Available online at www.sciencedirect.com

ScienceDirect

Physica A 387 (2008) 3862–3867

PHYSICA A

www.elsevier.com/locate/physa

Utility function estimation: The entropy approach

Andreia Dionisio[a,*], A. Heitor Reis[b,c], Luis Coelho[a]

[a] University of Evora, Center of Business Studies, CEFAGE-UE, Largo Colegiais, 2, 7000 Evora, Portugal
[b] University of Evora, Department of Physics, Portugal
[c] University of Evora, Evora Geophysics Center, Portugal

Available online 4 March 2008

Abstract

The maximum entropy principle can be used to assign utility values when only partial information is available about the decision maker's preferences. In order to obtain such utility values it is necessary to establish an analogy between probability and utility through the notion of a utility density function. In this paper we explore the maximum entropy principle to estimate the utility function of a risk averse decision maker.

Keywords: Maximum entropy; Utility functions; Preferences; Risk aversion

0. Introduction

The main goal of this research work is to explore the potential of the maximum entropy principle (ME) to estimate utility functions. In fact, utility functions are one of the most important concepts in decision analysis. They can be estimated empirically using partial information about the agent's preferences and its tolerance about the risk. In this paper we refer to partial information when we only have inferred the utility values based on observed decisions.

The main assumption taken to derive the utility function of an agent, using the ME principle is the correspondence between the concept of equilibrium in physics (statistical) and economics (mechanical). According to some authors (namely Foley [7], Candeal et al. [5], Darooneh [6]) economic equilibrium can be viewed as an asymptotic approximation to physical equilibrium and some difficulties with mechanical picture (economic) of the equilibrium may be eased by considering the statistical (physical) description of it.

In this paper we explore the ME principle to estimate the utility values of a risk averse investor. The rest of the paper is organized as follows. In Section 1 we present a brief discussion of the background theory, namely the ME principle and its applications to economics and more specifically to decision analysis. Section 2 presents the analogy between utility and probability, and utility and entropy. In order to explain in a better way these matters, we show a short example. Finally, Section 3 presents the main conclusions of this study.

* Corresponding author.
E-mail addresses: andreia@uevora.pt (A. Dionisio), ahr@uevora.pt (A.H. Reis), lcoelho@uevora.pt (L. Coelho).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2008.02.072

1. Background theory

Suppose that we have a set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$ and H is a measure of uncertainty. Shannon [18] developed a measure of uncertainty associated with an outcome from a set of symbols that satisfy the following properties (i) H should be continuous in $p_i, i = 1, \ldots, n$; (ii) if $p_i = 1/n$, then H should be a monotonic increasing function of n; (iii) H is maximized in a uniform probability distribution context; (iv) H should be additive; (v) H should be the weighted sum of the individual values of H.

According to Shannon [18] a measure that satisfies all these properties is the entropy which is defined as $H(X) = -\sum_i p_i \log p_i$. When the random variable has a continuous distribution, and $p_X(x)$ is the density function of the random variable X, the entropy (usually called differential entropy) is given by $H(X) = -\int p_X(x) \log p_X(x) \mathrm{d}x$.

The properties of the entropy of continuous (differential entropy) and discrete distributions are mainly alike. For continuous distributions, $H(X)$ is not scale invariant ($H(cX) = H(X) + \log|c|$) but is translation invariant ($H(c + X) = H(X)$). The differential entropy may be negative and infinite [18,19]. Entropy [$H(X)$] is a measure of the average amount of information provided by an outcome of a random variable and similarly, is a measure of uncertainty about a specific possible outcome before observing it [11].

Jaynes [14] introduced the maximum entropy (ME) principle as a generalization of Laplace's principle of insufficient reason. The ME principle appears as the best way when we intend to make an inference about an unknown distribution based only on few restriction conditions, which represent some moments of the distribution. According to several authors (see for example Refs. [20,11]) this principle uses only relevant information and eliminates all irrelevant details from the calculations by averaging over them.

