16. Making Simple Decisions

16.1 [almanac-game](Adapted from David Heckerman.) This exercise concerns the **Almanac Game**, which is used by decision analysts to calibrate numeric estimation. For each of the questions that follow, give your best guess of the answer, that is, a number that you think is as likely to be too high as it is to be too low. Also give your guess at a 25th percentile estimate, that is, a number that you think has a 25% chance of being too high, and a 75% chance of being too low. Do the same for the 75th percentile. (Thus, you should give three estimates in all—low, median, and high—for each question.)

1. Number of passengers who flew between New York and Los Angeles in 1989.
2. Population of Warsaw in 1992.
3. Year in which Coronado discovered the Mississippi River.
4. Number of votes received by Jimmy Carter in the 1976 presidential election.
5. Age of the oldest living tree, as of 2002.
6. Height of the Hoover Dam in feet.
7. Number of eggs produced in Oregon in 1985.
8. Number of Buddhists in the world in 1992.
9. Number of deaths due to AIDS in the United States in 1981.
10. Number of U.S. patents granted in 1901.

The correct answers appear after the last exercise of this chapter. From the point of view of decision analysis, the interesting thing is not how close your median guesses came to the real answers, but rather how often the real answer came within your 25% and 75% bounds. If it was about half the time, then your bounds are accurate. But if you're like most people, you will be more sure of yourself than you should be, and fewer than half the answers will fall within the bounds. With practice, you can calibrate yourself to give realistic bounds, and thus be more useful in supplying information for decision making. Try this second set of questions and see if there is any improvement:

1. Year of birth of Zsa Zsa Gabor.
2. Maximum distance from Mars to the sun in miles.
3. Value in dollars of exports of wheat from the United States in 1992.
4. Tons handled by the port of Honolulu in 1991.
5. Annual salary in dollars of the governor of California in 1993.
6. Population of San Diego in 1990.
7. Year in which Roger Williams founded Providence, Rhode Island.
8. Height of Mt. Kilimanjaro in feet.
9. Length of the Brooklyn Bridge in feet.
10. Number of deaths due to automobile accidents in the United States in 1992.

16.2 Chris considers four used cars before buying the one with maximum expected utility. Pat considers ten cars and does the same. All other things being equal, which one is more likely to have the better car? Which is more likely to be disappointed with their car's quality? By how much (in terms of standard deviations of expected quality)?

16.3 Chris considers five used cars before buying the one with maximum expected utility. Pat considers eleven cars and does the same. All other things being equal, which one is more likely to have the better car? Which is more likely to be disappointed with their car's quality? By how much (in terms of standard deviations of expected quality)?

16.4 [St-Petersburg-exercise] In 1713, Nicolas Bernoulli stated a puzzle, now called the St. Petersburg paradox, which works as follows. You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads. If the first heads appears on the nth toss, you win 2^n dollars.

1. Show that the expected monetary value of this game is infinite.
2. How much would you, personally, pay to play the game?
3. Nicolas's cousin Daniel Bernoulli resolved the apparent paradox in 1738 by suggesting that the utility of money is measured on a logarithmic scale (i.e., $U(S_n) = a\log_2 n + b$, where S_n is the state of having \$$n$). What is the expected utility of the game under this assumption?
4. What is the maximum amount that it would be rational to pay to play the game, assuming that one's initial wealth is \$$k$?

16.5 Write a computer program to automate the process in Exercise assessment-exercise. Try your program out on several people of different net worth and political outlook. Comment on the consistency of your results, both for an individual and across individuals.
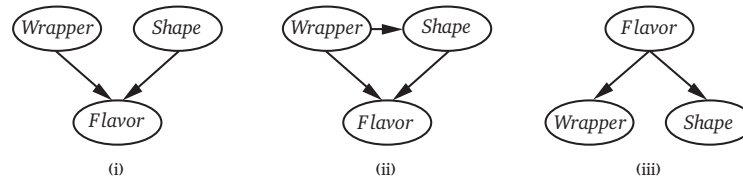
16.6 [surprise-candy-exercise] The Surprise Candy Company makes candy in two flavors: 75% are strawberry flavor and 25% are anchovy flavor. Each new piece of candy starts out with a round shape; as it moves along the production line, a machine randomly selects a certain percentage to be trimmed into a square; then, each piece is wrapped in a wrapper whose color is chosen randomly to be red or brown. 70% of the strawberry candies are round and 70% have a red wrapper, while 90% of the anchovy candies are square and 90% have a brown wrapper. All candies are sold individually in sealed, identical, black boxes.

