Economics Bulletin

Volume 36, Issue 4

Optimization over a collection of decision trees with three-valued outcomes

Lonnie Turpin
McNeese State University

Matiur Rahman
McNeese State University

Alberto Marquez
Lamar University

Abstract

This note considers decision trees with three-valued outcomes. The structure of the trees are represented in a familiar form, allowing for actions and states of nature where the states of nature are associated with objective probabilities. We discuss the partitioning of trees by path enumeration, and present a simple formula for calculating the probabilities of outcomes. Finally, we construct a linear programming model to optimize over the given probabilities to select the optimal partition tree representing the collection of actions that minimizes the potential for loss.

Citation: Lonnie Turpin and Matiur Rahman and Alberto Marquez, (2016) "Optimization over a collection of decision trees with three-valued outcomes", *Economics Bulletin*, Volume 36, Issue 4, pages 1959-1965
Contact: Lonnie Turpin - lturpin@mcneese.edu, Matiur Rahman - mrahman@mcneese.edu, Alberto Marquez - alberto.marquez@lamar.edu.
Submitted: October 04, 2016. **Published:** November 09, 2016.

1. Introduction

In the literature, loss aversion refers to the avoidance of outcomes with losses (as in Kobberling and Wakker (2005); Nygren et al. (1996); Tversky and Kahneman (1992), among others). A decision maker may prefer a lottery with lower expected value and higher variance over a lottery with higher expected value and lower variance but that involves an outcome with a loss in a non-conformable situation.

Define ψ as a three-valued sign associated with an outcome where $\psi \in \{-, 0, +\}$. Let us assume that a loss averse person seeks to identify the probability of loss associated with a decision by categorizing outcomes as negative ($\psi = (-)$), positive ($\psi = (+)$), or neutral ($\psi = (0)$). Consider the environment represented as a decision tree where a decision, or sequence of decisions, leads to a loss. This outcome is therefore represented as a negative sign ($\psi = (-)$) implying the potential for loss involved in that decision or sequence of decisions. By definition, the decision maker would avoid decisions where the outcome may result in a loss.

Now suppose there is no possible way to avoid a loss, yet there exists known probability values associated with each ψ. Call an all-encompassing decision tree environment the "main tree", where a smaller target tree is realized by first decomposing the decision tree into smaller trees via partitioning, then optimizing over the corresponding probabilities. This research presents a method for calculating the probability of a loss, and selecting among the enumerated sub-trees where the goal is to determine the tree (and thereby, the collection of decisions) that minimizes the potential for loss.

2. Scenario

We begin with a few basic definitions of decision trees. Although this research is self-contained, we note that classic fundamentals of decision trees are outlined in Raiffa (1968). When facing a decision tree, a decision maker makes a series of one or more choices to reach an outcome. Following Savage (1951), let A and S be finite sets representing alternatives and states of nature, respectively. The set A contains courses of action $a \in A$ that the decision maker has complete control over, and therefore have no probabilities. The set S has elements $s \in S$ representing states of the world that the decision maker has no control over, and therefore have probabilities.

Remark 1. *For exposition, we consider objective probabilities, i.e., probabilities that are objectively given with the decision tree rather assigned by the decision maker.*

To define the structure of the tree, let D represent the set of decision nodes, C represent the set of chance nodes, and E represent the set of end nodes where $D \subseteq A$ and $C \subseteq S$. Let a node $n \in C, D, E$.

Remark 2. *Define the Cartesian product space*

$$A \times S = \{(a, s) \mid a \in A, s \in S\}$$

with elements $(a, s) \in A \times S$ representing the ordered pairs of alternative a in state of nature s. To realize Remark 1, for any a, the elemental combination (a, s) has a probability $\pi((a, s)) \in [0, 1] \subseteq \mathbb{R}$ where $\sum_s \pi((a, s)) = 1$ axiomatically, given the probability is associated with state s only.

To indicate an order relation on the tree, consider any node n in the presence of preceding nodes p. The existence of any n directly implies the existence of its predecessors, indicating that a node cannot be analyzed unless its preceding nodes are included as well. We say the initial node is always a decision node $n \in D$ from which all following nodes stem, and the final nodes are end nodes $n \in E$. Let any ψ occur only at an end node $n \in E$.