The ME model is usually formulated to confirm the equality constraints on moments, or cumulative probabilities, of the distribution of the random variable X, where $h_j(X_i)$ is an indicator function over an interval for cumulative probability constraints and b_j are the moments j of the distribution.

$$
\begin{aligned}
& p^* = \arg\max -\sum_i p_i \log p_i, \quad \text{s.t.} \\
& \sum_i p_i = 1 \\
& \sum_i h_j(X_i)\, p_i = b_j \\
& p_i \geq 0 \; j = 1, \ldots, m, i = 1, \ldots, n.
\end{aligned}
\tag{1}
$$

The density that respects all the conditions of the model (1) is defined as *Entropy Density (ED)*. The Lagrangian of the problem is

$$
L = -\sum_i p_i \log p_i - \lambda_0 \left[\sum_i p_i - 1\right] - \sum_{j=1}^{m} \lambda_j \left[\sum_i h_j(X_i)\, p_i - b_j\right], \tag{2}
$$

where λ_0 and λ_i are the Lagrange multipliers for each probability or moment constraint. The solution to this problem is

$$
p_i = \exp\left[-\lambda_0 - 1 - \sum_{j=1}^{m} \lambda_j h_j(X_i)\right]. \tag{3}
$$

For small values of m it is possible to obtain explicit solutions [24]. If $m = 0$, meaning that no information is given, one obtains a uniform distribution. As one adds the first and the second moments, Golan et al. [10] recall that one obtains the exponential and the normal density, respectively. The knowledge of the third or higher moments does not yield a density in a closed form and only numerical solutions may provide densities.

In many cases, precise values for moments and probabilities are unavailable. In face of this problem Abbas [4] propose the use of the ME principle using upper and lower bonds in the moments constraints.

There are several research studies of ME applied to economics (see for example Refs. [12,16,21]).

The ME principle has been more recently applied in decision analysis, specially in the specification and estimation of utility values and utility functions. For example, Fritelli [9] derives the relative entropy minimizing martingale measure under incomplete markets and demonstrates the connection between it and the maximization of exponential utility. Herfert and La Mura [13] use a non-parametric approach based on the maximization of entropy to obtain a model of consumer preferences using available evidence, namely surveys and transaction data. In a different approach Abbas [3] presents an optimal question-algorithm to elicit von Neumann and Morgenstein utility values using the ME principle. Abbas [1] uses ME to assign utility values when only partial information is available about the decision maker's preferences and Abbas [2] uses the discrete form of ME principle to obtain a joint probability distribution using lower order assessments. Yang and Qiu [22] propose an expected utility-entropy measure of risk in portfolio management, and the authors conclude that using this approach it is possible to solve a class of decision problems which cannot be dealt with by the expected utility or mean-variance criterion.

Sandow et al. [17] use the minimization of cross-entropy (or relative entropy) to estimate the conditional probability distribution of the default rate as a function of a weighted average bond rating, concluding that the modeling approach is asymptotically optimal for an expected utility maximizing investor. Friedman et al. [8] explore an utility-based approach to some information measures, namely the Kullback–Leibler relative entropy and entropy, using the example of horse races. On the other way, Darooneh [6] uses the ME principle to find the utility function and the risk aversion of agents in a exchange market.

According to Abbas [2], the ME principle presents several advantages when we seek to construct joint probability distributions and assign utility values, namely: (i) it incorporates as much information as there is available at the time of making the decision; (ii) it makes no assumptions about a particular form or a joint distribution; (iii) it applies to both numeric and nonnumeric variables; and (iv) it does not limit itself to the use of only moments and correlation coefficients, which may be difficult to obtain in decision analysis practice.

2. Utility and entropy

When a decision problem is deterministic, the order of the prospects is enough to define the optimal decision alternative. However, when uncertainty is present, it is necessary to assign the von Neumann and Morgenstein utility values. One of the basic assumptions of decision theory is that an agent's observed behaviour can be rationalized in terms of the underlying preference ordering, and if the observed behaviour is consistent with the ordering we can make inferences about the utility function using the available data. Sometimes the observations are not sufficient to clearly identify the orderings and one needs more general inference methods. La Mura [15] presented a non-parametric method for preference estimation based on a set of axiomatic requirements: (i) no information; (ii) uniqueness; (iii) invariance; (iv) system independence, and (v) subset independence. The axioms characterize a unique inference rule, which amounts to the maximization of the entropy of the decision-maker's preference ordering.