Now you, the customer, have just bought a Surprise candy at the store but have not yet opened the box. Consider the three Bayes nets in Figure 3candy-figure.

1. Which network(s) can correctly represent $\mathbf{P}(Flavor, Wrapper, Shape)$?
2. Which network is the best representation for this problem?
3. Does network (i) assert that $\mathbf{P}(Wrapper|Shape) = \mathbf{P}(Wrapper)$?

4. What is the probability that your candy has a red wrapper?
5. In the box is a round candy with a red wrapper. What is the probability that its flavor is strawberry?
6. A unwrapped strawberry candy is worth s on the open market and an unwrapped anchovy candy is worth a. Write an expression for the value of an unopened candy box.
7. A new law prohibits trading of unwrapped candies, but it is still legal to trade wrapped candies (out of the box). Is an unopened candy box now worth more than less than, or the same as before?

Figure [3candy-figure] Three proposed Bayes nets for the Surprise Candy problem

16.7 [surprise-candy-exercise] The Surprise Candy Company makes candy in two flavors: 70% are strawberry flavor and 30% are anchovy flavor. Each new piece of candy starts out with a round shape; as it moves along the production line, a machine randomly selects a certain percentage to be trimmed into a square; then, each piece is wrapped in a wrapper whose color is chosen randomly to be red or brown. 80% of the strawberry candies are round and 80% have a red wrapper, while 90% of the anchovy candies are square and 90% have a brown wrapper. All candies are sold individually in sealed, identical, black boxes.

Now you, the customer, have just bought a Surprise candy at the store but have not yet opened the box. Consider the three Bayes nets in Figure 3candy-figure.

1. Which network(s) can correctly represent **P**(*Flavor, Wrapper, Shape*)?
2. Which network is the best representation for this problem?
3. Does network (i) assert that **P**(*Wrapper*|*Shape*) = **P**(*Wrapper*)?
4. What is the probability that your candy has a red wrapper?
5. In the box is a round candy with a red wrapper. What is the probability that its flavor is strawberry?
6. A unwrapped strawberry candy is worth s on the open market and an unwrapped anchovy candy is worth a. Write an expression for the value of an unopened candy box.
7. A new law prohibits trading of unwrapped candies, but it is still legal to trade wrapped candies (out of the box). Is an unopened candy box now worth more than less than, or the same as before?

16.8 Prove that the judgments $B \succ A$ and $C \succ D$ in the Allais paradox (page allais-page) violate the axiom of substitutability.

16.9 Consider the Allais paradox described on page allais-page: an agent who prefers B over A (taking the sure thing), and C over D (taking the higher EMV) is not acting rationally, according to utility theory. Do you think this indicates a problem for the agent, a problem for the theory, or no problem at all? Explain.

16.10 Tickets to a lottery cost 1. There are two possible prizes: a 10 payoff with probability 1/50, and a 1,000,000 payoff with probability 1/2,000,000. What is the expected monetary value of a lottery ticket? When (if ever) is it rational to buy a ticket? Be precise —show an equation involving utilities. You may assume current wealth of k and that $U(S_k) = 0$. You may also assume that $U(S_{k+10}) = 10 \times U(S_{k+1})$, but you may not make any assumptions about $U(S_{k+1,000,000})$. Sociological studies show that people with lower income buy a disproportionate number of lottery tickets. Do you think this is because they are worse decision makers or because they have a different utility function? Consider the value of contemplating the possibility of winning the lottery versus the value of contemplating becoming an action hero while watching an adventure movie.