Remark 3. *Following Remark 2, we let the probability measure $\pi(\{-, 0, +\}) = 1$ over the space $\{-, 0, +\}$ where $\pi(\psi) \in [0, 1]$ and $\sum_\psi \pi(\psi) = 1$. It follows immediately that if $\pi > 0$ the interval $(0, \pi)$ will be endowed with the Lebesgue measure $\mu((0, \pi)) = \pi$.*

Let K represent a finite set of paths with elements $k \in K$. Define a path $k = (n \mid p_n)$ recursively from $n \in E$. Here, the vertical bar $\mid$ represents the familiar conditioning of one node on another. Then, the probability of path $k \in K$ is the conditional product chain

$$\pi(k) = \prod_n \pi(n \mid p_n).$$

Finally, let an indicator function $\mathbf{1}_\psi$ be defined by $\mathbf{1}_\psi(k) \in \{0, 1\}$ where $\mathbf{1}_\psi(k) = 1$ if k leads to outcome sign ψ and zero, otherwise.

With the structure of the tree now understood, consider a collection of complete trees (with paths from the initial node to the end nodes) within the "main tree". Without loss of generality, we partition T by full path enumeration into $|T|$ trees. These individual trees $t \in T$ reflect the possible partitions of the main tree.

Remark 4. *Denote by $K(t)$ the set of all paths in a tree t. Let $i, j \in T$. For exposition, fix any $K(i) \cap K(j) = \emptyset$ where $i \neq j$. It follows directly that the concatenation of the paths $\langle K(T) \rangle$ admits the cardinal equivalence $|\langle K(T) \rangle| = |K|$.*

We propose an optimization model to search through the space of decision trees, and restrict the decision space to a single tree. Following the motivation of this research, the goal is to determine the tree that minimizes the potential for loss.

Remark 5. *The decision maker takes on only one outcome. We say the mapping $a\colon \text{decision maker} \to T$ represents an action, and the variable $a_t \in \{0, 1\}$ where $a_t = 1$ if action a is taken for tree t and zero, otherwise.*

Define the set of all actions as $\{a_t \mid t \in T\} \in \mathbb{Z}_2^{|T|}$ expressed simply as $\{a_t\}$. For readability, we let $\|\{a_t\}\|$ be the length of the vector $\{a_t\}$ where the cardinal equivalence $|T| = \|\{a_t\}\|$ holds.

3. Results

In traditional static optimization, the goal is to determine a single value for each decision variable, such that the objective function will be maximized or minimized. For a dynamic setting, time is considered and we encounter a dynamic optimization problem, often referred to as an intertemporal optimization problem (Walde, 2008). In this type of multi-period problem, we need to determine the optimal time-variant path of decisions and states over a planning period, represented here in a decision tree form. This multi-step decision process can be seen throughout literature in portfolio management (Costa and Nabholz, 2007; Chellathurai and Draviam, 2008), consumption-investment (Fama, 1970), and consumer spending decisions (Soman and Cheema, 2002), among others. However, in this research we are not narrowly deciding on the optimal path, instead we are focusing on the optimal sub-tree.

Remark 6. *The objective nature of the probabilities by Remark 1 combined with the nature of the decision variables allows for a static linear programming formulation of an otherwise intertemporal optimization problem.*

Theorem 1. *The goal can be formally defined as the following linear program.*

$$
\begin{array}{lrcll}
\min & \displaystyle\sum_t \pi(\psi)_t \, a_t & & & \\
\text{s.t.} & \pi(\psi)_t & = & \dfrac{\sum_k \pi(k)\,\mathbf{1}_\psi(k)}{\sum_k \pi(k)}, & \text{for all} \quad k \in K, t \in T \\
& \pi(k)_t & = & \displaystyle\prod_n \pi(n \mid p_n), & \text{for all} \quad k \in K, t \in T \\
& \mathbf{1}_\psi(k)_t & \in & \{0,1\}, & \text{for all} \quad k \in K, t \in T \\
& \pi(n)_t & \in & [0,1], & \text{for all} \quad n \in C, t \in T \\
& \pi(n)_t & = & 1, & \text{for all} \quad n \in D, t \in T \\
& \displaystyle\sum_t a_t & = & 1, & \text{for all} \quad t \in T \\
& a_t & \in & \{0,1\}, & \text{for all} \quad t \in T
\end{array}
$$

