Contents

The previous chapter modeled once-and-for-all decisions. The decider (or *player*, as we shall now say) chooses a feasible action, then Nature reveals the true state and the player experiences the final consequences.

Life is more dynamic than that. A person makes a decision, events unfold influenced in part by that decision, then new decisions loom with new consequences, followed by more events and more decisions.

In this chapter we offer simple models of these dynamic situations. The first part considers how to combine new information with old; the technique is called Bayesian updating. The rest of the chapter draws on Bayesian updating to model optimal decision making when different choices have not just different immediate consequences but also different impacts on future information and choice opportunities.

1 Bayes' Theorem

How should we combine different pieces of information? It turns out that probability theory has a clear answer, one that may or may not accord with naive intuition. To explain that answer, we begin with notation.

Notation	*Classic name*	*Bayesian name*	*Meaning*
s	random variable	state of Nature	possible true state
z	random variable	message	possible signal or message
$p(s)$	marginal prob	prior prob of state s	prob of s before message received
$p(z)$	marginal prob	message prob	overall probability of signal z
$p(s,z)$	joint probability	...of given signal and state	Sometimes written $p(z \cap s)$.
$p(z\|s)$	conditional prob	likelihood	measures accuracy of message
$p(s\|z)$	conditional prob	posterior probability	prob of s after z received.

The idea is that the player knows the possible states $s \in S$, and given all previous relevant information, she assesses the probabilities as $p(s)$. Then she receives a new message z and updates the probabilities to $p(s|z)$.

For example, you think that the probability of rain tomorrow is $p(r) = 0.4$ and the probability of sun (or, more precisely, of no measurable rainfall) is $p(\mathtt{s}) = 0.6$. But then you hear a forecast $z = \rho$ that it will rain. You update your belief that the true state will be rain to $p(r|\rho)$.

But what is the correct updated probability $p(r|\rho)$ of rain? Obviously it depends on the reliability of the forecast, e.g., on the likelihood $p(\rho|r)$.... but what else does it depend on, and how do we calculate the value?

The answer is obtained via classical probability algebra that goes back to the 1700s (worked out originally by Reverend Thomas Bayes and later by Pierre-Simon Laplace). Twentieth Century philosophers and statisticians developed the ideas (and terminology) to show how to leverage that answer.

Here's the algebra, along with commentary.

- Begin with the joint probabilities $p(s, z)$ (sometimes written $p(z, s)$ or $p(z \cap s)$) that the true state is s and the message received is z.

- For example, given true states $r, \mathtt{s} \in S$ and possible messages $\rho, \sigma \in Z$, there are four possible joint events (s, z), whose probabilities $p(s, z)$ sum to 1.0.

  - Say $p(\mathtt{s}, \sigma) = 0.42, p(\mathtt{s}, \rho) = 0.18, p(r, \rho) = 0.36, p(r, \sigma) = 0.04$.

- Summing over messages, we get the marginal[1] state probabilities

$$p(s) = \sum_{z \in Z} p(s, z). \tag{1}$$

- You can check that in the example $p(r) = .36 + .04 = 0.4$ and $p(s) = .42 + .18 = 0.6$.

[1]This classic terminology is, of course, inconsistent with standard econ jargon, where marginal refers to first derivatives, not integrals or sums.

- The Bayesian jargon emphasizes that these are the probabilities you have *prior* to the arrival of the message (sometimes called *the news* or *a signal*.)

- Summing over states, we get the marginal probabilities

$$p(z) = \sum_{s \in S} p(s, z). \tag{2}$$

- In the example $p(\rho) = .18 + .36 = 0.54$ and $p(\sigma) = .42 + .04 = 0.46$.

- The Bayesian jargon – "message probabilities" – reminds you that these are the probabilities of the possible messages that you might receive.

- By definition[2] our two sorts of conditional probabilities are:

$$p(s|z) = \frac{p(s, z)}{p(z)} \tag{3}$$

$$p(z|s) = \frac{p(s, z)}{p(s)} \tag{4}$$

  Cross-multiplying equation (4) gives us

$$p(s, z) = p(z|s)p(s) \tag{5}$$

  Substituting equation (5) into (3) gives us a first version of Bayes' Theorem (also called Bayes' rule):

$$p(s|z) = \frac{p(z|s)p(s)}{p(z)} \tag{6}$$

- This equation gives us what we sought.
  - The LHS is the Bayesian posterior probability of state s occurring after observing signal z.
  - The numerator of the RHS is the likelihood that the signal corresponds to the true state, times the prior probability.

[2]These formulas assume that none of the marginal probabilities are zero. If any of them is, that state or message will never occur, and can be dropped from the list.

