) and defined from one to zero. Utility functions can be defined as a measure of uncertainty due to a lack of perfect information. The disutility vector would be used in utility entropy to measure by the utility probability analogy.

Definition 2. *If we have an attribute X and its utility function is U, then the discrete disutility entropy* ($\overline{DUE}$) *can be formulated as.*

$$\overline{DUE}(\Delta\overline{U}_1(x),\Delta\overline{U}_2(x),\cdots,\Delta\overline{U}_n(x)) = -\sum_{i=1}^{n}\Delta\overline{U}_i(\mathrm{x})\log(\Delta\overline{U}_i(x)) \tag{7}$$

where $\overline{U}(x)$ is the disutility function. Let's assume we have 4 prospects and we assign least preferred prospect zero and most preferred prospect one. Other two prospects have 0.25 and 0.75 utility values. So, we have its disutility vector $\overline{DUE} = (1, 0.75, 0.25, 0)$. So, its entropy is calculated from disutility vector as $\overline{DUE} = (1, 0.75, 0.25, 0) \cong 0.2442$.

3.2. Continuous Case: Disutility function and differential entropy measure

In continuous case, the disutility vector is a disutility function, $\overline{U}(x)$, normalized from zero to one, and has the same mathematical properties as a complementary cumulative probability distribution function; both are non-increasing and range from zero to one.

Definition 3. *The entropy of disutility function can be formulated as.*

$$\overline{UE} = -\int_X \cdots \int \overline{U}(x_1,x_2,\cdots,x_n)\log(\overline{U}(x_1,x_2,\cdots,x_n))dx_1dx_2\cdots dx_n \tag{8}$$

where $\overline{U}(x_1, x_2, \cdots, x_n)$ is the disutility function and $x \in [a, b]$. The differential utility entropy has no lower bound and can be positive or negative. Nonetheless, there are several important properties of using disutility function in our approach; (i) our formulation has a lower bound for entropy of zero, (ii) differential entropy is only defined for distributions that has a density function which is difficult to elicit, but multiattribute utility function can be used directly without any assumption (iii) entropy of a discrete utility function is always positive, while differential entropy of continuous utility function may take any value positive or negative; but, entropy of disutility function is always nonnegative in both cases. Let's look at the following example. Assume that decision maker has an exponential utility function with risk aversion coefficient $\gamma = 1$. So, the disutility function is $\overline{U}(x) = e^{-x}$, and if $e^0 = 0$ and $x^* = 1$, then $\overline{UE} = 0,264$.

Proposition 1. *For any mutually independent attributes X and Y,*

$$max(\overline{UE}(U(x)),\overline{UE}(U(y))\le\overline{UE}(U(x),U(y)) \tag{9}$$

Proposition 2. $\overline{UE}(U(x)) \geq 0$*and equality holds if and only if utility of attributes,* $U(x_i) = 1$ *and.*$U(x_1) = \cdots = U(x_{i-1}) = U(x_{i+1}) = \cdots = U(x_n) = 0$

4. Maximum disutility entropy

In this section, we discuss the functional form of maximum disutility entropy when only partial preference information about the decision maker is available. Here we begin by eliciting preference information for the prospects from the decision maker. If we have only partial information and would like to find an estimated utility function of decision maker, we need to solve maximum disutility entropy function. We discuss the functional form of maximum disutility entropy function and then we show how to apply the maximum entropy disutility approach through several examples.

4.1. Maximum disutility entropy expression

The maximum disutility entropy formulation with constraints is.

$$\begin{gathered}\operatorname{argmax} -\int_X \cdots \int \overline{U}(x_1,x_2,\cdots,x_n)\log(\overline{U}(x_1,x_2,\cdots,x_n))dx_1dx_2\cdots dx_n \\ s.t. \\ \int \cdots_X \int r_i(x_1,x_2,\cdots,x_n)\overline{U}(x_1,x_2,\cdots,x_n)dx_1dx_2\cdots dx_n \quad i=1,2,\cdots,m\end{gathered} \tag{10}$$

where X is the domain of prospects, $r_i(x)$ is a given preference constraint, $\overline{U}$ s are given disutility values of decision maker.

In disutility entropy formulation, we don't need the normalization constraint and non-negativity constraint. Formulation (10) yields a disutility function of the form $\overline{U}_{maxent} = e^{-\sum_{i=1}^{n}\lambda_i r_i}$ where $\overline{U}_{maxent}$ is the maximum entropy disutility solution, $r_i(x)$ is a given preference constraint, λ_i s are the Lagrange multipliers for the given constraints.

Given a two-attribute utility function which is normalized zero to one, then cross moment of a disutility function is the expected utility of two attributes, and if one of the attributes is its maximum, then for the other attribute, it is the normalized expected utility for marginal distribution of an attribute. For instance, if $r(x, y) = xy$ then cross utility moment of.$\overline{U}(x, y)$

$$\int_{y^o}^{y^*}\int_{x^o}^{x^*} r(x,y)\overline{U}(x,y)dxdy = \int_{y^o}^{y^*}\int_{x^o}^{x^*} xy\overline{U}(x,y)dxdy = \mu_{xy} \tag{11}$$

is the expected utility for two attributes, and μ_{xy} is given by decision maker by assigning p to the uniform two attribute lottery Fig. 1.

$$\int_{y^o}^{y^*}\int_{x^o}^{x^*} xy\overline{U}(x,y)dxdy = \mu_{xy} = p_{xy}$$

If one of the attributes is its maximum $y^* = 1$, then.

$$\int_{y^o}^{y^*}\int_{x^o}^{x^*} r(x,y)\overline{U}(x,y)dxdy = \int_{y^o}^{y^*}\int_{x^o}^{x^*} x\overline{U}(x,y)dxdy = \mu_{x|y^*} \tag{12}$$

is the expected utility for attribute X, and $\mu_{x|y^*}$ is given by decision maker by assigning p to the uniform two attribute lottery where $\int_{y^o}^{y^*}\int_{x^o}^{x^*} x\overline{U}(x,y^*)dxdy = \mu_x = p_{x|y^*}$

Fig. 2 decision maker assigns utility values using the lotteries as defined above. Partial preference information means that new decisions are made for a decision maker based on some utility values are elicited from the decision maker and previous decisions made by decision maker.

4.2. Constraints and moments of disutility functions

In this section, our aim is to present moments of disutility functions and derive a formulation to represent moments using the expected utility of joint distribution and expected utilities of its marginal distributions. We also show an interpretation for first and cross moments of utility function, how to use them in practice, and how to receive these measures from the decision maker.

4.2.1. Definitions and explanations which are used in analysis

First, we assume that the attributes are preferentially independent.

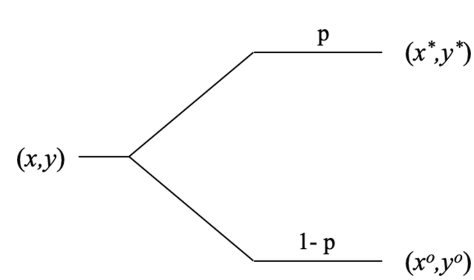

Fig. 1. Two attribute lottery to assign..μ_{xy}

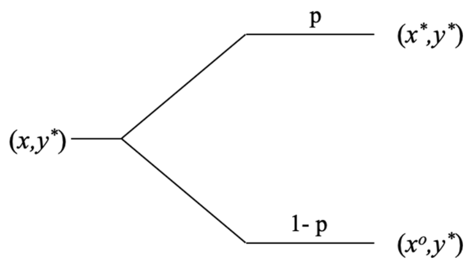

Fig. 2. Marginal utility lottery to assign μ_x when..$y = y^*$

So, for any x_1, x_2, y_1, y_2, we define.

$$if x_1 > x_2, then (x_1, y) \rangle (x_2, y) \forall y \in [y^o, y^*] \tag{13}$$

and.

$$if y_1 > y_2, then (x, y_1) \rangle (x, y_2) \forall x \in [x^o, x^*] \tag{14}$$

where a two attribute prospect or alternative is represented here as (x,y) and x^o, y^o shows the minimum values, and x^*, y^* shows the maximum values of attributes × and y respectively.

