Journal of the Korean
Statistical Society
Vol. 26, No. 4, 1997

Steady-state Probabilities under Non-additivity

Keuk-Ryoul Yoo[1]

Abstract

Uncertainty, which arises when little information is revealed, can be represented by a non-additive probability, while risk is described by an additive one. This paper demonstrates that in the presence of uncertainty a steady state probability exists, which implies that we can estimate an average over a long period even under uncertainty. It is also shown that the steady state probability may not be unique in the presence of uncertainty. This implies that the estimated average under uncertainty is less accurate than under risk.

Key Words : Uncertainty; Non-additive probability; Steady-state probability.

[1]Department of Business Administration, Dongduck Women's University, Seoul, Korea.

1. INTRODUCTION

In this paper, uncertainty is defined as a situation where a probability distribution is not well defined. Uncertainty can be represented by a non-additive probability. Non-additive probability is sometimes called capacity. A transition probabilities matrix is uncertain if at least one one-time unit conditional probability is not additive, while it is risky otherwise. Because the model in this paper assumes the possibility of non-additivity, it is generalization of a traditional additive case.

I demonstrate that in the presence of uncertainty a steady state probability exists, however the steady state probability may not be unique. The existence suggests that we can estimate an average over a long period even under uncertainty. However, the estimated average may be less accurate than under risk due to non-uniqueness.

In the classical probability theory, a probability is assumed to be additive. Savage(1954) developed the expected utility theory under the assumption that an individual faces an additive probability distribution. And in most literature, an uninformed decision maker who does not know the true state is assumed to have a single additive probability. This approach, however, does not distinguish between the situation where an individual has enough relevant information and the situation of little information. That is, the approach does not include the fact that people are more reluctant to choose a lottery when little information is reveled.

A distribution between risk and uncertainty was made by Knight(1921). A random variable is risky if its distribution is known, while it is uncertain if the distribution is unknown. He argued that uncertainty arises when decision maker's subjective opinions, without enough information, mainly determine the outcome. The Ellsberg paradox(1961) showed that an individual's behavior under uncertainty is different from under risk, and he is more reluctant to choose an uncertain asset than a risky one.

There are two main approaches to describe individual's behavior under uncertainty, the multiple priors method and the non-additive probability method. The former approach is developed by Bewley(1986) and Gilboa and Schmeidler(1989). The latter approach to uncertainty is elaborated by Dempster(1968), Shafer(1976), Schmeidler(1989), Gilboa(1987) and Wakker(1989). Gilboa and Schmeidler(1991) demonstrated that a preference ordering in the multiple priors approach can be preserved in the latter approach under some reasonable conditions. Yoo(1991) showed that there is dynamic inconsistency in the presence of uncertainty, i.e., the iterative law of expectation does not hold under uncertainty.

2. APPLICATIONS

There are many arguments which support the admission of uncertainty. Most important of these are when (i) a group of individuals are involved in decision making, (ii) there is no previous experience or total absence of relevant information(For more example, see Walley(1991)). I will give a few simple examples in this section. Consider a case where a group of individuals are involved in decision making. A group of individuals do not usually have deterministic beliefs even when each individual does, because of disagreement among them. Suppose that there are two individuals and the decision is finalized only when both agree. Suppose also that the present status gives utility of 140, and that the outcomes for the new alternative is 100 or 200. The probabilities assessed are 0.3 and 0.7 for the first decision maker, and 0.8 and 0.2 for the second one. Then the expected uncertainties for the new alternative are 170 and 120.

Therefore, the group will not take the new alternative because one disagrees to take it. In this case it is better to say that the appropriate expected utility for the second alternative is 120(=Min {170,120}). This utility level can be easily calculated using the non-additive approach by taking low probabilities 0.3 and 0.2(the non-additive probability approach yields the expected utility level $(0.2 \times 200) + \{(1 - 0.2) \times 100\} = 120$ using the Choquet integration which will be defined later). We know that the low probabilities are not additive.