- The denominator normalizes so that the posterior probabilities sum to 1.0. Of course, equation (3) tells us that the denominator is the message probability.
- Thus the posterior probability is larger when
  - the prior is larger
  - the message is more accurate (higher likelihood), and
  - the message is rarer (smaller message prob).
- Plugging equation (5) into the definition (2) of the message probability gives us a different expression for the denominator:

$$p(z) = \sum_{t \in S} p(t, z) = \sum_{t \in S} p(z|t)p(t) \tag{7}$$

1.1 Four versions of the theorem

Theorem 1 (Bayes) *Using notation introduced above, the following formulas are all valid.*

$$p(s|z) = \frac{p(z|s)p(s)}{p(z)} \tag{8}$$

$$p(s|z) = \frac{p(z|s)p(s)}{\sum_{t \in S} p(z|t)p(t)} \tag{9}$$

$$\frac{p(s|z)}{p(t|z)} = \frac{p(z|s)p(s)}{p(z|t)p(t)} \tag{10}$$

$$\ln \frac{p(s|z)}{p(t|z)} = \ln \frac{p(z|s)}{p(z|t)} + \ln \frac{p(s)}{p(t)}. \tag{11}$$

The proof is simple.

- Equation (8) is a repeat of (6), derived via classic probability algebra and interpreted in terms of updating prior probabilities in the light of new information.

- Equation (9) is just (8) with (7) substituted into the denominator. It is handy when the message probabilities $p(z)$ are not given, or needed explicitly.

- Equation (10) is just the quotient of equation (8) for one particular state s and the same equation for another particular state r.[3] The denominators in equation (8) cancel when you take the quotient.

- Equation (10) says that the posterior odds $\frac{p(s|z)}{p(t|z)}$ of any two possible states are the product of the likelihood ratio and the two states' prior odds.

- Equation (11) is the easiest one for us to remember. It says that, expressed in logit form, Bayes' theorem is just a simple sum:
  posterior log odds = log likelihood + prior log odds.

Bayes theorem is written above for finite state spaces S and finite message spaces Z. There are two important extensions that should be mentioned immediately, so that you can do your homework problems.

1. Multiple messages. For example, you might hear several different forecasts about tomorrow's weather, or get several different medical diagnostic test results. With two distinct messages, for example, you might receive messages $z_1 \in Z_1, \; z_2 \in Z_2$.

- One way to extend Bayes theorem is to apply it to conjunctions of messages, and work out the joint probabilities and likelihoods for each possible conjunction, e.g., $z = (z_1, z_2)$ for two messages.

- Another way is to apply Bayes theorem sequentially. Treat the posterior probabilities $p(s|z_1)$ after receiving the first message as the prior probabilities when processing the second message, etc.

[3] If you prefer, you can take quotients using (9) instead of (8), once you remember that t is just a dummy variable for summation in (9) rather than a particular state.

- Either way, the algebra is straightforward if the likelihoods of subsequent messages are not affected by those of earlier messages. The key property is called *conditional independence*:

$$p(z_1, z_2|s) = p(z_1|s)p(z_2|s) \quad \forall\ z_1 \in Z_1,\ z_2 \in Z_2,\ s \in S. \tag{12}$$

- With messages $z = (z_1, ..., z_n)$ conditionally independent, equation (10) becomes a simple product. Likewise, equation (11) is then a simple sum

$$\ln \frac{p(s|z)}{p(t|z)} = \ln \frac{p(s)}{p(t)} + \sum_{i=1}^{n} \ln \frac{p(z_i|s)}{p(z_i|t)} \tag{13}$$

2. Infinite State and Message Spaces. For example, the possible messages might be the number of geiger counter clicks per minute which, practically speaking, has no finite upper bound. Fortunately, the Bayes formulas given earlier require no modification when there are countably infinite states or messages.

Slight modifications are needed when we apply Bayes theorem to continuous state or message spaces.

- Suppose that we have joint probability density $f(s, z)$ over some rectangle $S \subset \Re$ of states and $Z \subset \Re$ of messages, and the rectangle $S \times R$ contains the support of the joint distribution.

- Replace sums by integrals in the definitions (1) and (2) of marginal probabilities, e.g., the prior density for states $s \in S$ is $f(s) = \int_{z \in Z} f(s, z)dz$.

- Provided that the joint density is continuous and positive at the realized signal and state, all of the formulas in Bayes Theorem remain valid when the symbol p (for probability mass) is replaced everywhere by f (for probability density function).

- Readers who enjoy formal mathematics will be able to prove this by chopping up the support of the joint distribution using a fine grid (mesh $\delta > 0$, say), applying the discrete Bayes formulas, taking the limit as $\delta \to 0$, and applying the definitions of density and continuity at (s, z).

1.2 Using Bayes theorem

One way to visualize the elements of Bayes theorem is to draw a tree in which Nature first determines the true state and then determines which message to send, as in Fig 1.