In this analysis, we also normalized the two-attribute utility function, U(x,y) such that.

$$0 \le U(x, y) \le 1 \forall x \in [x^o, x^*], \forall y \in [y^o, y^*] \tag{15}$$

If we don't normalize the utility function and its moments, then they are just arbitrary quantities and are meaningless without any arrangement. So, we need to scale them into a range. Since, a linear transformation does not affect the order of the alternatives, in our analysis; we scale them 0 to 1. We assign zero to the least preferred prospect, and one to the most preferred prospect, and normalize the remaining prospects between these two.

So, we can say from (5).

$U(x^o, y^o) = 0$ and $U(x^*, y^*) = 1$ (16).

4.2.2. *Finding the constraints of disutility entropy*

The moments of a disutility functions are defined as.

$$\iint f(x, y) \overline{U}(x, y) dxdy \tag{17}$$

If we rearrange (17), then.

$$= \iint f(x, y)[1 - U(x, y)]dxdy$$

$$= \iint f(x, y)dxdy - \iint f(x, y)U(x, y)dxdy \tag{18}$$

So, general functional form of moments of disutility functions is equal to.

$$\overline{UE}_{xy} = \iint f(x, y)\overline{U}(x, y)dxdy = 1 - EU_{xy} \tag{19}$$

We start calculating the first moments of disutility function and then the cross moment of the disutility function.

4.2.3. *First moments of disutility function*

First of all, we calculate the first moment when $f(x,y) = x$. (For proof and detailed calculation see Appendix-A).

$$\iint f(x, y)\overline{U}(x, y)dxdy = 1 - EU_{xy} \tag{20}$$

We know that expected utility of a function is calculated as

$$EU_{xy} = \iint f(x, y)U(x, y)dxdy, then \overline{UE}_x = 1 - \int_0^1 \int_0^1 xU(x, y)dxdy \tag{21}$$

$$\overline{UE}_x = 1 - [EU_{xy} - (1 - k_x)EU_{y|x^*} - k_x] \tag{22}$$

So, the first moment of disutility function when $f(x, y) = x$ is.

$$\iint x\overline{U}(x, y)dxdy = 1 + (1 - k_x)EU_{y|x^*} + k_x - EU_{xy} \tag{23}$$

where $k_x = U(x^*, y^o)$, which shows the corner point when × is its maximum, and y is its minimum value, and if attribute Y is a dominant attribute (when one of the attributes is its minimum, then the multi-attribute utility function is also its minimum) then $k_x = 0$.

So, in the attribute dominance case, the first moment simplifies to.

$$\iint x\overline{U}(x, y)dxdy = 1 + EU_{y|x^*} - EU_{xy} \tag{24}$$

We now calculate first moment when $f(x,y) = y$. (For proof and detailed calculation see Appendix-B). We do the same calculation but that time we use $f(x, y) = y$ instead of x.

$$= 1 - \int_0^1 \int_0^1 yU(x, y)dxdy \tag{25}$$

$$= 1 - [EU_{xy} - (1 - k_y)EU_{x|y^*} - k_y] \tag{26}$$

So, the first moment of the disutility function when $f(x, y) = y$ is.

$$\iint y\overline{U}(x, y)dxdy = 1 + (1 - k_y)EU_{x|y^*} + k_y - EU_{xy} \tag{27}$$

where$k_y = U(x^o, y^*)$, which shows the corner point when y is its maximum, and × is its minimum value. Here, $EU_{x|y*}$ is again the decision makers' indifference probability for an equivalent two outcome lottery that has only the best outcome EU_{xy}, and the worst outcome lottery EU_{xy}. We don't need to normalize EU_{xy} because we already assumed that it is normalized from zero to one. Moreover, $EU_{x|y^*}$ is the indifference probability for equivalent two outcome lottery of the marginal distributions when the attribute Y is its maximum. If attribute X is a dominant attribute (when one of the attributes is its minimum, then the multi-attribute utility function is also its minimum) then$k_y = 0$.

4.2.4. *Cross moments of disutility function*

In this part, we calculate the cross moments of disutility functions when$f(x, y) = xy$. (For proof and detailed calculation see Appendix-C). The cross moment of a bi-attribute disutility function is.

$$\iint xy\overline{U}(x, y)dxdy = 1 + EU_{xy} - (1 - k_x)EU_y - (1 - k_y)EU_x - (k_x + k_y) \tag{28}$$

So, here $EU_{xy} = \iint_{x,y \in [0,1]} f(x, y)U(x, y)dxdy$ is the expected utility of$F(x, y)$, EU_x and EU_y are the normalized expected utility for the marginal distributions of variables X and Y respectively. If we interpret the equation (28), it is the sum of one and expected utility of two variables minus the expected utilities of each marginal variables. To explain the meaning better, we first show the two-outcome lottery for two attribute case and then for each marginal attribute.

First, we can find the EU_{xy} by using the two-outcome lottery Fig. 3.

And then, the marginal expected utilities of each marginal can be found by Fig. 4.

Therefore, we rewrite the equations (24), (27), and (2) as.

$$\iint x\overline{U}(x, y)dxdy = 1 + (1 - k_x)p_y + k_x - p_{xy} \tag{29}$$

$$\iint y\overline{U}(x, y)dxdy = 1 + (1 - k_y)p_x + k_y - p_{xy} \tag{30}$$

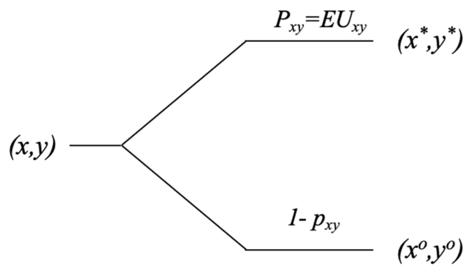

Fig. 3. Two outcome lottery for Expected utility for two attributes.

$$\iint xy\overline{U}(x,y)dxdy = 1 + p_{xy} - (1-k_x)p_y - (1-k_y)p_x - (k_x + k_y) \tag{31}$$

Moreover, we need to elicit the corner points from the decision maker. We again use the lotteries to find the decision maker's indifference probability assessments for the following two outcome lotteries Fig. 5.

4.2.5. Interpretation of the preference constraints used in the model

In the maximum disutility entropy formulations, the preference constraints are used in a constraint function as $\int_a^b r_i(x)\overline{U}(x)dx = \mu_i$. Here, $r_i(x)$ is a monotonically increasing function on the domain of the alternatives. As we know from the utility theory, if there are two alternatives, there exists an equivalent two-outcome lottery, with probability p of getting one of the alternatives and getting one another alternative with probability (1-p). In this lottery, the decision maker is indifferent between the two alternatives as shown in Fig. 6. If the utility function of the decision maker is normalized in a range [0,1], then the probability of the better alternative is equal to expected utility value of this lottery.