Uncertainty also applies when a decision maker has little experiences. Suppose a decision maker assesses the probability that a particular thumbtack will land pin-up on a specific toss. If he/she has no previous experience with this or similar thumbtacks, he/she may not estimate a unique probability. Instead, he/she can assess a range of probabilities for an event. Thus there exists lower and upper probabilities for the event. The lower probabilities which are relevant to decision making may not be additive. Note that the upper and lower probabilities will tend to converge as he/she accumulates information about its propensity to fall pin-up.

3. MODEL

Let Ω be a nonempty finite set of states of nature, $\Omega = \{1, 2, \ldots, N\}$. Let $v_A^{(n)}(B)$ be the possibly non-additive conditional probability that the random variables X, starting in event A, will be in event B after exactly n time units. For notational convergence, we will drop 1 if $n = 1$. The space of one time unit conditional probabilities and initial probabilities is $v = \{v : 2^{\Omega} \to \mathcal{R} | v(\Phi) = 0,\ v(\Omega) = 1,\ A \subseteq B \Rightarrow v(A) \leq v(B),\ A, B \subseteq \Omega\}$. It is also assumed that $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$(This condition is called convexity). Note that additivity is not assumed in v. Let $\mathbf{H}^{(n)}$ be the square transition probabilities matrix of $v_A^{(n)}(B) \in v$, where both the row events and column events are disjoint and exhaustive. Let $q_B^{(n)}(A)$ denote the value corresponding to event A in row and event B in column and let $q_B^{(n)}$ be a function taking values $q_B^{(n)}(A)$, where A is the row event of $\mathbf{H}^{(n)}$.

Since the probability may not be additive, we cannot use a traditional integral. To integrate with respect to a possibly non-additive probability, we use the Choquet integral(Choquet(1953)). The Choquet integral with respect to v_A is defined as follows(note that the first element in the right-hand-side equation is 0 in this paper since $q_B^{(n)}$ is non-negative.) :

$$E_{v_A}[q_B^{(n)}] = -\int_{-\infty}^{0} [1 - v_A(\{s : q_B^{(n)}(s) \geq \alpha\})]\, d\alpha + \int_{0}^{\infty} v_A(\{s : q_B^{(n)}(s) \geq \alpha\})]\, d\alpha.$$

We now turn to define a Chapman-Kolmogorov equation in the case of possibly non-additive probabilities. It is defined as :

$$v_A^{(n)}(B) = \int_{0}^{\infty} v_A(\{s : q_B^{(n-1)}(s) \geq \alpha\})]\, d\alpha.$$

The Chapman-Kolmogorov equation says that $v_A^{(n)}(B)$ is the expectation of $q_B^{(n-1)}$ with respect to the probability $v_A(\cdot)$. Notice first that if a probability is additive, the Chapman-Kolmogorov equation defined in this paper yields the same result as the traditional one, i.e., $\int_0^\infty v_A(\{s : q_B^{(n)}(s) \geq \alpha\})]\, d\alpha. = \sum_k v_A(k) q_B^{(n)}(k) = \sum_k v_A(k) v_k^{(n)}(B)$, where k is the kth column event of a row event A. Thus the Chapman-Kolmogorov equation defined in this paper is generalized of the traditional one. Notice second that if $v_A^{(n)}$ is strictly monotonic and non-additive, $v_A^{(n+1)}$ is possibly non-additive because $E_{v_A}[q_B] \neq E_{v_A}[q_E] + E_{v_A}[q_F]$ for a non-additive probability v_A and E, F, $B = E \cup F$, $E \cap F = \emptyset$. Notice third that the traditional Chapman-Kolmogorov equation is not adequate in the presence of uncertainty since all the steady-state probabilities are 0 under non-additivity. The intuition behind is as follows. $v_A^{(n)}(B)$ is the weighted average of a random variables

$q_B^{(n-1)}$. If we use a traditional equation under non-additivity, the sum of the weights is less than 1, which makes relatively smaller the m time unit conditional probability for a larger m. If we use it in a Choquet fashion, however, the sum of the weights is always 1. The following example demonstrates this point.