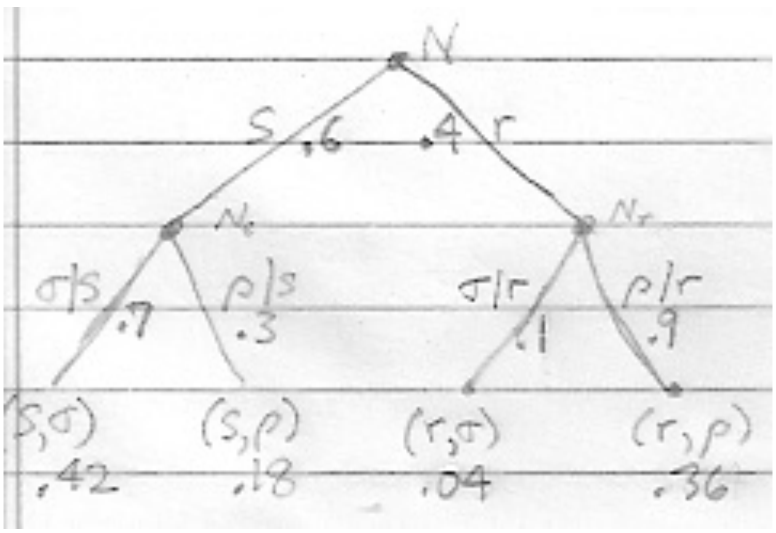

Figure 1: Nature determines the state (with prior probability $p(\mathtt{s}) = .6$) and then determines the message (with likelihoods $p(\sigma|\mathtt{s}) = .7$ and $p(\rho|r) = .9$). Those given probabilities imply the joint probabilities and posterior probabilities.

To return to our earlier example,

- the given prior probabilities of the states are $p(r) = 0.4$, $p(\mathtt{s}) = 0.6$. The message accuracy is given in the form of the likelihoods $p(\rho|r) = 0.90$ and $p(\sigma|\mathtt{s}) = 0.70$.
- Of course, the complementary likelihoods are $p(\rho|\mathtt{s}) = 1 - .70 = .30$ and $p(\sigma|r) = 1 - .90 = .10$.
- You can find the joint probabilities by following tree branches to the end, which is the interpretation of equation (5).

– For example, follow the $\mathtt{s}$ branch and the $\sigma|\mathtt{s}$ branch to obtain the joint probability $p(\mathtt{s}, \sigma) = p(\mathtt{s})p(\sigma|\mathtt{s})$.

– To spell it all out:

$$
\begin{aligned}
p(r, \rho) &= p(\rho|r)p(r) = (0.9)(0.4) = 0.36 \\
p(r, \sigma) &= p(\sigma|r)p(r) = (0.1)(0.4) = 0.04 \\
p(s, \rho) &= p(\rho|\mathtt{s})p(\mathtt{s}) = (0.3)(0.6) = 0.18 \\
p(s, \sigma) &= p(\sigma|\mathtt{s})p(\mathtt{s}) = (0.7)(0.6) = 0.42 \\
sum\ of\ joints &= 1
\end{aligned} \tag{14}
$$

- Read the first line as: "the joint probability of a rainy forecast and a rainy day is equal to the likelihood of a rainy forecast conditional on a rainy day, times the probability of a rainy day."

- Note that the sum of the joint probabilities equals one since they cover all possible combinations.

Next, you can obtain the unconditional probability of a given message (ex. a rainy forecast ρ or a sunny forecast σ).

- Recall that it is found by summing the joint probabilities containing that signal across the possible states.

$$p(z) = \sum_{s \in S} p(s, z) \tag{15}$$

- Thus, the message probabilities in the example are:

$$
\begin{aligned}
p(\rho) &= \sum_{s \in S} p(\rho, s) = p(\rho, r) + p(\rho, \mathtt{s}) = 0.36 + 0.18 = 0.54 \\
p(\sigma) &= \sum_{s \in S} p(\sigma, s) = p(\sigma, r) + p(\sigma, \mathtt{s}) = 0.04 + 0.42 = 0.46
\end{aligned} \tag{16}
$$

- Note that even though there is a higher prior for sunny days ($p(\mathbf{s}) = 0.60$), there is a lower prior for a sunny forecast ($p(\sigma) = 0.46$). The reason is that a rainy forecast is more accurate: $p(\rho|r) = 0.90$. There is only a 10% chance of observing a sunny forecast on a rainy day, but a 30% of observing a rainy forecast on a sunny day. Sunny days are more likely (prior of 60%), and a rainy forecast occurs 30% of the time on a sunny day. This makes a sunny forecast more believable, since it only occurs 10% of the time on rainy days.