In the maximum disutility entropy case, the preference constraint can be interpreted as one minus decision maker's indifference probability for a given two-outcome lottery. So, the preference constraint is $\int_a^b r_i(x)\overline{U}(x)dx = 1 - p$ and can be interpreted as the decision maker's indifference probability for its least preferred outcome. Therefore, if we can learn the indifference probability of the decision maker by using lottery method, then the maximum disutility entropy method determines an exponential utility function of the decision maker.

5. Maximum disutility entropy for single and Multi-Attribute cases

In this section, we discuss the maximum disutility entropy for univariate, bivariate, and tri-variate cases. We first discuss the single attribute case, then two-attribute case with different scenarios. Finally, a three-attribute case will be discussed in this section.

5.1. Maximum disutility entropy for single attribute cases

The first definition is a single attribute case of maximum disutility entropy. The maximum disutility entropy for single attribute case is defined as.

$$\begin{aligned} &\max \quad -\int_{x^o}^{x^*} \overline{U}(x)\log(\overline{U}(x))dx \\ &\qquad\qquad s.t. \\ &\int_{x^o}^{x^*} r_i(x)\overline{U}(x)dx = \mu_i \qquad i = 1, 2, \cdots, n \end{aligned} \tag{32}$$

Formulation (32) yields a disutility function of the form $\overline{U}_{maxent} = e^{-\sum_{i=1}^{n}\lambda_i r_i(x)}$ where $\overline{U}_{maxent}$ is the maximum entropy disutility solution, $r_i(x)$ is a given preference constraint,λ_i's are the Lagrange multipliers that are calculated to satisfy the given constraints, andμ_i's are the given constants by the decision maker.

Example 5.1. *Consider the first and the second moments of the utility function is known. The maximum entropy disutility formulation with given constraints is.*

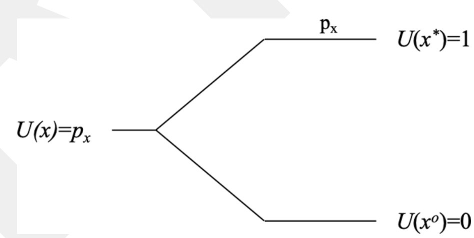

Fig. 6. Two outcome lottery for corner point assessments.

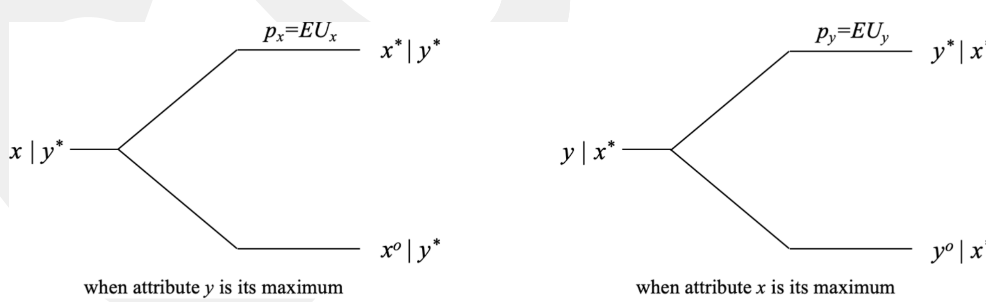

Fig. 4. Two outcome lottery for Expected utility for each marginal.

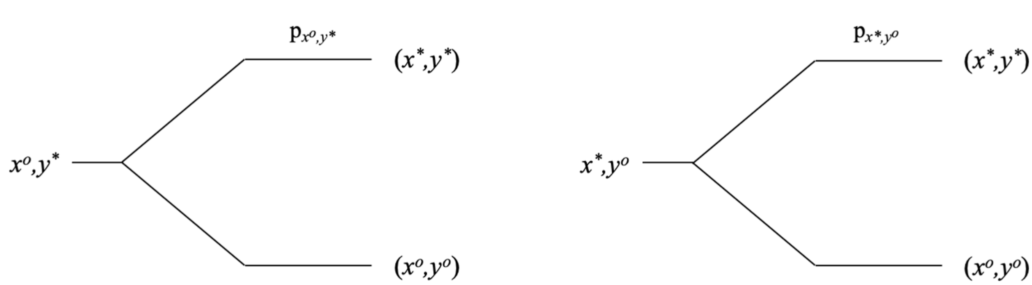

Fig. 5. Two outcome lottery for corner point assessments.

$$\max \quad -\int_{x^o}^{x^*} \overline{U}(x)\log(\overline{U}(x))dx$$

$$s.t. \qquad (33)$$

$$\int_{x^o}^{x^*} \overline{U}(x)dx = (x^* - x^o)(1-p) + x^o$$

$$\int_{x^o}^{x^*} x\overline{U}(x)dx = \frac{\left[(x^*)^2 - (x^o)^2\right](1-p_x) + (x^o)^2}{2}$$

Here, the x^o and x^* values show the maximum and minimum values of attribute ×, respectively. Also, p_x is the decision makers' indifference probability of getting the best prospect in an equivalent two-outcome lottery that only has two prospects; best and worst. Fig. 7 shows the two-outcome lottery.

The maximum disutility entropy solution with given constraints has the form.

$$\overline{U}(x) = e^{-\lambda_1 - \lambda_2 x} \qquad (34)$$

5.2. Maximum disutility entropy for two attribute cases

In this part, we define two attribute maximum disutility entropies. The two-attribute disutility function is defined as.

$$\overline{U}(x,y) = U(x^*,y^*) - U(x,y^*) - U(x^*,y) + U(x,y) \qquad (35)$$

The maximum disutility entropy for two attribute case is defined as.

$$\max \quad -\int_{y^o}^{y^*}\int_{x^o}^{x^*} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy$$

$$s.t. \qquad (36)$$

$$\int_{y^o}^{y^*}\int_{x^o}^{x^*} r_i(x,y)\overline{U}(x,y)dxdy = \mu_i \qquad i = 1,2,\cdots,n$$

The maximum disutility entropy has the form $\overline{U}_{maxent} = e^{-\sum_{i=1}^{n}\lambda_i r_i(x,y)}$.

In general, the constraints of a two-attribute disutility function are defined as.

$$\begin{aligned} &= \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)\overline{U}(x,y)dxdy \\ &= \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)[U(x^*,y^*) - U(x,y^*) - U(x^*,y) + U(x,y)]dxdy \\ &= \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)U(x^*,y^*)dxdy - \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)U(x,y^*)dxdy \\ &\quad - \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)U(x^*,y)dxdy + \int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y)U(x,y)dxdy \\ &= 1 + EU_{xy} - (1-k_x)EU_y^n - (1-k_y)EU_x^n - (k_x+k_y) \end{aligned} \qquad (37)$$

where EU_{xy} is the expected utility of bivariate distribution of attributes X and Y, $EU_x^n$is the normalized expected utility for marginal distribution of attribute X, EU_y^n is the normalized expected utility for marginal distribution of attribute Y, $k_x = U(x^*,y^o)$ and $k_y = U(x^o,y^*)$ are the corner points that one attribute has its maximum value and the other attribute has its minimum value.

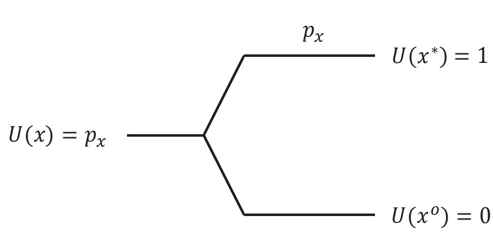

Fig. 7. Best and Worst Outcome Lottery.