16.11 [assessment-exercise]Assess your own utility for different incremental amounts of money by running a series of preference tests between some definite amount M_1 and a lottery $[p, M_2; (1 - p), 0]$. Choose different values of M_1 and M_2, and vary p until you are indifferent between the two choices. Plot the resulting utility function.

16.12 How much is a micromort worth to you? Devise a protocol to determine this. Ask questions based both on paying to avoid risk and being paid to accept risk.

16.13 [kmax-exercise] Let continuous variables $X_1, ..., X_k$ be independently distributed according to the same probability density function $f(x)$. Prove that the density function for $\max \{X_1, ..., X_k\}$ is given by $kf(x)(F(x))^{k-1}$, where F is the cumulative distribution for f.

16.14 Economists often make use of an exponential utility function for money: $U(x) = -e^{-x/R}$, where R is a positive constant representing an individual's risk tolerance. Risk tolerance reflects how likely an individual is to accept a lottery with a particular expected monetary value (EMV) versus some certain payoff. As R (which is measured in the same units as x) becomes larger, the individual becomes less risk-averse.

1. Assume Mary has an exponential utility function with $R = \$500$. Mary is given the choice between receiving $500 with certainty (probability 1) or participating in a lottery which has a 60% probability of winning $5000 and a 40% probability of winning nothing. Assuming Marry acts rationally, which option would she choose? Show how you derived your answer.
2. Consider the choice between receiving $100 with certainty (probability 1) or participating in a lottery which has a 50% probability of winning $500 and a 50% probability of winning nothing. Approximate the value of R (to 3 significant digits) in an exponential utility function that would cause an individual to be indifferent to these two alternatives. (You might find it helpful to write a short program to help you solve this problem.)

16.1 Economists often make use of an exponential utility function for money: $U(x) = -e^{-x/R}$, where R is a positive constant representing an individual's risk tolerance. Risk tolerance reflects how likely an individual is to accept a lottery with a particular expected monetary value (EMV) versus some certain payoff. As R (which is measured in the same units as x) becomes larger, the individual becomes less risk-averse.

1. Assume Mary has an exponential utility function with $R = \$400$. Mary is given the choice between receiving \$400 with certainty (probability 1) or participating in a lottery which has a 60% probability of winning \$5000 and a 40% probability of winning nothing. Assuming Marry acts rationally, which option would she choose? Show how you derived your answer.
2. Consider the choice between receiving \$100 with certainty (probability 1) or participating in a lottery which has a 50% probability of winning \$500 and a 50% probability of winning nothing. Approximate the value of R (to 3 significant digits) in an exponential utility function that would cause an individual to be indifferent to these two alternatives. (You might find it helpful to write a short program to help you solve this problem.)

16.15 Alex is given the choice between two games. In Game 1, a fair coin is flipped and if it comes up heads, Alex receives \$100. If the coin comes up tails, Alex receives nothing. In Game 2, a fair coin is flipped twice. Each time the coin comes up heads, Alex receives \$50, and Alex receives nothing for each coin flip that comes up tails. Assuming that Alex has a monotonically increasing utility function for money in the range [\$0, \$100], show mathematically that if Alex prefers Game 2 to Game 1, then Alex is risk averse (at least with respect to this range of monetary amounts).

Show that if X_1 and X_2 are preferentially independent of X_3, and X_2 and X_3 are preferentially independent of X_1, then X_3 and X_1 are preferentially independent of X_2.

16.16 [airport-au-id-exercise]Repeat Exercise airport-id-exercise, using the action-utility representation shown in Figure airport-au-id-figure.

16.17 For either of the airport-siting diagrams from Exercises [airport-id-exercise] and [airport-au-id-exercise], to which conditional probability table entry is the utility most sensitive, given the available evidence?

16.18 Modify and extend the Bayesian network code in the code repository to provide for creation and evaluation of decision networks and the calculation of information value.