Example. Suppose the transition probabilities matrix is

$$\mathbf{H} = \begin{pmatrix} .1 & .6 \\ .7 & .2 \end{pmatrix}$$

Using the traditional equation $\sum_{\{k\in\Omega\}} v_i(k)v_k(j)$ yields

$$\begin{pmatrix} .43 & .18 \\ .21 & .46 \end{pmatrix}$$

And, $\lim_{n\to\infty} \sum_{k\in\Omega} v_i(k)v_k^{(n-1)}(j) = 0$, for $i, j = 1, 2$.
However, using the definition in this paper,

$$\mathbf{H}^{(2)} = \begin{pmatrix} .46 & .24 \\ .22 & .48 \end{pmatrix}$$

where $0.46 = (0.6)(0.7) + (1 - 0.6)(0.1)$.
And

$$\mathbf{H}^{(3)} = \begin{pmatrix} .24 & .38 \\ .39 & .29 \end{pmatrix}$$

$$\mathbf{H}^{(4)} = \begin{pmatrix} .33 & .30 \\ .27 & .36 \end{pmatrix}$$

$$\mathbf{H}^{(5)} = \begin{pmatrix} .30 & .31 \\ .29 & .33 \end{pmatrix}$$

$$\mathbf{H}^{(8)} = \begin{pmatrix} .29 & .32 \\ .29 & .32 \end{pmatrix}$$

By iteration,

$$\mathbf{H}^{(\infty)} = \begin{pmatrix} .29 & .32 \\ .29 & .32 \end{pmatrix}$$

4. EXISTENCE OF A STEADY-STATE PROBABILITY UNDER NON-ADDITIVITY

The following theorem demonstrates that exists a steady-state probability for a finite Markov chain if the transition probability is strictly monotonic. $v_A(\cdot)$ is strictly monotonic if $B \subset$ (proper subset) C, $B, C \subseteq \Omega$, implies $v_A(B) < v_A(C)$. Note that without uncertainty a steady-state probability exists if the transition probability is strictly positive.

Theorem. Let $\mathbf{H}$ be a (possibly non-additive) transition probabilities matrix. If Ω is finite and $v_i(\cdot)$ is strictly monotonic for $i \in \Omega$, then there exists a (possibly non-additive) probability π on Ω such that

$$\pi(j) = lim_{n \to \infty} v_i^{(n)}(j)$$

Proof. Define

$$z_{ij}^{(n)}(k) = v_i(\{s : q_j^{(n)}(s) \geq q_j^{(n)}(k)\}) - v_i(\{s : q_j^{(n)}(s) > q_j^{(n)}(k)\}), \quad i, j, k \in \Omega.$$

Notice that $\sum_{k \in \Omega} z_{ij}^{(n)}(k) = 1$ and

$$v_i^{(n)}(j) = \int_0^\infty v_i(\{s : q_j^{(n-1)}(s) \geq \alpha\}) d\alpha = \sum_{k \in \Omega} z_{ij}^{(n-1)}(k)\ v_k^{(n-1)}(j).$$

Define $\mathbf{Z} = \{v_i(A \cup k) - v_i(A) : A \subseteq \Omega - k,\ i, k \in \Omega\}$. Then $z_{ij}^{(n)}(k) \in \mathbf{Z}$, $n = 1, 2, \ldots$, and since $\mathbf{Z}$ is finite and $v_i(\cdot)$ is strictly monotonic, there exists δ such that $\delta = min_{i,j,k \in \Omega}\{\mathbf{Z}\} > 0$.

Fix arbitrary i, h and j.

Define $Q(n) = \{k \in \Omega : z_{ij}^{(n)}(k) > z_{hj}^{(n)}(k)\}$ and $S(n) = \{k \in \Omega : z_{ij}^{(n)}(k) \leq z_{hj}^{(n)}(k)\}$.