Example 5.2. *Let's assume a decision situation has two attributes X and Y. We want to find a representative utility function with given constraints $r_1 = x$, and $r_2 = y$. The maximum disutility entropy formulation for two attribute case is.*

$$\begin{aligned} &\underset{}{\text{argmax}} - \iint_{x,y\in[0,1]} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy \\ &\qquad s.t. \\ &\iint x\overline{U}(x,y)dxdy = \mu_1 \\ &\iint y\overline{U}(x,y)dxdy = \mu_2 \end{aligned} \qquad (38)$$

We can calculate the above equation by Lagrange multipliers as.

$$\begin{aligned} L(\overline{U}(x,y)) = -\iint_{x,y\in[0,1]} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy - \lambda_1\left\{\iint x\overline{U}(x,y)dxdy - \mu_1\right\} \\ -\lambda_2\left\{\iint y\overline{U}(x,y)dxdy - \mu_2\right\} \end{aligned} \qquad (39)$$

Taking the partial derivative and equating it to zero gives.

$$\frac{\partial^2 \overline{U}(x,y)}{\partial x \partial y} = -\log(\overline{U}(x,y)) - \lambda_1 x - \lambda_2 y = 0 \qquad (40)$$

When we rearrange the above equation, it gives.

$$\begin{aligned} \overline{U}(x,y) &= e^{-\lambda_1 x - \lambda_2 y} \\ &= 1 - U(x,y) \\ &= \overline{U}(x,y) \end{aligned} \qquad (41)$$

which is independent.

Also, if we have an additional constraint $\iint xyu_{xy}(x,y)dxdy = \mu_{12}$ and put it into the maximum entropy formulation, then.

$$\begin{aligned} L(\overline{U}(x,y)) = -\iint_{x,y\in[0,1]} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy - \lambda_1\left\{\iint_{x,y} x\overline{U}(x,y)dxdy - \mu_1\right\} \\ -\lambda_2\left\{\iint_{x,y} y\overline{U}(x,y)dxdy - \mu_2\right\} - \lambda_{12}\left\{\iint_{x,y} xy\overline{U}_{xy}(x,y)dxdy - \mu_{12}\right\} \end{aligned} \qquad (42)$$

The partial derivative gives expression below when we equate it to zero.

$$\frac{\partial^2 \overline{U}(x,y)}{\partial x \partial y} = -\log(\overline{U}(x,y)) - \lambda_1 x - \lambda_2 y - \lambda_{12} xy = 0 \qquad (43)$$

When we rearrange the above equation, it gives.

$$\begin{aligned} \overline{U}(x,y) &= e^{-\lambda_1 x - \lambda_2 y - \lambda_{12} xy} \\ &= e^{-\lambda_1 x} e^{-\lambda_2 y} e^{-\lambda_{12} xy} \end{aligned} \qquad (44)$$

The above expression for multi-attribute disutility function incorporates utility dependence. This functional form is an extension of the product of two exponential marginal utility functions and it includes the utility dependence among them with the third constraint which is the cross-moment constraint. When the Lagrange multiplier λ_{12} in (44) equals to zero, then the solution is same as the product of two exponential marginal utility function which are mutually utility independent.

As a result, if we only assess a utility function with mutual utility independence among attributes, we need the first moments of marginal utility functions of two attributes which are the expected utility of each attribute. If we want to incorporate utility dependence between attributes, then we can assess the utility cross moment as an additional constraint which is the expected utility of two attributes.

In all of the analysis, we assume that the utility functions are normalized from one (worst outcome) to one (best outcome). Here, in the constraint formulation EU_{xy} is the decision maker's indifference probability of an equivalent two-outcome lottery that has only two outcomes; the best (x^*, y^*) and the worst(x^o, y^o) . Moreover, EU_x and EU_{xy} are the decision maker's indifference probability for equivalent two-outcome lotteries of the marginal distributions of X and Y respectively. Finally, $k_x = U(x^*, y^o)$ and $k_y = U(x^o, y^*)$ are the corner points that one attribute has its maximum value and the other attribute has its minimum value. Now, we show the two-outcome lotteries of EU_{xy}, EU_x, EU_y and corner points k_x and k_y that we need to elicit from the decision maker.

Fig. 8 shows the equivalent two-outcome lottery that has only two outcomes; the best (x^*, y^*) and the worst(x^o, y^o). Here, the probability p_{xy} is the expected utility of bivariate distribution of attributes X and Y.

Finally, Fig. 9 shows the equivalent two outcome lotteries of corner points. The equivalent two-outcome lottery that has only two outcomes; the best (x^*, y^*) and the worst(x^o, y^o). In that case we ask the decision maker what is the probability that you are indifferent between a two-outcome lottery that has only the best (x^*, y^*) and the worst (x^o, y^o) outcomes and an outcome that one attribute has its maximum value and the other attribute has its minimum value.

5.3. The functional forms of two attribute maximum disutility functions

We discuss the functional forms of two attribute maximum disutility functions in two different cases. First, we discuss the case that no further information about the multiattribute disutility functions is available expect the bounds on each attribute and the constraints that include information about only one attribute,

$f_1(x, y) = x$ and $f_2(x, y) = y$ (45).

In the second case, in addition to the constraint $f_1(x, y)$ and$f_2(x, y)$, a third constraint which includes information from both attributes, cross-moments are incorporated into the maximum entropy disutility formulation.

$$f_3(x, y) = xy \tag{46}$$

5.3.1. First Case:

Here, first we consider the case where only $f_1(x, y) = x$ and $f_2(x, y) = y$ moment constraints are known. In this case, the maximum disutility entropy formulation is.

$$\begin{aligned} &\operatorname{argmax} - \int_{y^o}^{y^*}\int_{x^o}^{x^*} \overline{U}(x, y)\log(\overline{U}(x, y))dxdy \\ &\qquad s.t. \\ &\int_{y^o}^{y^*}\int_{x^o}^{x^*} x\overline{U}(x, y)dxdy = \frac{1 + EU_{xy} - (1 - k_y)EU_x - k_y}{2} \\ &\int_{y^o}^{y^*}\int_{x^o}^{x^*} y\overline{U}(x, y)dxdy = \frac{1 + EU_{xy} - (1 - k_x)EU_y - k_x}{2} \end{aligned} \tag{47}$$

and this formulation yields a disutility function of the form$\overline{U}_{maxent} = e^{-\lambda_0 - \lambda_1 x - \lambda_2 y}$.

$$\begin{aligned} \overline{U}(x, y) &= e^{-\lambda_0} e^{-\lambda_1 x} e^{-\lambda_2 y} \\ &= \overline{U}_x(x)\overline{U}_y(y) \end{aligned} \tag{48}$$

The solution to this case is well-known in maximum entropy literature and yields a two-attribute disutility function equal to the product of the marginal survival functions.

If we derive the two-attribute disutility function, then we have the joint utility density function.

$$u(x, y) = u_x(x)u_y(y) \tag{49}$$

The first case shows that the maximum entropy disutility function assumes utility independence between attributes if the constraints include information from only one attribute. If more information is available about the attributes and utility function, such as cross moments, it can also be added to maximum entropy formulation. In the second case, we show the maximum entropy disutility formulation if we additionally know the cross moments of disutility function.