Proof. Fix $\psi = (-)$. Using the probability variables as objective coefficients directly implies the objective $\min \sum_t \pi(\psi)_t a_t$. Adjust Remark 3 by letting $\pi(\psi)_t > 0$ for all t to allow for a solution expressed in the form $\sum_t a_t$. Though forcing the solution set $\sum_t a_t \geq 1$ would achieve robustness, to satisfy Remark 5 we use $\sum_t a_t = 1$ since the objective $\min \pi(\psi)_t a_t$ together with $\sum_t a_t = 1$ makes the constraint $\sum_t a_t \geq 1$ redundant. With the additional constraints obvious, it only remains to show that

$$\pi(\psi) = \frac{\sum_k \pi(k)\,\mathbf{1}_\psi(k)}{\sum_k \pi(k)}$$

for every k. The recursion starts with the terminal node $n \in E$ where there exists a ψ such that the indicator function $\mathbf{1}_\psi(k) \in \{0,1\}$ for every k. Let $|A| > 1$ for any $n \in D$ such that S for any alternative is collectively exhaustive implying $\pi(S) = 1$ by Remark 2. Then for any tree t, the probability over K may be $\pi(K) \neq 1$ forcing the denominator $\sum_k \pi(k) \neq 1$ and requiring the derivation of $\pi(\psi)$ to take the form of a weighted average. See the Appendix for clarification of $\pi(K) \neq 1$. Without loss of generality, fix $\pi(n) = 1$ for every $n \in D$ to preserve the calculation of $\pi(k)$, and the result follows easily by backward recursion. $\square$

Corollary 1. *Though a necessary condition for logical structure, excluding Remark 4 would have no mathematical impact on the result in Theorem 1. The proof of this notion is obvious.*

References

Chellathurai, T., & Draviam, T. (2008, April). "Markowitz principles for multi-period portfolio selection problems with moments of any order." *In Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 464, No. 2092, pp. 827-854). The Royal Society.

Costa, O. L. V., & Nabholz, R. B. (2007). "Multiperiod mean-variance optimization with intertemporal restrictions." *Journal of Optimization Theory and Applications*, 134(2), 257-274.

Fama, E. F. (1970). "Multiperiod consumption-investment decisions." *The American Economic Review*, 60(1), 163-174.

Kobberling, V., & Wakker, P. P. (2005). "An index of loss aversion." *Journal of Economic Theory*, 122(1), 119-131.

Nygren, T. E., Isen, A. M., Taylor, P. J., & Dulin, J. (1996). "The influence of positive affect on the decision rule in risk situations: Focus on outcome (and especially avoidance of loss) rather than probability." *Organizational Behavior and Human Decision Processes*, 66(1), 59-72.

Raiffa, H. (1968). *Decision analysis: introductory lectures on choices under uncertainty*, Addison-Wesley.

Savage, L. J. (1951). "The theory of statistical decision." *Journal of the American Statistical Association*, 46(253), 55-67.

Soman, D., & Cheema, A. (2002). "The effect of credit on spending decisions: The role of the credit limit and credibility." *Marketing Science*, 21(1), 32-53.

Tversky, A., & Kahneman, D. (1992). "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 5(4), 297-323.

Walde, K. (2008). *Applied Intertemporal Optimization*. Books. Business School-Economics, University of Glasgow.

Appendix

Consider the sample decision tree in Figure 1 where we represent the decision nodes as $\square$, the chance nodes as $\bigcirc$, and the end nodes as $\triangle$ with the probabilities $\pi(n)$ labeled on each branch where $n \in C, D$. Full path enumeration for the tree in Figure 1 is shown in Table 1.

