Decision Analysis Module
(2 credits)

Tutor: Dave Weimer

This module provides a self-paced introduction to decision analysis, the basic framework for evaluating alternatives in situations involving risk. Your first goal is to master the basic components of decision analysis: constructing decision trees, using Bayes= theorem to complete the specification of probabilities, assigning monetary or utility values to outcomes, solving for optimal decisions through backward-induction, and determining the sensitivity of the solution to changes in assumptions. Your second goal is to become familiar with computer software (DATA 3.0 by TreeAge) for implementing decision analysis. Finally, your third goal is to move beyond stylized illustrations to practice the application of what you have learned to an actual decision problem.

Your grade in the module will be based on two products. The first product is a journal, including DATA files, that gives your answers to questions and solutions to problems posed to you as you work through the basics of decision analysis. You should be able to explain the contents of your journal clearly if called upon to do so. The second product, which will become an appendix to this module, is a project that applies decision analysis to a problem that interests you. The most apparent applications are likely to be in the areas of health, safety, and environmental policy, but you are not limited to these areas. Your project will be judged on the appropriateness of the decision analysis framework for the problem you have selected, your diligence in making use of available evidence to inform probabilities, payoffs, and other model inputs, your skill in interpreting your results, and the quality of your presentation. The actual report should be concisely written and submitted in WordPerfect format with accompanying DATA files as appropriate.

Most of the readings are from Howard Raiffa, ***Decision Analysis*** (Reading, MA: Addison-Wesley, 1968).

Fundamentals

1. Introduction: Decision Trees

Read the following:

Raiffa, Chapters 0, 1, and up to page 14 in Chapter 2.

David L. Weimer and Aidan R. Vining, ***Policy Analysis: Concepts and Practice*** (Englewood Cliffs, NJ: Prentice-Hall, 1992), Chapter 12: ABenefit-Cost Analysis in Bureaucratic Settings: The Strategic Petroleum Reserve"

Exercise 1: As a hands-on introduction to DATA, the following steps walk you through constructing and solving the e_0 branch of Raiffa=s AYour Basic Problem.@

1. Open the DATA 3.0 window.

2. Name the branch by typing AE0 Branch@ in the box at the upper left of the worksheet.

3. Position your cursor on the shaded square at the bottom right corner of the box and double click with your mouse. You should now have two branches coming from the box.

4. Position your cursor on the upper circle and single click. Type Asigma 1" in the box that appears. Repeat with lower circle and type Asigma 2" in the box.

5. Go back to the upper circle and double click to get the next part of the tree. Repeat for lower circle.

6. Click on the upper most of the four circles to get a box. Type Aurn 1" to label the branch then hit tab on your keyboard to move below the branch. Now type the probability for that branch, 0.8. Move to the second of the four circles and single click. Type Aurn 2" and then hit tab to position cursor below the branch. Now type 0.2. Repeat this process so that the lower branch looks like the upper branch. Note: As the probabilities assigned to the branches coming from a change node must sum to one, you can have complementary probabilities automatically calculate for you if you like. To do so follow the sequence **Edit|Preferences|Other_Calc_Prefs** and put a check in the box next to ACalculate complementary probabilities automatically.@

7. Go back to the upper most of the four circles and click the icon on the tool bar that shows a square, triangle, and circle, the symbols used to describe different types of decision nodes. Now click on the triangle, to change the node from a chance note (the default drawn by the computer when you create branches) to a terminal node. When you do so, a menu will appear with spaces to provide payoffs. Enter A40" under payoff 1. This is how much you will receive if you guess urn type 1 and your guess turns out to be correct. Repeat the process for the next three end nodes, entering A-20,@ A-5,@ and A100" to complete the specification of payoffs.

8. Before you solve the tree, you might want to check your numeric format to make sure that we will print decimal places. Follow the sequence **Edit|Numeric_Preferences** and type A2" in the box following ADecimal places.@ If you did not calculate complementary probabilities, then you should check them as follows: **Analysis|Verify_Probabilities**.

9. You can solve the tree as follows: **Analysis|Roll_Back**. This will solve the tree. Note that values in parentheses are negative numbers. Also note that the sub-optimal path is pruned and the expected value of the optimal strategy, sigma 1, is $28 dollars. To return to the unsolved tree, which could be modified if necessary, simply use **Analysis|Roll_Back** again.

10. Save your tree as follows: **File|Save** (or click save icon) and use window to name it and place it in your directory.