5.3.2. Second Case:

Now, we consider the case we additionally know the constraint$f_3(x, y) = xy$. In this case, we attach the third constraint.

$$\int_{y^o}^{y^*}\int_{x^o}^{x^*} xy\overline{U}(x, y)dxdy = 1 + EU_{xy} - (1 - k_x)EU_y - (1 - k_y)EU_x - (k_x + k_y) \tag{50}$$

to the maximum disutility formulation and we have.

$$\begin{aligned} &\operatorname{argmax} - \int_{y^o}^{y^*}\int_{x^o}^{x^*} \overline{U}(x, y)\log(\overline{U}(x, y))dxdy \\ &\qquad s.t. \\ &\int_{y^o}^{y^*}\int_{x^o}^{x^*} x\overline{U}(x, y)dxdy = \frac{1 + EU_{xy} - (1 - k_y)EU_x - k_y}{2} \\ &\int_{y^o}^{y^*}\int_{x^o}^{x^*} y\overline{U}(x, y)dxdy = \frac{1 + EU_{xy} - (1 - k_x)EU_y - k_x}{2} \\ &\int_{y^o}^{y^*}\int_{x^o}^{x^*} xy\overline{U}(x, y)dxdy = 1 + EU_{xy} - (1 - k_x)EU_y - (1 - k_y)EU_x - (k_x + k_y) \end{aligned} \tag{51}$$

and this formulation yields a disutility function of the form$\overline{U}_{maxent} = e^{-\lambda_0 - \lambda_1 x - \lambda_2 y - \lambda_3 xy}$.

If we derive the two-attribute disutility function, then we have the joint utility density function.

$$u(x, y) = \left[\lambda_1\lambda_2 + \lambda_1\lambda_3 x - \lambda_3 + \lambda_2\lambda_3 y + (\lambda_3)^2 xy\right]e^{-\lambda_1 x - \lambda_2 y - \lambda_3 xy} \tag{52}$$

shows that there is a relationship between attributes and they are not independent.

5.4. Disutility entropy for three attribute cases

In this section, we calculate the maximum disutility entropy for three attribute cases. First of all, we write down the three-attribute disutility function as.

$$\begin{aligned} \overline{U}(x, y, z) = &U(x^*, y^*, z^*) - U(x^*, y^*, z) - U(x^*, y, z^*) - U(x, y^*, z^*) \\ &+ U(x^*, y, z) + U(x, y^*, z) + U(x, y, z^*) + U(x, y, z) \end{aligned} \tag{53}$$

Then, we consider a three-attribute case and solve the three-attribute case by using maximum disutility entropy approach by using our proposed approach. Here, we assume that we elicit the information from the decision maker and use this information as our constraints in the model by using the lottery method as below decision tree and ask questions to decision maker to elicit his preferences over best and worst alternatives respect to other alternatives using the decision tree below (Fig. 10).

In general, the constraints of a three-attribute disutility function are

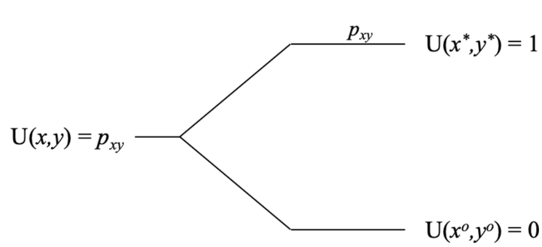

Fig. 8. Equivalent two outcome lotteries of marginal distributions.

GCRIS

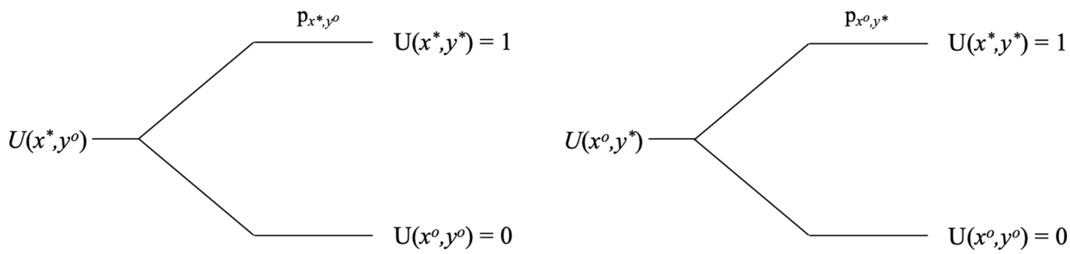

Fig. 9. Two outcome lotteries of corner points.

defined as.

$$
\begin{aligned}
&= \int_{z^o}^{z^*}\int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y,z)\overline{U}(x,y,z)dxdydz \\
&= \int_{z^o}^{z^*}\int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y,z)[U(x^*,y^*,z^*) - U(x^*,y^*,z) - U(x^*,y,z^*) - U(x,y^*,z^*)]dxdydz \\
&+ \int_{z^o}^{z^*}\int_{y^o}^{y^*}\int_{x^o}^{x^*} f(x,y,z)[U(x^*,y,z) + U(x,y^*,z) + U(x,y,z^*) + U(x,y,z)]dxdydz \\
&= 1 - EU_{xyz} + (1-k_x)EU^n_{yz} + (1-k_y)EU^n_{xz} + (1-k_z)EU^n_{xy} \\
&-(1-k_{yz})EU^n_x - (1-k_{xz})EU^n_y - (1-k_{xy})EU^n_z + (k_x+k_y+k_z) - (k_{xy}+k_{xz}+k_{yz})
\end{aligned} \tag{54}
$$

After calculating the constraints of the model based on the information elicited from the decision maker, we write down the maximum disutility entropy model as.

$$
\begin{aligned}
&\max \quad -\int_{z^o}^{z^*}\int_{y^o}^{y^*}\int_{x^o}^{x^*} \overline{U}(x,y,z)\log(\overline{U}(x,y,z))dxdydz \\
&\qquad s.t. \\
&\int_{z^o}^{z^*}\int_{y^o}^{y^*}\int_{x^o}^{x^*} f_i(x,y,z)\overline{U}(x,y,z)dxdydz = \mu_i \qquad i = 1,2,\cdots,n
\end{aligned} \tag{55}
$$

Thus, the maximum disutility entropy has the form $\overline{U}_{maxent} = e^{-\sum_{i=1}^{n}\lambda_i f_i(x,y,z)}$. The model can be extended to multiattribute cases in the same analogy. Here, the advantage of our model is to use all kind of utility functions in our model due to the advantage of using cumulative utility functions instead of using utility density function as the analogy between probability functions and utility functions. Therefore, in our approach we use multiattribute utility functions in entropy expression which are always nonnegative which helps to apply entropy principle to multiattribute cases without any assumption.

We now compare the maximum entropy utility and maximum disutility entropy principle by a two-attribute utility function.