16.19 Consider a student who has the choice to buy or not buy a textbook for a course. We'll model this as a decision problem with one Boolean decision node, B, indicating whether the agent chooses to buy the book, and two Boolean chance nodes, M, indicating whether the student has mastered the material in the book, and P, indicating whether the student passes the course. Of course, there is also a utility node, U. A certain student, Sam, has an additive utility function: 0 for not buying the book and -\$100 for buying it; and \$2000 for passing the course and 0 for not passing. Sam's conditional probability estimates are as follows:

$$P(p\,|\,b, m) = 0.9 \qquad P(m\,|\,b) = 0.9$$
$$P(p\,|\,b, \neg m) = 0.5 \qquad P(m\,|\,\neg b) = 0.7$$
$$P(p\,|\,\neg b, m) = 0.8$$
$$P(p\,|\,\neg b, \neg m) = 0.3$$

You might think that P would be independent of B given M, But this course has an open-book final—so having the book helps.

1. Draw the decision network for this problem.
2. Compute the expected utility of buying the book and of not buying it.
3. What should Sam do?

16.20 [airport-id-exercise]This exercise completes the analysis of the airport-siting problem in Figure airport-id-figure.

1. Provide reasonable variable domains, probabilities, and utilities for the network, assuming that there are three possible sites.
2. Solve the decision problem.
3. What happens if changes in technology mean that each aircraft generates half the noise?
4. What if noise avoidance becomes three times more important?
5. Calculate the VPI for *AirTraffic*, *Litigation*, and *Construction* in your model.

16.21 [car-vpi-exercise] (Adapted from Pearl [@Pearl:1988].) A used-car buyer can decide to carry out various tests with various costs (e.g., kick the tires, take the car to a qualified mechanic) and then, depending on the outcome of the tests, decide which car to buy. We will assume that the buyer is deciding whether to buy car c_1, that there is time to carry out at most one test, and that t_1 is the test of c_1 and costs \$50.

A car can be in good shape (quality q^+) or bad shape (quality q^-), and the tests might help indicate what shape the car is in. Car c_1 costs \$1,500, and its market value is \$2,000 if it is in good shape; if not, \$700 in repairs will be needed to make it in good shape. The buyer's estimate is that c_1 has a 70% chance of being in good shape.

1. Draw the decision network that represents this problem.
2. Calculate the expected net gain from buying c_1, given no test.
3. Tests can be described by the probability that the car will pass or fail the test given that the car is in good or bad shape. We have the following information:

$P(pass(c_1, t_1)\,|\,q^+(c_1)) = 0.8$

$P(pass(c_1, t_1)\,|\,q^-(c_1)) = 0.35$

Use Bayes' theorem to calculate the probability that the car will pass (or fail) its test and hence the probability that it is in good (or bad) shape given each possible test outcome.

4. Calculate the optimal decisions given either a pass or a fail, and their expected utilities.
5. Calculate the value of information of the test, and derive an optimal conditional plan for the buyer.

16.22 [nonnegative-VPI-exercise]Recall the definition of *value of information* in Section VPI-section.

1. Prove that the value of information is nonnegative and order independent.
2. Explain why it is that some people would prefer not to get some information—for example, not wanting to know the sex of their baby when an ultrasound is done.
3. A function f on sets is **submodular** if, for any element x and any sets A and B such that $A \subseteq B$, adding x to A gives a greater increase in f than adding x to B:

$$A \subseteq B \implies (f(A \cup \{x\}) - f(A)) \geq (f(B \cup \{x\}) - f(B)) .$$

Submodularity captures the intuitive notion of *diminishing returns*. Is the value of information, viewed as a function f on sets of possible observations, submodular? Prove this or find a counterexample.

The answers to Exercise almanac-game (where M stands for million): First set: 3M, 1.6M, 1541, 41M, 4768, 221, 649M, 295M, 132, 25,546. Second set: 1917, 155M, 4,500M, 11M, 120,000, 1.1M, 1636, 19,340, 1,595, 41,710.