2. Bayesian Updating

Read the following:

Raiffa, Chapter 2, pages 14 to 38

Exercise 2: Consider a disease whose treatment costs $20,000 once symptoms appear. If a prophylactic costing $1,500 is administered prior to the onset of symptoms, then treatment costs only $10,000. (The treatment modalities are equivalent from the patient's perspective.) A test has been developed to help identify those who have the disease. It has a false positive rate of 20 percent and a false negative rate of 30 percent. The fraction of people in the target population who have the disease is 10 percent.

You have been asked by a hospital to advise them about the course of treatment that has the lowest expected cost. Design an optimal testing-treatment strategy for the hospital under that assumption that the test costs $100. What is the maximum price that the hospital should be willing to pay for the test?

Using pencil and paper, set up a decision tree for this problem and determine the optimal strategy for the hospital. After you have solved the tree by hand, implement it in DATA as follows:

1. The initial decision node has three branches: wait to see if symptoms appear; treat with the prophylactic before symptoms appear; and administer the test before symptoms appear to gain more information. To set up three branches in DATA, first double click on lower right corner of box to get two branches, and then use **Options|Add_Branch** to add the third. The waiting and treating branches are followed directly by chance nodes with branches for Adisease@ and Ano disease.@ Testing gives either a positive or negative result. Following each of these results is a decision node with the branches Atreat@ and Await.@ Each of these decisions is followed by a chance node with branches for disease and Ano disease.@

2. Let p and n stand for positive and negative test results, respectively. Let D and ND stand for Adisease@ and Ano disease,@ respectively. We are given the following information:

False positive rate: $P(p|ND) = .20$, which implies that $P(n|ND) = .80$
False negative rate: $P(n|D) = .30$, which implies that $P(p|D) = .70$
Frequency of disease in target population: $P(D) = .10$, which implies that $P(ND) = .90$

We use Bayes= Rule to compute the following:
$P(D|p) = P(p|D)P(D)/[P(p|D)P(D)+P(p|ND)P(ND)] = (.7)(.1)/[(.7)(.1)+(.2)(.9)] = .28$
Therefore, $P(ND|p) = .72$

$P(D|n) = P(n|D)P(D)/[P(n|D)P(D)+P(n|ND)P(ND)] = (.3)(.1)/[(.3)(.1)+(.8)(.9)] = .04$
Therefore, $P(ND|n) = .96$

Further, $P(p) = P(p|D)P(D)+P(p|ND)P(ND) = .25$ so that $P(n) = .75$.
Make sure that you understand the derivation of these probabilities.

3. If we put payoffs in terms of cost, then we will maximize expected monetary value by minimizing expected cost. In terms of DATA, we can indicate this through the following menu choices:
Edit|Preferences|Calculation Method|Optimal Path|Low

4. We can write payoffs at the terminal nodes as follows: Total_cost = Test_cost + Proph_cost + Dis_cost where Test_cost = $100 for scenarios involving tests, Proph_cost = $1,500 for scenarios involving prophylactic treatment, and Dis_cost = $20,000 for those who have the disease without prophylactic treatment and Dis_cost = $10,000 for those who have the disease with prophylactic treatment. One of the most useful, but somewhat tricky, features of DATA is the ability to define payoffs in terms of variables. The advantage of this is that you can then easily change payoffs by simply giving the variables new values. The procedure requires one to go through three steps.

First, create the variables that you will use. Click on the root brach then open **Values|Define_Values|New|Variable** then type Test_cost and OK**.** Repeat for Dis_cost, Proph_cost, and Total_cost.

Second, set the initial values of each of these variables to zero. Go to the root branch. Use **Values|Define_Values|Value|At_Selected_Node(s)**. Click on Test_cost and type the value 0, because we haven=t incurred any test costs yet. Repeat for Proph_cost and Dis_cost, assigning each one the value zero. DATA will treat the value of these variables as 0 until it reaches a branch where a different value is assigned. Now assign the value Total_cost =Test_cost +Proph_cost +Dis_cost. All four variables now have the value zero.

Third, chance the values of variables as appropriate at nodes. Move to the branch that indicates Atest.@ Use **Values|Define_Values|Value|At_Selected_Node(s)** to change the value of Test_cost to 100. Next move to one of the branches that involve prophylactic treatment. Depress the shift key and click the other nodes at which prophylactics are administered. Now use **Values|Define_Values|Value|At_Selected_Node(s)** to change the value of Proph_cost to 1500. Finally, use this procedure to change the values of Dis_cost at the terminal nodes to either 10000 or 20000 depending on whether or not the scenario involves prophylactic treatment.

Third, click on all the terminal nodes (using the shift key) and use **Values|Change_Active_Payoff(1)** to define payoffs as Total_cost.