Maximum entropy disutility	Maximum entropy utility
$\text{argmax} - \iint_{x,y\in[0,1]} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy$	$\text{argmax} - \iint_{x,y\in[0,1]} u(x,y)\log(u(x,y))dxdy$
$s.t.$	$s.t.$
$\iint x\overline{U}(x,y)dxdy = \mu_1$	$\iint xu(x,y)dxdy = \mu_1$
$\iint y\overline{U}(x,y)dxdy = \mu_2$	$\iint yu(x,y)dxdy = \mu_2$
$\iint xy\overline{U}_{xy}(x,y)dxdy = \mu_{12}$	$\iint xyu_{xy}(x,y)dxdy = \mu_{12}$
	$\iint u(x,y)dxdy = 1 \quad u(x,y)\geqslant 0$
by Lagrange multipliers.	by Lagrange multipliers.
$L(\overline{U}(x,y)) = -\iint_{x,y\in[0,1]} \overline{U}(x,y)\log(\overline{U}(x,y))dxdy - \lambda_1\left\{\iint x\overline{U}(x,y)dxdy - \mu_1\right\} -\lambda_2\left\{\iint y\overline{U}(x,y)dxdy - \mu_2\right\} - \lambda_{12}\left\{\iint xy\overline{U}_{xy}(x,y)dxdy - \mu_{12}\right\}$	$L(u(x,y)) = -\iint_{x,y\in[0,1]} u(x,y)\log(u(x,y))dxdy - \lambda_1\left\{\iint xu(x,y)dxdy - \mu_1\right\} -\lambda_2\left\{\iint yu(x,y)dxdy - \mu_2\right\} - \lambda_{12}\left\{\iint xyu(x,y)dxdy - \mu_{12}\right\} - \lambda_0\left\{\iint u(x,y)dxdy - 1\right\}$
The partial derivative gives expression below when we equate it to zero.	The partial derivative gives expression below when we equate it to zero.
$\frac{\partial^2 L(\overline{U}(x,y))}{\partial x \partial y} = -\log(\overline{U}(x,y)) - 1 - \lambda_1 x - \lambda_2 y - \lambda_{12}xy = 0$	$\frac{\partial^2 L(u(x,y))}{\partial x \partial y} = -\log(u(x,y)) - 1 - \lambda_1 x - \lambda_2 y - \lambda_{12}xy - \lambda_0 = 0$
When we rearrange the above equation, it gives.	When we rearrange the above equation, it gives.
$\overline{U}(x,y) = e^{-\lambda_1 x - \lambda_2 y - \lambda_{12}xy - 1} = e^{-\lambda_1 x}e^{-\lambda_2 y}e^{-\lambda_{12}xy}e^{-1}u(x,y) = \left[\lambda_1\lambda_2 + \lambda_1\lambda_3 x - \lambda_3 + \lambda_2\lambda_3 y + (\lambda_3)^2 xy\right]e^{-\lambda_1 x - \lambda_2 y - \lambda_3 xy}$	$u(x,y) = e^{-\lambda_1 x - \lambda_2 y - \lambda_{12}xy - 1 - \lambda_0} = e^{-\lambda_1 x}e^{-\lambda_2 y}e^{-\lambda_{12}xy}e^{-1-\lambda_0}$

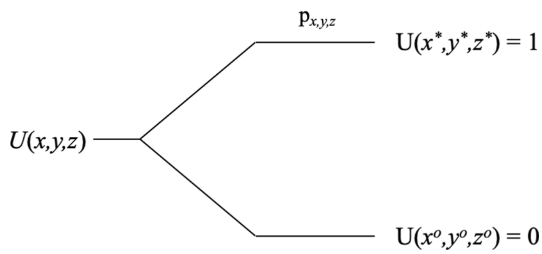

Fig. 10. Three outcome lottery for corner point assessments.

6. Application of maximum disutility entropy

We now apply our method to a party problem which was first introduced by Howard in MS & E 252 Class notes at Stanford University. In this example, a student at Stanford University is planning to arrange a party with three decision alternatives as outdoor, porch or indoor. On the other hand, Kim is uncertain about what the weather will be like on the party day. The weather on the party day can be sunny or rainy which are used as state variable in the decision problem. The corresponding decision tree of Kim's party problem is shown in Fig. 11.

In this party decision problem, Kim has 6 different alternatives and she assign values and corresponding utilities to each alternative as shown in Table 2. In this example, Kim prefers to arrange a party outdoor. If the weather is sunny than this alternative is the most favorable alternative for Kim. On the other hand, if the weather is rainy once Kim choose to arrange the party outside, this alternative is the least favorable option for him. Based on his preferences, utility value of "1″ is assigned to best alternative (O-S), and utility value of "0" is assigned to worst alternative (O-R). Other 4 alternatives are assigned based on the lottery method.

In this example, normally Kim has a utility function but let's assume that we don't know his utility function and our aim is to decide his maximum disutility entropy function over the values of the alternatives. So, the maximum disutility entropy formulation is determined by using the utility values of the alternatives that he is facing as a nonlinear optimization problem and formulized as.

$$\overline{U}_{maxent}(x) = argmax\left(-\int_0^{100}\overline{U}(x)ln\overline{U}(x)dx\right)$$

s.t.

$$\int_0^{50}\overline{U}(x)dx = 0,33 \int_0^{90}\overline{U}(x)dx = 0,05 \quad (56)$$

$$\int_0^{40}\overline{U}(x)dx = 0,43 \int_0^{20}\overline{U}(x)dx = 0,68$$

$$0 \leq \overline{U}(x) \leq 1$$

where the right-hand side values of equation are calculated by using the utility values of Kim from Table-2. For instance, the "indoor-rainy" option has a value of 50 and its corresponding utility value is 0,67. So, the disutility value of $\mu_{50} = 1 - 0,67 = 0,33$.

When the nonlinear optimization model in Equation (56) is solved, we find the solution in exponential form as.

$$e^{-\lambda_0\alpha_0 - r_{50}(x)\lambda_1 - r_{90}(x)\lambda_2 - r_{40}(x)\lambda_3 - r_{20}(x)\lambda_4 - 1} \quad (57)$$

where λ_i is the Lagrange multiplier for the given constraint and r_i is the the given utility preference constraint. Based on the Kim's utility values and corresponding degree of satisfaction values, we compare Kim's maximum disutility entropy function and his utility function based on the given alternatives.

We can easily see from Fig. 12 that the maximum disutility entropy function of Kim is consistent with his utility function. In the degree of satisfaction range of [20,60], maximum disutility entropy values show a bid risk averse behavior then original utility function, however maximum disutility entropy approach assume a good approximation to Kim's utility function.

The maximum disutility approach provides a unique utility function that does not make any prior assumptions about the functional form of the utility function and the structure of utility function is determined by available preference information constraints r_i(x).

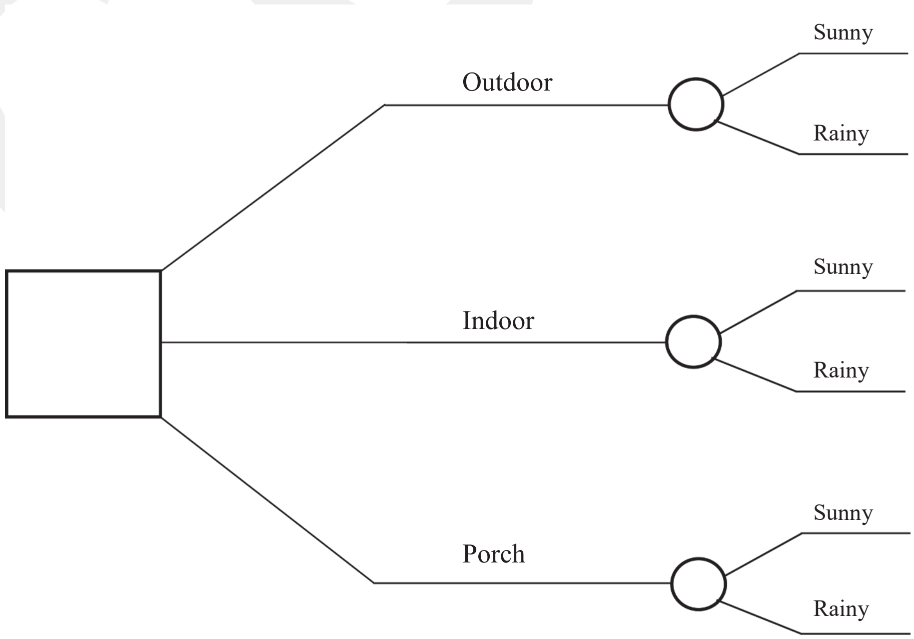

Fig. 11. Decision tree of Kim's party problem.