Finally, you get to the main goal — to find the Bayesian posterior probabilities of each possible state having seen a particular message.

- Just use version 1 of Bayes Rule:

$$p(s|z) = \frac{p(s,z)}{p(z)} = \frac{p(z|s)p(s)}{p(z)}.$$

- In the current example,

$$\begin{aligned} p(r|\rho) &= \frac{p(r,\rho)}{p(\rho)} = \frac{0.36}{0.54} = 0.67 \\ p(s|\rho) &= \frac{p(s,\rho)}{p(\rho)} = \frac{0.18}{0.54} = 0.33 \\ p(r|\sigma) &= \frac{p(r,\sigma)}{p(\sigma)} = \frac{0.04}{0.46} = 0.087 \\ p(s|\sigma) &= \frac{p(s,\sigma)}{p(\sigma)} = \frac{0.42}{0.46} = 0.913 \end{aligned} \tag{17}$$

- Note that the sum of the posteriors for any given message is 1.

To do these sorts of calculations routinely, download from the class website a spreadsheet that spells it all out.

1. Begin by filling in your given information regarding prior probabilities and message likelihoods.

2. Use those givens to calculate complementary prior probabilities and likelihoods.

3. Note that likelihoods over all messages for a given state sum to 1.0, but that the likelihoods over all states for a given message can sum to more or less than 1.0.

4. Then use the spreadsheet to compute the joint probabilities.

5. Alternatively, if joint probabilities are given, then enter them and proceed with the remaining steps.

6. Calculate the message probabilities as column sums. The spreadsheets do this automatically.

7. Finally, calculate all the posterior probabilities by dividing joints by the appropriate message probs.The spreadsheets do this automatically.

An example. UCSC student Z. Wang has a project that includes a good spreadsheet exercise. He begins with two urns, A and B. Urn A has 6 red balls and 4 green, while urn B has 6 green and 4 red. A flip of a fair coin not visible to the player determines the urn to be used subsequently. The player sees the colors of 15 balls drawn with replacement from that urn, estimates the probability that the urn is A, then sees the colors of 10 more draws with replacement from the same urn.

1. Suppose that the first set of draws has 9 green balls of 15. What are the Bayesian prior and posterior probabilities that the balls come from urn A?

2. Suppose that the second set of draws has 6 green of 10. If the player didn't see the first set of draws, what is her Bayesian posterior?

3. What is the Bayesian posterior after seeing both draws just described? Is the intuition valid that it lies between the posteriors from the previous two items?

4. What is the general Bayesian posterior after seeing n green balls of 15 and m green balls of 10? How does it compare the the posteriors of seeing only n green balls of 15 or only m green balls of 10?

2 Decision trees

Decision trees are a great device for posing and solving dynamic decision problems. These trees involve one player, plus Nature. Decision trees consist of:

1. an initial node, sometimes called a root.

2. subsequent nodes connected to their predecessor nodes by branches.

3. each non-terminal node is owned either by Nature or the player, and is so labelled.

4. terminal nodes (sometimes called leaves) give the utility of the outcomes they represent.

Example: rain and umbrellas, Figure 2 .

- Action set $A = \{\text{carry umbrella } (c), \text{ don't carry umbrella } (d)\}$.

- Set of states $S = \{\text{rain } (r), \text{ sun}(\mathtt{s})\}$.

- Nature (N) determines which state is realized according to $p(r) = 0.4$, $p(\mathtt{s}) = 0.6$.

- The fact that Nature moves after the Player in this tree indicates that the Player does not know Nature's move when choosing his action.

2.1 Solving trees

The general method for solving trees is called **backwards induction (BI)**:

1. Starting with terminal nodes, (expected) utility is induced at preceding nodes, step by step.

2. Induction at Nature nodes is done by taking the expected value across branches.

3. Induction at player nodes is done by taking the maximum across branches.

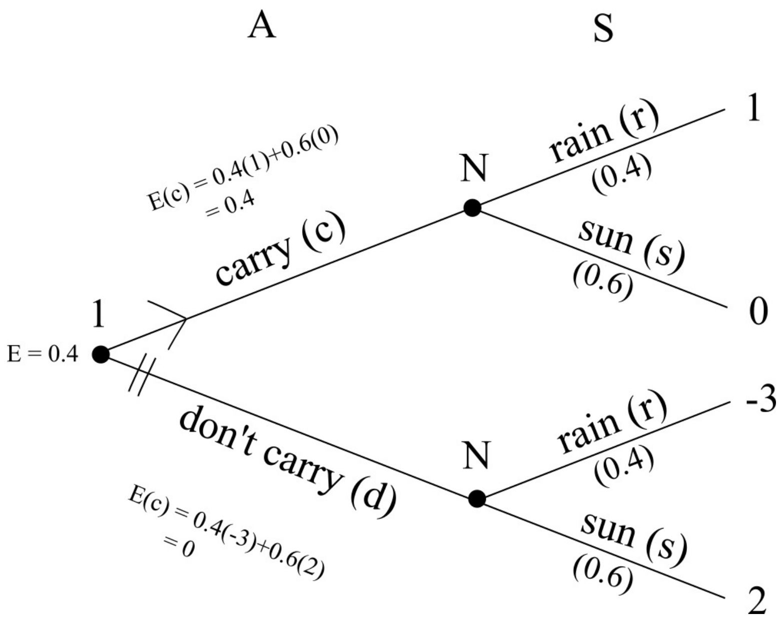

Figure 2: A simple decision tree, solved [typo to fix: bottom E(c) = 0 should be E(d) = 0, or $E(u|d) = 0$].

4. The value at the root node obtained in the final step is the value of the (optimal) decision.

5. The player node choices together constitute the complete contingency plan, or solution of the problem.

In the Figure 2 example:

- Step 1: at top nature node (c) the expected utility of carry is: $E(u|c) = 0.4(1) + 0.6(0) = 0.4$.

  At the bottom nature node (d) the expected utility of don't carry is: $E(u|d) = 0.4(-3) + 0.6(2) = 0$.

- Step 2: Therefore, at the decision node (initial node, root in this case) $E(u|c) > E(u|d)$, so the optimal decision is to carry the umbrella.

- Step 3: Value of the decision is then obtained at the initial node, $E(u|c) = 0.4$. (Note that this is an expectation, not a realization. Given this action, realized utility will be either 0 or 1.)

What if we redrew the tree so that Nature moved first, then the player, but the payoffs were the same. Then there would be no uncertainty, since the player moves after the true state is known.

- With no uncertainty, the decision maker would have $u = 1$ if rain since he would have the umbrella and $u = 2$ if sun since he would have no umbrella.
- The expected value of the decision with no uncertainty is $0.4(1) + 0.6(2) = 1.6$.

Now, consider the more complicated decision tree when you don't know the true state but have heard a forecast. Using the example from the previous section, the decision tree is as in Figure 3.

We solve this more complicated decision problem as follows.

- First use Bayes Theorem to find the (message and) posterior probabilities.
  - Figure 2 used the prior probabilities, but here we use
  - the posteriors given message $z = r$ on the top 2 final branches, and
  - the posteriors given message $z = \mathtt{s}$ on the bottom 2 final branches, because
  - the top initial branch is for when message $z = r$ is received and the bottom initial branch is for $z = \mathtt{s}$, so those branches have the respective message probabilities $(.54, .46)$.
- With the tree now fully specified, we solve it via BI.
  - Take the expected utilities at the 4 N nodes that just precede the terminal nodes.