We extend an approach developed by Abbas [1] and also used before in a similar way by Herfert and La Mura [13], the *maximum entropy utility principle*, where a utility function is normalized to the range between zero and unity and the utility density function is the first derivative of a normalized utility function. Based on such a definition, the utility density function has two main properties: (i) is non-negative; and (ii) integrates to unity. The two properties allows the analogy between utility and probability, and consequently, with entropy [1].

For the discrete case, the utility vector has K elements, defined as

$$U \triangleq (u_0, u_1, \ldots, u_{K-2}, u_{K-1}) = (0, u_1, \ldots, u_{K-2}, 1). \tag{4}$$

This vector of dimension K can be represented as a point in a $(K-2)$ dimensional space, which is defined by $0 \leq u_1 \leq \cdots \leq u_{K-2} \leq 1$. This region, called utility volume, has a volume equal to $1/(K-2)!$.

In the utility increment vector (ΔU) the elements are equal to the difference between consecutive elements in the utility vector, it has $K-1$ elements and is defined by

$$\Delta U \triangleq (u_1 - 0, u_2 - u_1, \ldots, 1 - u_{K-2}) = (\Delta u_1, \Delta u_2, \ldots, \Delta u_{K-1}).$$

The coordinates of ΔU are all non-negative and sum to unity.

According to Abbas [1] the knowledge of the preference order alone does not give any information at all about the location of the utility increment vector. In these conditions it is reasonable to assume that the respective location is uniformly distributed over the domain. The assumption gives equal likelihood to all utility values and satisfy the agent's preference order, adding no further information than the knowledge of the order of the prospects.

For the continuous case, the concepts are similar, but the number of prospects K can be infinite. Is this case the utility vector is a utility curve $[U(x)]$, and has the same mathematical properties as a cumulative probability distribution. The utility increment vector (or in this case, utility density function) is now a derivative of the utility curve

$$u(x) \triangleq \frac{\partial U(x)}{\partial x} \tag{5}$$

which is non-negative and integrates to unity.

Given the analogy between utility and probability, the concept of entropy can be used as a measure of spread for the coordinates of the utility increment vector

$$H(\Delta u_1, \Delta u_2, \ldots, \Delta u_{K-1}) = -\sum_{i=1}^{K-1} \Delta u_i \log \Delta u_i. \tag{6}$$

The utility increment vector that maximizes this measure is the uniform distribution. There are other measures that can be used to spread the utility increment vector, although, the entropy satisfies the following 3 axioms: (1) the measure of spread of the utility increment vector is a monotonically increasing function of the number of prospects K, when the utility increments are all equal; (2) the measure of spread of a utility increment vector should be a continuous function of the increments; (3) the order in which we calculate the measure of spread should not influence the results.

The differential entropy can also be applied to a utility density function

$$H(u(x)) = -\int_a^b u(x) \log u(x) \mathrm{d}x,$$

and this function is maximized when $u(x) = 1/(b-a)$. The uniform density integrates to a linear (risk neutral) utility function.

The maximum entropy utility problem is described by

$$\begin{aligned} & u_{\max ent}(x) = -\int_a^b u(x) \log u(x) \mathrm{d}x, \quad \text{s.t.} \\ & \int_a^b u(x) \mathrm{d}x = 1 \\ & \int_a^b h_i(x) u(x) \mathrm{d}x = b_i \\ & u(x) \geq 0, i = 1, \ldots, n. \end{aligned} \tag{7}$$

Abbas [1] used a CARA utility density to show that the differential entropy has a unique maximum, that occurs exactly when the agent is risk neutral.