Path k	Description	$\pi(k)$		Sign ψ	$\mathbf{1}_{\psi}(k)$
1	$D_1 \to S_1 \to D_2 \to S_3 \to E_1$	$1 \times 0.4 \times 1 \times 0.2$	$= 0.080$	$(-)$	1
2	$D_1 \to S_1 \to D_2 \to S_3 \to E_2$	$1 \times 0.4 \times 1 \times 0.5$	$= 0.200$	$(+)$	0
3	$D_1 \to S_1 \to D_2 \to S_3 \to E_3$	$1 \times 0.4 \times 1 \times 0.3$	$= 0.120$	(0)	0
4	$D_1 \to S_1 \to D_2 \to S_4 \to E_4$	$1 \times 0.4 \times 1 \times 0.7$	$= 0.280$	$(-)$	1
5	$D_1 \to S_1 \to D_2 \to S_4 \to E_5$	$1 \times 0.4 \times 1 \times 0.3$	$= 0.120$	$(+)$	0
6	$D_1 \to S_1 \to E_6$	1×0.3	$= 0.300$	(0)	0
7	$D_1 \to S_1 \to E_7$	1×0.3	$= 0.300$	$(+)$	0
8	$D_1 \to D_3 \to E_8$	1×1	$= 1.000$	$(-)$	1
9	$D_1 \to D_3 \to S_5 \to E_9$	$1 \times 1 \times 0.5$	$= 0.500$	$(+)$	0
10	$D_1 \to D_3 \to S_5 \to E_{10}$	$1 \times 1 \times 0.25$	$= 0.250$	$(-)$	1
11	$D_1 \to D_3 \to S_5 \to E_{11}$	$1 \times 1 \times 0.25$	$= 0.250$	$(-)$	1
12	$D_1 \to S_2 \to E_{12}$	1×0.45	$= 0.450$	$(-)$	1
13	$D_1 \to S_2 \to E_{13}$	1×0.25	$= 0.250$	$(+)$	0
14	$D_1 \to S_2 \to D_4 \to S_6 \to E_{14}$	$1 \times 0.3 \times 1 \times 0.25$	$= 0.075$	$(-)$	1
15	$D_1 \to S_2 \to D_4 \to S_6 \to E_{15}$	$1 \times 0.3 \times 1 \times 0.75$	$= 0.225$	(0)	0

Table 1: Path Enumeration of Figure 1

Fix the partition for T into $|T| = 3$ trees where

- $t = 1 \equiv \{k = 1, k = 2, k = 3, k = 4, k = 5, k = 6, k = 7\}$
- $t = 2 \equiv \{k = 8, k = 9, k = 10, k = 11\}$
- $t = 3 \equiv \{k = 12, k = 13, k = 14, k = 15\}$

and the remaining metrics are shown in Table 1. We first make clear the point that the probability $\pi(K) \neq 1$ by

$$\begin{aligned}\pi(K) &= \sum_{k=1}^{15} \pi(k) \\ &= 0.08 + 0.2 + 0.12 + 0.28 + \ldots + 0.225 \\ &= 4.4.\end{aligned}$$

In accordance with the goal, fix $\psi = (-)$. The probability $\pi(\psi = (-))_T$ representing the loss coefficient

$$\begin{aligned}\pi\left(\psi=(-)\right)_{\sum_{t=1}^{3}} &= \frac{\sum_{k=1}^{15}\pi\left(k\right)\mathbf{1}_{\psi=(-)}\left(k\right)}{\sum_{k=1}^{15}\pi\left(k\right)}\\ &= \frac{0.08\left(1\right)+0.2\left(0\right)+0.12\left(0\right)+0.28\left(1\right)+\ldots+0.225\left(0\right)}{0.08+0.2+0.12+0.28+\ldots+0.225}\\ &= \frac{2.385}{4.4}\\ &= 0.542045\end{aligned}$$

follows easily. By Remark 4, $K\left(t=1\right)\cap K\left(t=2\right)\cap K\left(t=3\right)=\emptyset$. For $t=1$, we calculate

$$\begin{aligned}\pi\left(\psi=(-)\right)_{t=1} &= \frac{\sum_{k=1}^{7}\pi\left(k\right)\mathbf{1}_{\psi=(-)}\left(k\right)}{\sum_{k=1}^{7}\pi\left(k\right)}\\ &= \frac{0.08\left(1\right)+0.2\left(0\right)+0.12\left(0\right)+0.28\left(1\right)+0.12\left(0\right)+0.3\left(0\right)+0.3\left(0\right)}{0.08+0.2+0.12+0.28+0.12+0.3+0.3}\\ &= \frac{0.36}{1.4}\\ &= 0.257143.\end{aligned}$$

With $\pi\left(\psi=(-)\right)_{t=1}\approx 0.26$, applying this logic to $t=2$ and $t=3$, we get $\pi\left(\psi=(-)\right)_{t=2}=0.75$ and $\pi\left(\psi=(-)\right)_{t=3}\approx 0.53$. By Theorem 1, the resulting solution set $\{a_{t=1}=1, a_{t=2}=0, a_{t=3}=0\}$ yields the objective function value $\min\sum_{t}\pi\left(\psi=(-)\right)_{t}a_{t}\approx 0.26$.