Because $\sum_{k \in \Omega} z_{ij}^{(n)}(k) = \sum_{k \in \Omega} z_{hj}^{(n)}(k) = 1$,

$$\sum_{k \in Q(n)} z_{ij}^{(n)}(k) + \sum_{k \in S(n)} z_{ij}^{(n)}(k) = \sum_{k \in Q(n)} z_{hj}^{(n)}(k) + \sum_{k \in S(n)} z_{hj}^{(n)}(k)$$

and

$$\sum_{k\in Q(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] = -\sum_{k\in S(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] \tag{4.1}$$

And also,

$$\begin{aligned}&\sum_{k\in Q(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] \\ &\quad = 1 - \sum_{k\in S(n)} z_{ij}^{(n)}(k) - \sum_{k\in Q(n)} z_{hj}^{(n)}(k) \leq 1 - N\delta.\end{aligned} \tag{4.2}$$

Let $M_j^{(n)}$ and $m_j^{(n)}$ be the maximum and minimum of the elements of the jth column in $\mathbf{H}^{(n)}$.

Then

$$M_j^{(1)} - m_j^{(1)} \leq 1 - (N-1)\delta - \delta \leq 1 - N\delta. \tag{4.3}$$

$$\begin{aligned}&v_i^{(n+1)}(j) - v_h^{(n+1)}(j) \\ &= \sum_{k\in\Omega} z_{ij}^{(n)}(k) v_k^{(n)}(j) - \sum_{k\in\Omega} z_{hj}^{(n)}(k) v_k^{(n)}(j) \\ &= \sum_{k\in\Omega} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] v_k^{(n)}(j) \\ &= \sum_{k\in Q(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] v_k^{(n)}(j) + \sum_{k\in S(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] v_k^{(n)}(j) \\ &\leq \sum_{k\in Q(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] M_j^{(n)} + \sum_{k\in S(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)] m_j^{(n)} \\ &= \sum_{k\in Q(n)} [z_{ij}^{(n)}(k) - z_{hj}^{(n)}(k)][M_j^{(n)} - m_j^{(n)}] \quad \text{by} \quad (4.1) \\ &\leq [M_j^{(n)} - m_j^{(n)}][1 - N\delta] \quad \text{by} \quad (4.2)\end{aligned} \tag{4.4}$$

(4.3) and (4.4) yields

$$v_i^{(n+1)}(j) - v_h^{(n+1)}(j) \leq [1 - N\delta]^{(n+1)}. \tag{4.5}$$

It is easy to see that $M_j^{(n)}$ and $m_j^{(n)}$ are weakly decreasing($M_j^{(n)}$ is weakly decreasing(increasing) if $M_j^{(n)} \leq (\geq) m_j^{(n)}$.) and weakly increasing functions, respectively. $M_j^{(n)}$ and $m_j^{(n)}$ are bounded above by 1 and below by 0, and therefore both have the same limit, say $\pi(j)$. Since $m_j^{(n)} \leq \pi(j) \leq M_j^{(n)}$, $m_j^{(n)} - M_j^{(n)} \leq v_i^{(n)}(j) - \pi(j) \leq M_j^{(n)} - m_j^{(n)}$, which, together with (4.5), completes the proof.

To calculate $E_\pi(X)$ for a random variable X, we also need $\pi(A)$, $A \subseteq \Omega$. Let $v(\cdot)$ be the initial probability, and let $v(B|A)$ be the conditional probability of the event B given A. There are two updating rules – (i) Bayes' rule, (ii) Dempster-Shafer rule.

(i) Bayes' rule : $v(B|A) = v(A \cap B)/v(A)$ for all $A \neq \Phi$.

(ii) Dempster-Shafer rule : $v(B|A) = [v(B \cup A^c) - v(A^c)]/[1 - v(A^c)]$ for all $A \neq \Phi$.

The former is commonly used in the case of additivity. The latter is a pessimistic decision rule assuming that all the unoccurred events are the best possible ones. The Dempster-Shafer rule is usually used since decision makers are in general averse to uncertainty.

Corollary. Under the same conditions in the above theorem with anyone of two updating rules, there exists a (possibly non-additive) probability π on Ω such that

$$\pi(B) = lim_{n \to \infty} v_A^{(n)}(B), \quad A, B \subseteq \Omega.$$

Proof. $v_A(B)$, $A, B \subseteq \Omega$, is defined as

$$v_A(B) = \int_0^\infty v(\{s : q_B(s) \geq \alpha\}|A) d\alpha. \tag{4.6}$$

Using the same method in the proof of the above theorem completes the proof.