This approach is also defended by Darooneh [6], who considers that the equilibrium condition may be expressed by the maximum entropy utility, since the risk of the market induces the randomness. The solution for this problem is given by the following expression

$$u_{\max ent}(x) = \exp[-\lambda_0 - 1 - \lambda_1 h_1(x) - \lambda_2 h_2(x) - \cdots - \lambda_n h_n(x)], \tag{8}$$

where $[a, b]$ are the domain of the prospects, $h_i(x)$ is a given preference constraint, $b_i' s$ are a given sequence of utility values or moments of the utility function and λ_i is the Lagrangian multiplier for each utility value. The uniform utility density is a special case of Eq. (8) where the constraints $h_i(x)$ do not exist. When $h_1(x) = x$ and the remaining constraints are zero, the maximum entropy utility is a CARA utility on the positive domain. When $h_1(x) = x$ and $h_2(x) = x^2$ the maximum entropy utility is a Gaussian utility density, which integrates to a S-shaped prospect theory utility function on the real domain.

The risk aversion parameter (γ), using the Arrow–Pratt definition, of the agent is given by

$$\gamma_{\max ent}(x) = -\frac{\partial \ln[u_{\max ent}(x)]}{\partial x} = \lambda_1 h_1'(x) + \lambda_2 h_2'(x) + \cdots + \lambda_n h_n'(x), \tag{9}$$

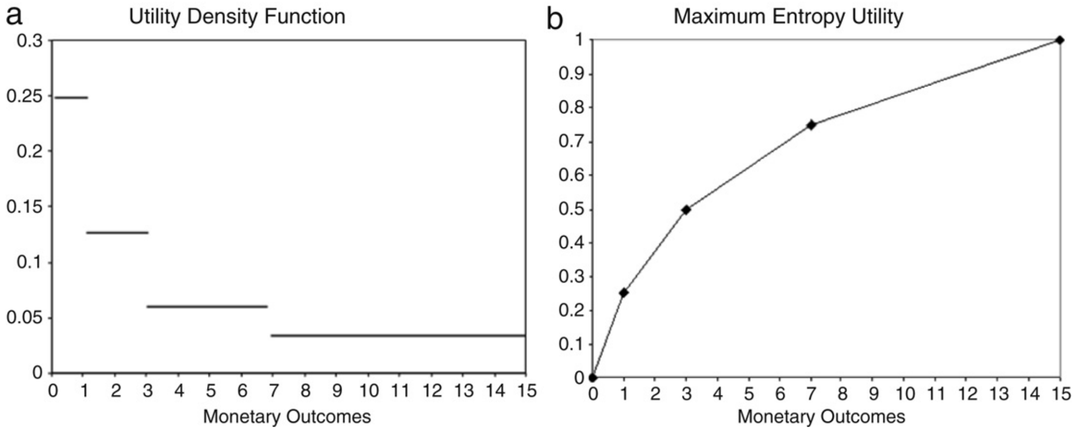

Fig. 1. Utility density function (a) and utility function (b).

Table 1
Elicitation values for utility

$u(x)$	0	0.25	0.50	0.75	1.0
x	0	1	3	7	15

x are the monetary outcomes.

where $h_i'(x) = \partial h_i(x)/\partial x$. The Eq. (9) shows the linear effect contributed by the derivative of each preference constraint on the overall risk aversion function.

Abbas [1] presents several examples of application of maximum entropy utility principle, namely for cases when we know some utility values, cases when we need to infer utility values by observing decisions and for the case of multiattribute utility. For all the cases explored, Abbas [1] concludes that the maximum entropy utility principle presents advantages and satisfies the important assumption of utility and probability independence that stems from the foundations of normative utility theory.

The illustration of the maximum entropy utility problem can be made when we have a small number of utility values. Sometimes when we elicited the values of some prospects (using for example the Certainty Equivalent Method) the decision maker is not interested to give many answers [23]. In this case, it is difficult to find the best curve for the utility and some statistical indicators are not useful.

Suppose that for some decision maker it is only possible to pose three questions to elicit the utility (first and last utility values are fixed). Table 1 shows the available information.