Table 2
Utility Values for Alternatives of Kim's decision problem.

Alternatives	Value (Degree of Satisfaction)	Utility Value
Outdoor-Sunny (O-S)	100	1
Outdoor-Rainy (O-R)	0	0
Indoor-Sunny (I-S)	40	0,57
Indoor-Rainy (I-R)	50	0,67
Porch-Sunny (P-S)	90	0,95
Porch-Rainy (P-R)	20	0,32

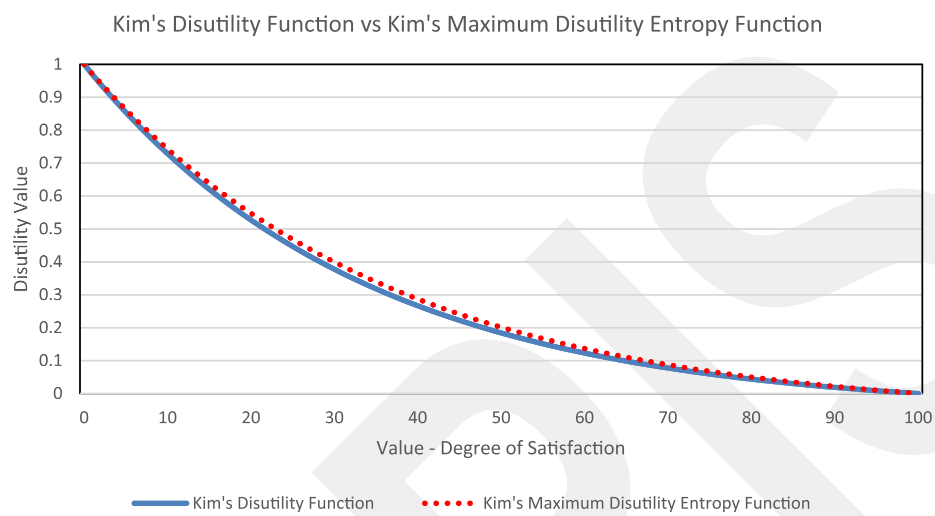

Fig. 12. Comparison of Kim's Disutility Function and Maximum Disutility Entropy Function.

7. Conclusion and future work

In this paper, we have expressed the analogy between utility and probability and defined maximum entropy distribution over utility functions. Because the sign of the cross derivative of multiattribute utility function can be negative, we showed how to use multiattribute utility functions in entropy expression instead of using cross derivatives of utility functions. Also, we don't need to assign a utility function over a value function for multiattribute cases. We can directly use the multi-attribute utility function in this new maximum utility entropy formulation.

We showed the discrete and continuous cases of the maximum disutility entropy expression and gave an example to familiarize the decision makers with maximum disutility entropy approach when there are two attributes. The analogy between utility and probability and new entropy measurement's ability to apply multiattribute cases provides several directions for future research.

In the future work, the maximum disutility model will be extended to various approximations as the number of attributes will be increased, the number of outcomes of each attribute will be increased and different level of dependence structures will be analyzed. Also, different combination of these cases will be examined and sensitivity analysis will be made among attributes and utility functions for these cases.

In addition, the proposed model provides insight regarding the independence and dependence in different scenarios with partial information. The maximum disutility entropy model will be analyzed to see effects of dependence on the accuracy of the model and its results in decision problems by simulating a different number of attributes and its outcomes.

Future research will also be concerned with probability-utility analogy to apply probability rules to utility, using the cumulative form of utility functions. For instance, mutual disutility information and Kullback-Leibler disutility divergence formulations can be expressed in the same analogy. Mutual information formulation and Kullback-Leibler divergence require cross-derivatives of functions and a new formulation based on utility functions would be defined using the cumulative disutility entropy function. Another direction for the future research would be to estimate higher order utility functions when lower order utility assessments are available.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

Calculation of First Moment when $f(x,y) = x$.

$$\because \iint f(x,y)\overline{U}(x,y)dxdy = \iint f(x,y)[1-U(x,y)]dxdy \\ \because = \iint f(x,y)dxdy - \iint f(x,y)U(x,y)dxdy \\ \because = 1 - \iint f(x,y)U(x,y)dxdy \tag{58}$$

The right-hand side of the equation (58) is the one minus the expected utility. We can write down the equation (58) separately to solve it for each attribute separately.

$$\because \int_0^1 \int_0^1 f(x,y)U(x,y)dxdy = \int_0^1 \left\{ \int_0^1 xU(x,y)dx \right\} dy \tag{59}$$

First of all, we work on the equation between brackets. And we use integration by parts and we assign $u = x \rightarrow du = dx$ and $dv = U(x,y)dx \rightarrow v = U(x,y)$.

$$\because \int_0^1 xU(x,y)dx = xU(x,y)|_{x^o=0}^{x^*=1} - \int_0^1 U(x,y)dx \\ \because = U(x^*,y) - \int_0^1 U(x,y)dx \tag{60}$$

We insert the equation (60) and we have.

$$\because = \int_0^1 \left\{ U(x^*,y) - \int_0^1 U(x,y)dx \right\} dy \\ \because = \int_0^1 U(x^*,y)dy - \int_0^1 \int_0^1 U(x,y)dxdy \tag{61}$$

The first part of the eq. (61) is not normalized; so, we first normalize it. Using the definition of marginal utility.

$$\because U(y|x^*) = \frac{U(x^*,y) - U(x^*,y^0)}{U(x^*,y^*) - U(x^*,y^0)} = \frac{U(x^*,y) - k_x}{1-k_x} \qquad where\ \ k_x = U(x^*,y^0) \\ \because U(x^*,y) = (1-k_x)U_{y|x^*}(y) + k_x \tag{62}$$

We now make the following substitution into (58) and with calculus, we get.

$$\because \int_0^1 \int_0^1 f(x,y)U(x,y)dxdy = 1 - \left[(1-k_x)EU_{y|x^*} + k_x\right] - \left[1 - EU_{xy}\right] \\ \because = EU_{xy} - (1-k_x)EU_{y|x^*} - k_x \tag{63}$$

And finally, we have.

$$\because \iint f(x,y)\overline{U}(x,y)dxdy = 1 + (1-k_x)EU_{y|x^*} + k_x - EU_{xy} \tag{64}$$

Appendix B

By symmetry, and using the same logic as Appendix 1, we can also write equation (26) by exchanging $f(x,y) = x$ with $f(x,y) = y$.

Appendix C

Calculation of First Moment when $f(x,y) = xy$.

$$\because \iint f(x,y)\overline{U}(x,y)dxdy = \iint f(x,y)[1-U(x,y)]dxdy \\ \because = \iint f(x,y)dxdy - \iint f(x,y)U(x,y)dxdy \\ \because = 1 - \iint f(x,y)U(x,y)dxdy \tag{65}$$

The right-hand side of the equation (65) is the one minus the expected utility.
We first calculate the second term in (65) to solve it for each attribute separately.

$$\because \int_0^1 \int_0^1 f(x,y)U(x,y)dxdy = \int_0^1 y\left\{ \int_0^1 xU(x,y)dx \right\} dy \tag{66}$$

First of all, we work on the equation between brackets. And we use integration by parts and we assign $u = x \rightarrow du = dx$ and $dv = U(x,y)dx \rightarrow v = U(x,y)$.