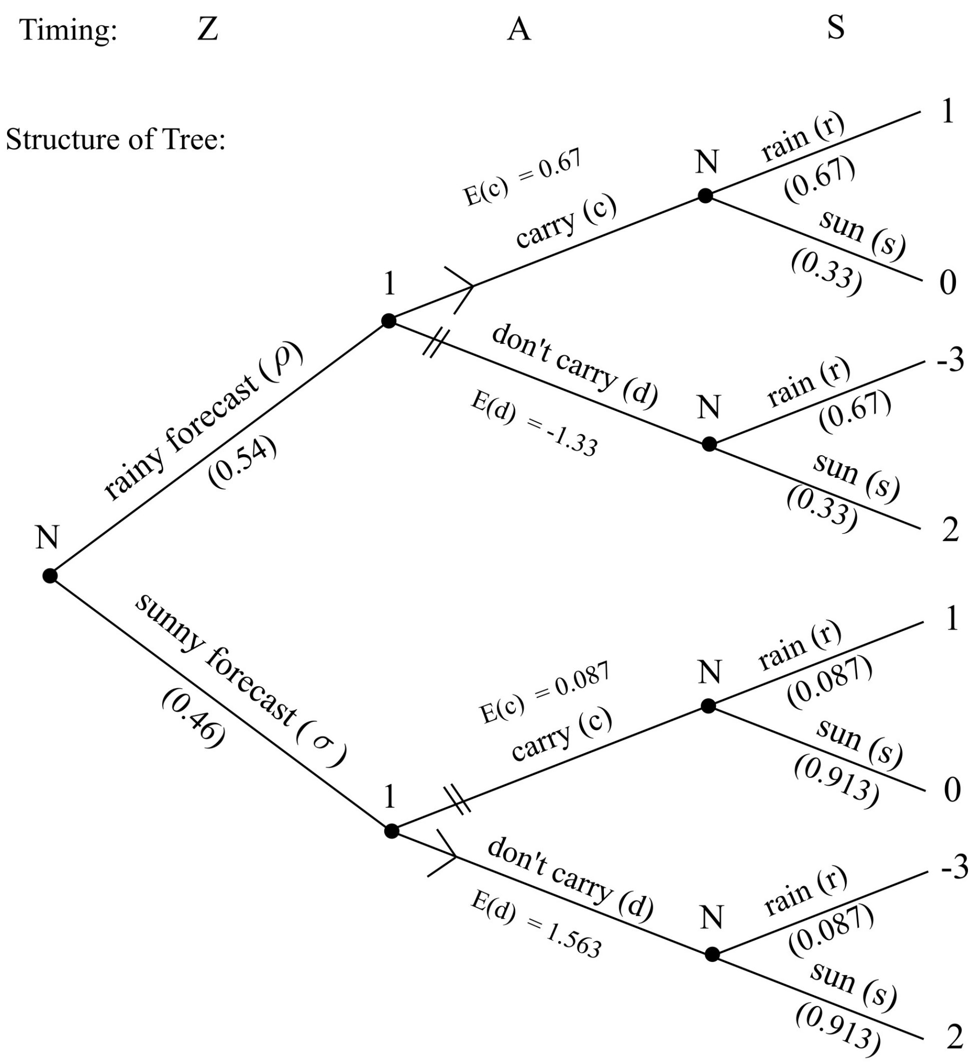

Figure 3: Forecast decision tree, solved.

- Thus $E(u|c,\rho) = 0.67, E(u|d,\rho) = -1.33, E(u|c,\sigma) = 0.087, E(u|d,\sigma) = 1.563$.
- Working backward, we next consider the preceding nodes. These are all player's decision nodes, so we will maximize his expected utility.
- On the upper branch (where the message is ρ), he chooses c since $E(u|c,\rho) = 0.67 > -1.33 = E(u|d,\rho)$. The maximized expected value 0.67 is attached to that decision node.
- On the lower branch (where the message is σ), he chooses d since $E(u|d,\sigma) =$

$1.563 > 0.087 = E(u|c, \sigma)$. The maximized expected value 1.563 is attached to that decision node.

  - Finally, we work back to the initial node, owned by N. We take the expected value, $Eu = (.54)(.67) + (.46)(1.563) = 1.08$.

Here's how to read the solution.

- The decision rule is an optimal action in each contingency. Here the contingencies are which message (forecast) is received. We obtained:

$$\text{Decision Rule} : \{c|\rho, d|\sigma\} \tag{18}$$

  - The decision rule states: "carry an umbrella if there is a rainy forecast, don't carry if there is a sunny forecast."

- The expected value of the tree is: $E = (0.54)(0.67) + (0.46)(1.563) = 1.08$.

  - This is the probability of a signal times the expected utility of that signal, summed across all possible signals.

- Recall the optimal decision with no information was to carry the umbrella and the expected payoff if this optimal uninformed decision was $E_c = 0.4$.

- If there was perfect information (signal fully reveals the state, i.e. Bayesian posterior probability is either 0 or 1) then the expected value is: $E_{PI} = 0.6(2) + 0.4(1) = 1.6$.

  - Thus, the value of perfect information is the value of the informed tree minus the value of the uninformed tree: $VPI = E_{PI} - E_c = 1.6 - 0.4 = 1.2$.

- Our signals do not fully reveal the states (likelihoods of being correct less than 1).

  - Thus the value of this imperfect information is: $E - E_c = 1.08 - 0.4 = 0.68$.

2.2 Value of information.

With this example in hand, we can state some general principles. For now, assume that the utilities are monetary payoffs and the player is risk neutral.

The *value of information* is the difference between the expected payoff with and without the information.

- This difference is the maximum that the person would be willing to pay for the information. Precise (but rather cumbersome) mathematical expressions appear below in equations (21, 22).

- A crucial insight: If your optimal action doesn't depend on which message you receive, then the information has no value.

- In the example, you always carry the umbrella with no information, but you leave it home when the forecast is σ. If the likelihoods or payoffs were a bit different so that it would be optimal to carry the umbrella even if the message were σ, then the expected payoff would be the same with as without information, and the value of information would be 0.

To make this more precise mathematically, begin with the usual notation for states $s \in S$, messages $z \in Z$ and joint probabilities p.

- Denote the optimal uninformed decision as a_o and the optimal informed decision given message z as a_z.

- Denote the value of the optimal uninformed decision in state s as $u_o^*(s) \equiv u(a_o, s)$ and the corresponding value of the optimal informed decision, given signal z, as $u_z^*(s) \equiv u(a_z, s)$.