The maximum entropy formulation for the density function is

$$
\begin{aligned}
&u_{\max ent}(x) = -\int_0^{15} u(x)\log u(x)\mathrm{d}x, \quad \text{s.t.} \\
&\int_0^{15} u(x)\mathrm{d}x = 1; \quad \int_0^{7} u(x)\mathrm{d}x = 0.75; \quad \int_0^{3} u(x)\mathrm{d}x = 0.5; \quad \int_0^{1} u(x)\mathrm{d}x = 0.25 \\
&u(x) \geq 0.
\end{aligned}
\tag{10}
$$

If we compare the preference constraints $h_i(x)$ of Eqs. (7) and (10), we find indicator functions over the intervals. The solution of the maximization problem has the form

$$u_{\max ent}(x) = \exp\left[-\lambda_0 - 1 - \lambda_1 I_1(x) - \lambda_2 I_2(x) - \lambda_3 I_7(x)\right], \quad 0 \leq x \leq 15,$$

where I_j are indicator functions for the interval j This equation is a staircase utility density function, as we see in Fig. 1. It also shows the maximum entropy utility function is a piecewise linear function connecting the given utility values.

3. Conclusions

This paper presents an efficient alternative way to estimate the utility function of any agent when there is only partial available information about the decision maker's preferences. The maximum entropy approach, here presented, provides a unique utility function that makes no assumptions about the structure, unless there is preference information to support it.

Based on the recent literature on this area of research, we show that the analogy probability — utility can be explored in order to use information theory measures, and obtain a more robust estimation of the utility function.

Acknowledgement

Financial support from Fundacao da Ciencia e Tecnologia, Lisbon, is gratefully acknowledged by the authors, under the contract PTDC/GES/70529/2006.

References

[1] A. Abbas, Operations Research 54 (2006).
[2] A. Abbas, IEEE Transactions on Engineering Management 53 (2006) 146–159.
[3] A. Abbas, IEEE Transactions on Systems, Man and Cybernetics 34 (2004) 169–178.
[4] A. Abbas, Proceedings of the 25th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, 2005.
[5] J. Candeal, J. De Miguel, E. Indurain, G. Mehta, Utility and entropy, Economic Theory 17 (2001) 233–238.
[6] A. Darooneh, Entropy 8 (2006) 18–24.
[7] D. Foley, Journal of Economic Theory 62 (1994) 321.
[8] C. Friedman, J. Huang, S. Sandow, Entropy 9 (2007) 1–26.
[9] M. Fritelli, Mathematical Finance 10 (2000) 39–52.
[10] A. Golan, G Judge, D. Miller, Maximum Entropy Econometrics: Robust Estimation with Limited Data, John Wiley Sons, New York, 1996.
[11] A. Golan, Journal of Econometrics 107 (2002) 1–15.
[12] B. Gulko, The entropy pricing theory — market beliefs, valuation, incomplete markets, asset pricing, Ph.D. Dissertation, UMI – Dissertations Services, 1998.
[13] M. Herfert, P. La Mura, Estimation of consumers preferences via ordinal decision-theoretic entropy, Working-paper 64, Leipzig Graduate School of Management, 2004.
[14] E. Jaynes, Phys. Rev. 106 (1957) 620.
[15] P. La Mura, Proceedings of the 9th Conference on Theoretical Aspects of Rationality and Knowledge, Bloomington, IN, USA, 2003.
[16] D. Samperi, Entropy and statistical model selection for asset pricing and risk management, 1999, preprint in http://papers.ssrn.com.
[17] S. Sandow, C. Friedman, M. Gold, P. Chang, Journal of Banking and Finance 30 (2006) 679–693.
[18] C. Shannon, Bell Systems Technical journal 27 (1948) 379–423; 623–656.
[19] E. Soofi, Journal of the American Statistical Association 89 (1994) 1243–1254.
[20] E. Soofi, Journal of the American Statistical Association 95 (2000) 1349–1353.
[21] M. Stuzer, Journal of Finance 51 (1996) 1633–1652.
[22] J. Yang, W. Qiu, European Journal of Operational Research 164 (2005) 792–799.
[23] P. Wakker, D. Deneffe, Management Science 42 (8) (1996) 1131–1150.
[24] A. Zellner, Journal of Econometrics 75 (1996) 51–68.