GCRIS

$$\begin{aligned} \because \int_0^1 xU(x,y)dx = xU(x,y)\big|_{x^o=0}^{x^*=1} - \int_0^1 U(x,y)dx \\ \because \qquad = U(x^*,y) - \int_0^1 U(x,y)dx \end{aligned} \tag{67}$$

We insert the equation (67) into (66) and we have.

$$\begin{aligned} \because = \int_0^1 y\left\{U(x^*,y) - \int_0^1 U(x,y)dx\right\}dy \\ \because = \int_0^1 yU(x^*,y)dy - \int_0^1\int_0^1 yU(x,y)dxdy \end{aligned} \tag{68}$$

Again, we apply integration by parts second term of RHS of (68) and we assign $u = y \rightarrow du = dy$ and$dv = U(x,y)dy \rightarrow v = U(x,y)$.

$$\begin{aligned} \because \int_0^1 yU(x,y)dy = yU(x,y)\big|_{y^o=0}^{y^*=1} - \int_0^1 U(x,y)dy \\ \because \qquad = U(x,y^*) - \int_0^1 U(x,y)dy \end{aligned} \tag{69}$$

Then, we combine (68) and (69) and.

$$\begin{aligned} \because = \int_0^1 yU(x^*,y)dy - \int_0^1 \left[U(x,y^*) - \int_0^1 U(x,y)dy\right]dx \because = \int_0^1 yU(x^*,y)dy - \int_0^1 U(x,y^*)dx + \int_0^1 U(x,y)dydx \because \\ = (1-k_x)EU_{y|x^*} + k_x - \left[1 - ((1-k_y)EU_{x|y^*} + k_y)\right] + \left[1 - EU_{xy}\right] \because = (1-k_x)EU_{y|x^*} + k_x + (1-k_y)EU_{x|y^*} + k_y - EU_{xy} \end{aligned} \tag{73}$$

So, cross moment of disutility function is.

$$\because 1 - \iint f(x,y)U(x,y)dxdy = 1 - (1-k_x)EU_{y|x^*} - k_x - (1-k_y)EU_{x|y^*} - k_y + EU_{xy} \tag{74}$$

And finally, we have.

$$\because \iint f(x,y)\overline{U}(x,y)dxdy = 1 + EU_{xy} - (1-k_x)EU_{y|x^*} - (1-k_y)EU_{x|y^*} - (k_x + k_y) \tag{75}$$

References

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. Journal, 27*, 623–656.

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics. *Physical review, 106*(4), 620.

Jaynes, E. T. (1957). Information Theory and Statistical Mechanics II. *Physical Review, 108*(2), 171.

Rao, M. (2005). More on a new concept of entropy and information. *Journal of Theoretical Probability, 18*(4), 967–981.

Smith, S. A. (1974). A derivation of entropy and the maximum entropy criterion in the context of decision problems. *IEEE Transactions on systems, Man, and Cybernetics, 2*, 157–163.

Jessop, A. (1999). Entropy in multiattribute problems. *Journal of Multi-Criteria Decision Analysis, 8*(2), 61–70.

Berhold, M. H. (1973). The use of distribution functions to represent utility functions. *Management Science, 19*(7), 825–829.

Barron, F. H. (1992). Selecting a best multiattribute alternative with partial information about attribute weights. *Acta psychologica, 80*(1–3), 91–103.

Hadfi, R., Ito, T. (2012). Asymptotic Maximum Entropy Principle for Utility Elicitation under High Uncertainty and Partial Information. *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference*, Marco Island, Florida, USA.

Herfert, M., La Mura, P. (2004). Estimation of Consumer Preferences via Ordinal Decision-Theoretic Entropy. Working Paper 64, *Leipzig Graduate School of Management (HHL)*.

Abbas, A. E. (2004). Entropy methods for adaptive utility elicitation. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 34*(2), 169–178.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* (2nd ed.). Princeton, NJ: Princeton University Press.

Abbas, A. E. (2006). Entropy methods for joint distributions in decision analysis. *Engineering Management, IEEE Transactions on, 53*(1), 146–159.

Abbas, A. E. (2006). Maximum entropy utility. *Operations Research, 54*(2), 277–290.

Pires, C., Dionísio, A., & Coelho, L. (2013). Estimating utility functions using generalized maximum entropy. *Journal of Applied Statistics, 40*(1), 221–234.

Kim, E. Y., & Ahn, B. S. (2019). Some Notes on Maximum Entropy Utility. *Entropy, 21*(7), 637.

Bajgiran, A. H., Mardikoraem, M., & Soofi, E. S. (2021). Maximum entropy distributions with quantile information. *European journal of operational research, 290*(1), 196–209.

Maeda, Y., Nguyen, H. T., & Ichihashi, H. (1993). Maximum entropy algorithms for uncertainty measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 1*(01), 69–93.

Chen, X., & Dai, W. (2011). Maximum entropy principle for uncertain variables. *International Journal of Fuzzy Systems, 13*(3), 232–236.

Harju, M., Liesiö, J., & Virtanen, K. (2019). Spatial multi-attribute decision analysis: Axiomatic foundations and incomplete preference information. *European Journal of Operational Research, 275*(1), 167–181.

Wang, C., & Bier, V. M. (2013). Expert elicitation of adversary preferences using ordinal judgments. *Operations Research, 61*(2), 372–385.

Dionisio, A., Reis, A. H., & Coelho, L. (2008). Utility function estimation: The entropy approach. *Physica A: Statistical Mechanics and its Applications, 387*(15), 3862–3867.

Gao, B., Wu, C., Wu, Y., & Tang, Y. (2015). Expected Utility and Entropy-Based Decision-Making Model for Large Consumers in the Smart Grid. *Entropy, 17*(10), 6560–6575.

Al-Omari, A. I. (2016). A new measure of entropy of continuous random variable. *Journal of Statistical Theory and Practice, 10*(4), 721–735.

Chaji, A., Fukuyama, H., & Shiraz, R. K. (2018). Selecting a model for generating OWA operator weights in MAGDM problems by maximum entropy membership function. *Computers & Industrial Engineering, 124*, 370–378.

Sutcu, M. (2020). Effects of total cost of ownership on automobile purchasing decisions. *Transportation Letters, 12*(1), 18–24.

Montiel, L. V., & Bickel, J. E. (2013). Approximating joint probability distributions given partial information. *Decision Analysis, 10*(1), 26–41.

Montiel, L. V., & Bickel, J. E. (2014). A generalized sampling approach for multilinear utility functions given partial preference information. *Decision Analysis, 11*(3), 147–170.

Ebrahimi, N. (2000). The maximum entropy method for lifetime distributions. *Sankhyā: The Indian Journal of Statistics, Series A*, 236–243.

Rao, M., Chen, Y., Vemuri, B. C., & Wang, F. (2004). Cumulative residual entropy: A new measure of information. *IEEE Trans. Inform. Theory, 50*(6), 1220–1228.

Abbas, A. E. (2002). An entropy approach for utility assignment in decision analysis. *Proc. 22nd International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering. AIP Conference Proceedings*. Moscow, ID.