- By EUH the value $u_o^*(s)$ of the optimal uniformed decision depends on the priors and satisfies:

$$\sum_{s\in S} p(s)u_o^*(s) \geq \sum_{s\in S} p(s)u(a,s) \qquad \forall a \neq a_o \in A \tag{19}$$

- By EUH the value of the optimal informed decision $u_z^*(s)$ depends on the Bayesian posteriors and satisfies:

$$\sum_{s\in S} p(s|z)u_z^*(s) \geq \sum_{s\in S} p(s|z)u(a,s) \qquad \forall a \neq a_z \in A \tag{20}$$

Risk Neutral Case. If the player is risk neutral, we can take the utilities to be monetary payoffs. In this case the value of information structure $I = [S, Z, p]$ is the weighted sum of the improved payoff for each message:

$$V_I = \sum_{z\in Z} p(z) \sum_{s\in S} p(s|z) \left[u_z^*(s) - u_o^*(s)\right]. \tag{21}$$

- If $a_z = a_o$ then the information does not change the action and the value of information for that message is zero.
- If $a_z \neq a_o$ then on average the payoff is improved, but not necessarily in all states.
  - For example, suppose that $p(s|z) = 0.80$ for this message and some state s.
  - It may be for this state (with probability 80%) the information causes you to "correctly" change your action, yielding a higher payoff in that state.
  - However, it may be that the other 20% of the time the information causes you to change a correct action into an incorrect one, lowering your payoff. Thus, a_o would have been a correct choice, but the signal z caused you to change your action to a_z, which is not optimal in the realized state. This will have a negative impact on the value of information.
- Equation (20) ensures that the weighted sum is nonnegative. It says that when message z is received, the action a_z yields the highest expected payoff over all other possible actions, including a_o.

- That is, for each message you improve your expected payoff (or leave it unchanged).

- Finally, equation (21) weights the improvements by the probability $p(z)$ of receiving that message.

General Case. For a player with Bernoulli function u, the value V_I of information structure $I = [S, Z, p]$ is the willingness to pay to get the optimal informed expected utility instead of the optimal uninformed expected utility. One way to write it is to say that V_I is the solution x to:

$$\sum_{z \in Z} \sum_{s \in S} p(z) p(s|z) u(a_z^* - x, s) = \sum_{s \in S} p(s) u(a_o^*, s) \tag{22}$$

The LHS is the expected utility of the action with information given payment x, and the RHS is the EU without information. Since Bernoulli functions are, in practice, seldom known for specific players, the risk neutral case, in equation (21), is usually more useful in applications.

Upper Bound. In many applications we don't actually know some of the details of the information structure before making a decision of whether to purchase it. It is then helpful to have an upper bound on V_I.

- The value of perfect information, denoted VPI, refers to the value of information, V_I, when the signal z always fully reveals the true state s.

- The information is said to be perfect when the Bayesian posterior probability is either zero or one for all signals:

$$p(s|z) \in \{0, 1\} \quad \forall z \in Z \tag{23}$$

- The VPI provides an upper bound on V_I when the information is imperfect.

  - Note that VPI requires at least as many signals as states, or using the cardinality of a set, VPI requires: $|Z| \geq |S|$.

– If the posterior $p(s|z) = 1$ (and if $|Z| = |S|$) then the appropriate likelihood $p(z|s)$ must also be one and all other likelihoods for that state must be zero.

– This also means that the prior of the signal must equal the prior of the state that it perfectly reveals (if $|Z| = |S|$).

With these formal definitions in hand, it is time for some more intuition. Information has the most value when (note that this is not a mutually exclusive list):

- The prior probabilities of the states are close together.
- The larger the cost from making the wrong decision.
- The more responsive the actions are to the information. This depends on how different the Bayesian posteriors are from the priors for a given state.
- The higher the likelihoods. This tells you how accurate the signals are.
- The higher the probability of observing a signal that changes your action.

3 Dynamic programming and beyond

Dynamic programming is used to solve dynamic optimization problems that are additively separable and in finite, discrete time.

- Let $a(t)$ be the choice variable and let $y(t)$ be the state variable.
- A general form of the problem is:

$$\begin{aligned} V(y_0, 0) &= \max_{a_t \in A_t} \sum_{t=0}^{T} F(a_{t,} y_t, t) \\ \text{subject to } y_{t+1} &= y_t + Q(a_t, y_t, t) \\ \text{and } G(a_t, y_t, t) &\geq 0 \end{aligned} \tag{24}$$

where:

– F is the period-by-period payoff,

– Q is a deterministic function which determines the evolution of the state variable.

– G is the constraint on the set of allowable actions.

– The initial value of the state variable y_o is typically given.

– The time horizon may be finite (T is a positive integer) or infinite ($T = \infty$).

- The solution is a series of actions which maximize the value function V starting at time zero.