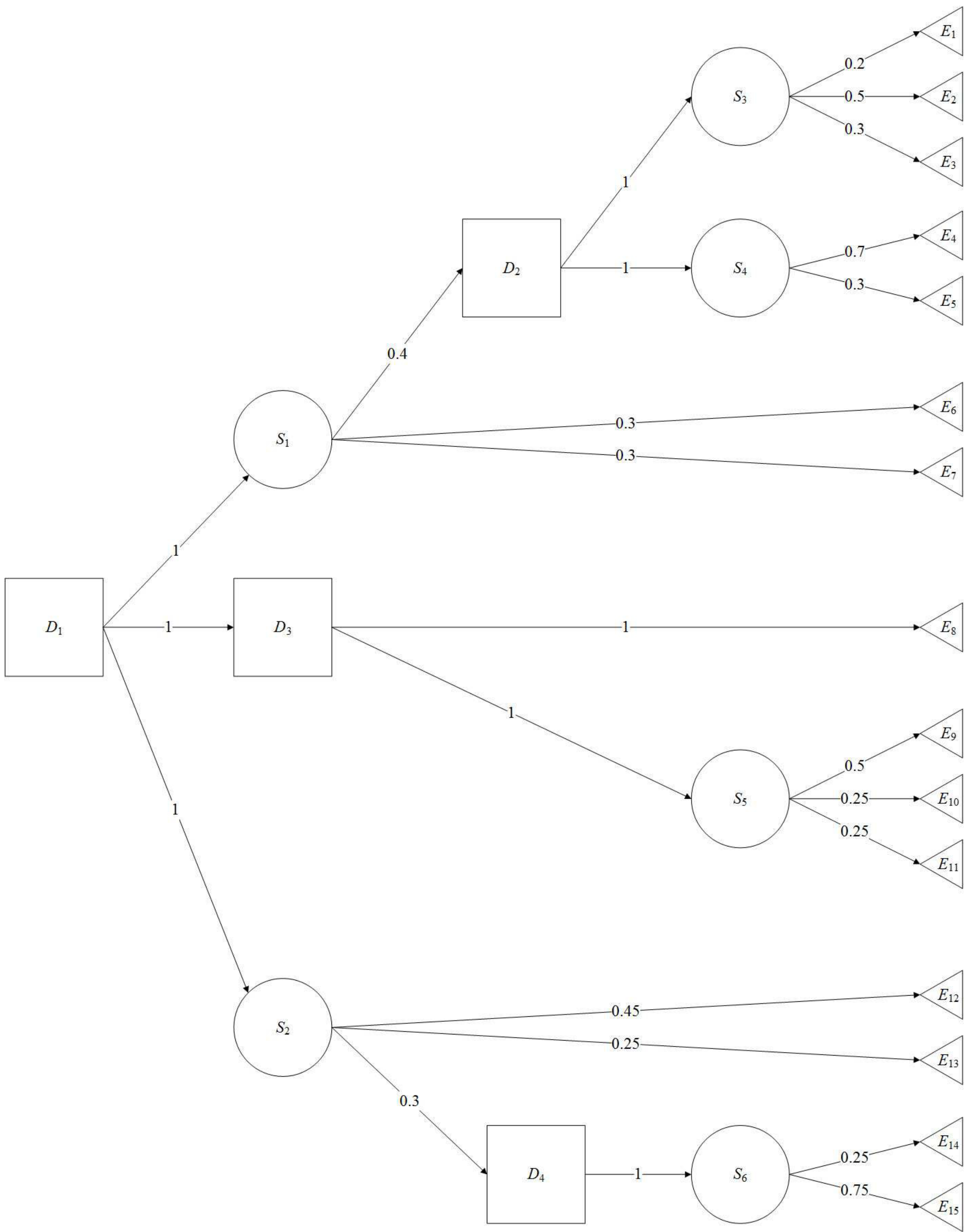

Figure 1: Sample Decision Tree