Without uncertainty, it is well known that $\pi(\cdot)$ is additive and unique, and $\pi(\cdot)$ satisfies

$$\pi(j) = \int_0^\infty \pi(\{s : q_j(s) \geq \alpha\}) d\alpha.$$

where π is the steady-state probability. In the presence of uncertainty, however, $\pi(\cdot)$ does not hold the above properties.

Therefore $E_\pi(X)$ may not be unique. The following example shows that $\pi(\cdot)$ may not be additive and unique.

Example. Suppose that $v_i(2 \cup 3) = 0.6$ and $v_i(1 \cup 2) = 0.7$, $i = 1, 2, 3$, $v(2) = v(3) = 0.2$, $v(2 \cup 3) = 0.4$ and that transition probabilities matrix is

$$H = \begin{bmatrix} 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.2 \end{bmatrix}$$

Then by calculation, $\pi(1) = 0.17$ and $\pi(2) = \pi(3) = 0.2$. By (4.6) and the assumptions, using Bayes' rule, $v_1(1) = 0.2$, $v_{2\cup 3}(1) = 0.15$ and $v_1(2 \cup 3) = v_{2\cup 3}(2 \cup 3) = 0.6$. Thus we obtain a new transition probabilities matrix :

$$H = \begin{bmatrix} 0.2 & 0.6 \\ 0.15 & 0.6 \end{bmatrix}$$

Iteration yields $\pi'(1) = 0.158$ and $\pi'(2 \cup 3) = 0.6$. It is easy to see that $\pi'(2 \cup 3) \neq \pi(2) + \pi(3)$ and also $\pi'(1) \neq \pi(1)$. A more complicated example can show $\pi(j) \neq \int_0^\infty \pi_i(\{s : q_j(s) \geq \alpha\})d\alpha$.

5. CONCLUSION

In this paper, I have shown in the presence of uncertainty a steady-state probability exists. However, the steady-state probability is not unique. Therefore, a unique long-run expectation may not exist under uncertainty, while without uncertainty it does. This implies that a long-run expectation provides rougher idea under uncertainty than under risk.

REFERENCES

(1) Bewley, T.(1986). *Knightian Decision Theory : Part I*, Yale University, working paper.

(2) Bewley, T.(1989). *Market Innovation and Entrepreseneurship : Knightian View*, Yale University, working paper.

(3) Choquet, G.(1954). Theory of Capacities, *Annals de l'Institut Fourier* **5**, 131-295.

(4) Dempster, A.(1968). A Generalization of Bayesian Inference, *Journal of Royal Statistics Society, Series B*, **30**, 205-247.

(5) Ellsberg, D.(1961). Risk, Ambiguity and the Savage Axioms, *Quarterly Journal of Economics*, 643-669.

(6) Gilboa, I.(1987). Expected Utility Theory with Purely Subjective Non-additive probabilities, *Journal of Mathematical Economics*, 141-153.

(7) Gilboa, I. and Schmeidler, D.(1989). Maxmin Expected Utility with a Non-unique Prior, *Journal of Mathematical Economics*, **16**, 141-153.

(8) Gilboa, I. and Schmeidler, D.(1991). *Updating Ambiguous Beliefs*, Northwestern University, working paper.

(9) Knight, F.(1921). *Risk, Uncertainty and Profit*, Houghton Mifflin, Boston & New York.

(10) Savage, L.(1954). *The Foundation of Statistics*, Wiley, New York.

(11) Schmeidler, D.(1989). Subjective Probability and Expected Utility without Additivity, *Econometrica*, **57**, 571-587.

(12) Shafer, G.(1976). *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, N. J.

(13) Wakker, P.(1989). *Additive Representations of Preferences, A New Foundation of Decision Analysis*, Kluwer Academic Publishers, Dordrecht.

(14) Walley, P.(1991). *Statistics Reasoning with Imprecise Probabilities*, Chapman and Hall, London.

(15) Yoo, K.(1991). The Iterative Law of Expectation and Non-additive Probability Measure, *Economics Letters*, October, 145-149.