- But the actions shouldn't be chosen in isolation ("myopically") because they are linked via Q.

- The value function $V(y_t, t)$ takes the same form as in equation (24) but the sum starts at t instead of at $t = 0$.

- In many applications, F and V are independent of t, except perhaps at the (finite) last period $t = T < \infty$.

Dynamic programming problems (DPPs) like (24) can be thought of as a discrete time version of optimal control theory. DPPs are deterministic in that there are no "Nature" moves.

The Bellman equation often helps to solve DPPs. It defines the value function recursively:

$$\begin{aligned} V(y_t, t) &= \sup_{a_t \in A_t} \{F(a_t, y_t, t) + V(y_{t+1}, t+1)\} \\ \text{for } t &= 0, \ldots, T-1 \end{aligned} \tag{25}$$

- The sup in equation (25) refers to a supremum, which is the highest attainable value.

- A supremum always exists, even if a maximum doesn't. (For example, 1 is the supremum of the interval $(0,1)$, which has no maximum. 1 is both the supremum and the maximum of the interval $(0,1]$, so they coincide when the supremum is an element of the set.)

- Equation (25) says to choose the action which yields the highest sum of the payoff this period, the F term, plus the continuation value, the V term.

- The Bellman equation can be solved in the finite horizon case via backward induction. Indeed, the equation is essentially a statement of induced value via BI.

- The first order condition for this type of problem captures the trade-off between the value today and the discounted value tomorrow.

- Proposition 3b in Kreps says that for a nonstochastic problem solving the Bellman equation gives you a solution to a dynamic programming problem and visa-versa.

There are different solution techniques for DPPs.

- Brute force: You can backwards induct when you have a problem which is in finite time ($T < \infty$).

- Use a Bellman equation to verify a conjecture.

- Pontryagin maximization principle. You can differentiate the Bellman equation to convert the functional equation into a differential equation.

- Manipulate the Bellman equation to obtain the necessary conditions and insights, but not actually solve the problem.

Other issues in solving DPPs.

- With an infinite horizon, you may require a transversality condition that constrains asymptotic behavior.

- Continuous time when the control, or action, variable a_t is a jump variable (see Dixit).

Stochastic problems

- If the problem contains a stochastic process it can be difficult to model.
- It usually helps to assume that the stochastic process is stationary. This means that the probabilities are time independent. A time independent stochastic process is called a Markov process. Now the problem becomes:

$$\begin{aligned} & \max_{a_t \in A_t} \mathrm{E} \sum_{t=0}^{\infty} d^t F(a_t, y_t) \\ \text{given} \quad & p(y_{t+1}|y_t, a_t) \end{aligned} \tag{26}$$

  Where $p(y_{t+1}|y_t, a_t)$ is a Markov transition probability matrix that depends on a_t and $d \equiv \frac{1}{1+r}$ is a discount factor where $r > 0$ is the discount rate.

  - $p(y_{t+1}|y_t, a_t)$ gives the probability of being in a particular state next period conditional on the present state and the chosen action.
  - The solution to this problem is a complete contingency plan. "If y_t has value __, then take action __".

- Here the Bellman equation is written in terms of expected values. Let α_t be a contingency plan given by the above quote. Then the Bellman equation is:

$$V(y) = \sup_{\substack{a_t \in A_t \\ t > t_0}} \left\{ F(a_t, y_t) + d \sum_{y \in S} V(y) p(y|y_{t_o}, \alpha_t) \right\} \tag{27}$$

- So the value now (LHS) is the maximum current payoff plus the discounted expected value next period.
- That expected future value is the state-contingent value times the posterior probability of that state occurring given the current state and current action.

- There is path dependence in the sense that the current Markov probabilities are a result of the past actions which have influenced (stochastically) the current state.
- This sort of problem is sometimes called dynamic stochastic decision problem, or a markov decision process, or stochastic optimal control.
- There are also continuous time versions of this problem, but we won't go into them here. Dixit is a good first source.

4 Behavioral Considerations

To be added in the next draft.

Zika example. Base rate neglect. Confirmation bias. Giverenzer approach

Perhaps mention quasihyperbolic discounting. (especially if the first chapter mentions prospect theory.)

also, the chapter should perhaps mention the Scoring Rule literature.

maybe add problems on real options, e.g., angel investing and beta tests, or stages of FDA approval for new drugs.

also problems on perfect vs useless info. if z independent of s, then show posterior = prior.

also problem on multiple messages that aren't conditionally independent.

5 Readings

new cites

Savage, Leonard J. 1954. The Foundations of Statistics. New York, Wiley.

Camerer, Colin. Behavioral Game Theory book.

Cheung, Stephen L. "Recent developments in the experimental elicitation of time preference." Journal of Behavioral and Experimental Finance 11 (2016): 1 – 8.