5. After you find the optimal decision, use **Analysis|Roll Back** to go back to the unsolved tree. In order to investigate the question of how high the test cost would have to be before it should be abandoned as part of the optimal solution, we can move the mouse to the branch indicating the strategy to test and use **Analysis|Sensitivity_Analysis|One_Way.** With it, we can specify a range of, say, 0 to 1000 with 20 steps for Test_cost. We will have to answer several questions related to what happens at subsequent nodes C we only want the value of Test_cost to change at this node. After we have done this, DATA will print a graph that will allow us to determine optimal strategies at various values for Test_cost.

Exercise 3: Use DATA to solve Raiffa=s AYour Basic Problem.@

3. Expected Utility

Read:

Raiffa, Chapter 4.

M. R. Gold et al., AIdentifying and Valuing Outcomes,@ in Martha R. Gold, Joanna E. Siegel, Louise B. Russell, and Milton C. Weinstein, eds., ***Cost-Effectiveness in Health and Medicine*** (New York, NY: Oxford University Press, 1996), 82-134.

Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," ***Science*** 185:4157 (1974), 1124-1131.

W. Kip Viscusi, "Prospective Reference Theory: Toward an Explanation of the Paradoxes," ***Journal of Risk and Uncertainty*** 2:3 (1989), 235-264.

Exercise 4: On pages 34 to 36 Raiffa poses the Awildcatter problem@ in terms of expected monetary value. On pages 100 to 101 he provides information that you can use to construct a utility for money curve for the wildcatter. First, set-up and solve the wildcatter problem in terms of expected monetary value using DATA. Second, redo the analysis assuming that the wildcatter maximizes expected utility. Note that in constructing the wildcatter=s utility for money curve it is important to express utility in terms of total assets. You will have to use the given information to graph a few points and then interpolate for the particular monetary values that you need to convert to utilities. Be sure to specify the dollar values at all terminal nodes so that they include starting assets and all costs incurred during the scenarios that produce them. Convert these final dollar amounts to utilities that you enter as payoffs in your DATA decision tree. By the way, item c on page 100 often causes confusion C feel free to ignore it.

Exercise 5: Find some people who are willing to answer questions about their perceptions of various health conditions. Your objective is to try to elicit their utilities for the conditions by offering them lotteries in which one prize is Aperfect health@ and the other is Adeath.@ Assign a utility of one to Aperfect health@ and a utility of zero to Adeath.@ Make a list of between five and ten conditions that you expect to differ markedly in terms of utility. Carefully write out specifications of the conditions that you can read to your respondents. Do the utilities you elicit make intuitive sense? How consistent are they across individuals?

Exercise 6: Summarize Aprospective reference theory@ and discuss how effectively it resolves some of the apparent paradoxes of expected utility theory.

4. Sensitivity Analysis: Normal Form and Monte Carlo Analysis

Read:

Raiffa, Chapter 6

Anthony Boardman, David Greenberg, Aidan Vining, and David Weimer, ***Cost-Benefit Analysis: Concepts and Practice*** (Englewood Cliffs, NJ: Prentice Hall, 1996), Chapter 6: ADealing with Uncertainty: Expected Value and Sensitivity Analysis@

Exercise 7: Return to the problem presented in *Exercise 2*. Imagine that you were uncertain about the frequency of disease in the target population. Construct a normal form diagram for the problem and use it to determine the disease frequencies at which the optimal strategy for the hospital changes.

Exercise 8: Return to the problem presented in *Exercise 2*. Conduct a Monte Carlo simulation to get a sense of the distribution of outcomes under the optimal strategy. Move to the root branch and use **Analysis|Monte_Carlo_Simulation**. Set ANumber of Trials@ to 1000 and click ABegin.@ You can then view a histogram of the realized costs as well as obtain a text report. As the graph is not very effective in this case, and you cannot change it within DATA, you should export the text report to a spreadsheet where you can produce a more effective histogram.

Exercise 9: Conduct a Monte Carlo simulation for Raiffa=s AYour Basic Problem.@ Discuss the riskiness of the optimal strategy.

Exercise 10: Explore some of the other capabilities of DATA. Construct an example that demonstrates the use of one of these capabilities.

Final Project

You now have familiarity with the basic elements of decision analysis. It is time to try to apply it to a real problem through a project of your choice. An ideal project sets out a realistic problem for which decision analysis is an appropriate tool and takes full advantage of available information and empirical evidence to implement it. Candidate topics can be found in medicine, health and safety regulation, investment, and a variety of other fields. **Please consult with the tutor about your project choice**. If you are at all risk averse, then you should probably review your decision tree with the tutor early enough to allow yourself time to make substantial revisions if they are necessary.

Your report should include both a WordPerfect file with the text of your report and DATA files with your decision trees. These files will be made available under your name to future students who do this module.