**Statistical Decision Theory: Incorporating additional information**

Recall that under the conditions of **risk** which is characterized by the ability of the decision maker to assign objective probabilities to the states of nature, the near-universal decision criterion is the expected monetary value. Each course of action is evaluated by the weighted average of pay-offs under each state of nature where the weights are the state probabilities. Obviously, the more information we have about the likelihoods of the states of nature, the better the decisions will be. Thus there is a great deal of motivation to obtain additional information to recognize the states of nature that are more likely to occur. However, in most cases additional information is not free but comes at a cost. In addition, most information will not be perfect to completely eliminate the uncertainty about the states of nature; it will simply reduce it. Such imperfect information can be obtained from various sources such as market surveys, questionnaires, statistics from a random sample, geological measurements to assess the potential of an oil field or many medical diagnostic tests. It is therefore very important to judge the value of additional information to see if it is worth its cost.

Suppose in the oil field development example we did in class, we are offered the use of a “crystal ball” that will tell us what the condition of the field is before we make our commitment to *drill*, *unconditionally* *lease* or *conditionally lease.*  We will use this imaginary device to assess the value of *perfect information.*  Although such perfect information may not exist is reality, it is conceptually useful to place an upper bound on the value of any information—no information may be more valuable than the perfect information.

**Expected value *of* perfect information (EVPI)**

Let’s reproduce the oil drilling problem:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **State Probs.** | **.12** | **.16** | **.24** | **.48** | **EMV** |
| **States** | ***S*1:500K** | ***S*2:200K** | ***S*3:50K** | ***S*4:dry** |  |
| **Drill** | $6.5M | $2.0M | $-0.25M | $-0.75M | $0.68M |
| **Unc. Lease** | $0.45M | $0.45M | $0.45M | $0.45M | $0.45M |
| **Cond. Lease** | $0.25M | $1.0M | 0 | 0 | $0.46M |

Using the criterion of expected monetary value (EMV) or Bayes’ rule we choose the drill option with an optimal expected pay-off of $0.68. We may call this the *expected value without perfect information*. To value perfect information, we calculate the expected value we could have obtained had we used the crystal ball. We will call this the *expected value* ***with*** *perfect information.*  EVPI then is simply the difference between these two expected values since the difference is the additional value that can be attributed to the crystal ball. To calculate the expected value with perfect information we reason as follows: Before we look at the crystal ball our best guess is that it will say the field is rich (500K) with 12% probability. If that should happen, the best option would be to drill and obtain $650K. Thus with the use of the crystal ball we have a shot to a payoff of 650 with 12% probability. Likewise 16% chance crystal ball will indicate a mediocre (200K) field in which case the best option is to drill and obtain $200. Thus the value with the crystal ball is the sum of the best pay-offs under each state of nature weighted by the state probabilities or .12\*6.5+.16\*2.0+.24\*0.45+.48\*0.45 = 1.424M.. Thus the difference between this value with perfect information and that without such information is $1.424M – $0.68M = $0.744M. This is the maximum any information, perfect or otherwise, is worth. And therefore can also be considered as the *cost of uncertainty* in this problem.

**Expected value of sample (imperfect) information (EVSI)**

Alas! In the real World we do not have perfectly reliable crystal balls; we get our information from sources that are not always truthful. We may conduct a market research on a sample of potential customers, but what we learn from the sample may not reflect exactly how the market as a whole will react to a new product. Or a physician may administer a test to look for a ‘marker gene’ whose presence may strongly correlate with a certain medical condition but may not necessarily imply the existence of it. In all of these cases the outcome of the additional information may correlate with certain states of nature increasing our belief in their likelihood such as the gene increasing the likelihood of a medical condition or the discovery of a “closed-structure” through seismic tests to indicate a stronger possibility of a rich oil field. To value such information we make use of the well-known Bayes’ Law.

First, some terminology: In addition to the distinct known set of *alternatives* (options), *Ak* and *states of natures*, *Sj,* we now have the outcomes of the new information source. We call the results of the new information *events*. Whatever process is used to gain additional information, it is assumed to result in a known set of discrete outcomes (events) denoted by *Ei*. In the oil field problem suppose there is a seismic test costing $10K that can reveal more information on the potential of the field. Specifically the test may result in one of three events: *E*1: “closed geological structure;” *E*2: “some structure” and *E*3: “no structure.” It is known from past experience that a closed-structure is very favorable for a rich field while no structure is a good indicator of a poor field. More specifically from the historical performance of the test we know the *reliability* of the information in the form of the following conditional probabilities, *P*(*Ei*|*Sj*):

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Priors: | .12 | .16 | .24 | .48 |
|  | *S*1:500K | *S*2:200K | *S*3:50K | *S*4:dry |
| *E*1: Closed Structure | 7/12 | 9/16 | 11/24 | 9/48 |
| *E*2: Some Structure | 4/12 | 4/16 | 8/24 | 18/48 |
| *E*3: No Structure | 1/12 | 3/16 | 5/24 | 21/48 |

For example, in 7 out of 12 fields that have proven rich (*S*1) in the past, a closed-structure has been observed, while 21 out of 48 dry fields (*S*4) have shown no structure, etc. Bayes Law reasons that if we should discover a closed structure through this seismic test, our belief in a rich field ought to go up; while the discovery of no structure should increase the likelihood of poor fields.

The state probabilities *P*(*Sj*), before any additional information are called ***priors*** (probability of state of nature *j* before any additional information). Bayes’ Law allows us to ***revise*** these prior probabilities based on the result of the seismic test. The revised probabilities are called the ***posterior*** probabilities, (*e*.g*.*, the probability of a 500K field **if** we observed a closed structure, *etc*.) Notice from this characterization that the posterior probabilities are also conditional probabilities, now in the form of *P*(*S*j|*E*i), not to be confused with the reliability probabilities of the form *P*(*Ei*|*Sj*). Caution *P*(*S*j|*E*i) ≠ *P*(*Ei*|*Sj*)! At the heart of the Bayes’ Law is the basic **multiplication rule** of probabilities which we briefly review:

**Digression:** **multiplication rule:**

Let *A* and *B* be two events. Then observing both events together*, i.e.*, the **join**t probability of these events, P(*A* and *B*) = P(*A*) \* P(*B*|*A*) = P(*B*) \* P(*A*|*B*). Notice that this rule allows events to be dependent, in the sense that probability of event *A* may depend on whether or not event *B* occurred or *vice versa*. To illustrate consider event *A*: an ace from a deck; event *B*: another ace from the same deck. Probability of *B* depends on whether event *A* occurred (in which case it is 3/51; otherwise it is 4/51). Thus the probability of two consecutive aces, P(*A* and *B*) = 4/52\*3/51. Notice that 3/51 is P(*B*|*A*).
**End digression**.

To apply the Bayes’ Law, we first compute) the joint probabilities, P(*Sj* and *Ei*) as P(*Sj*)\*P(*Ei*|*Sj*) as in table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Priors:** | **.12** | **,16** | **.24** | **.48** |  |
|  | ***S*1:500K** | ***S*2:200K** | ***S*3:50K** | ***S*4:dry** | **Marginals** |
| ***E*1: Closed Structure** | 0.07 | 0.09 | 0.11 | 0.09 | 0.36 = *E*1 |
| ***E*2: Some Structure** | 0.04 | 0.04 | 0.08 | 0.18 | 0.34 = *E*2 |
| ***E*3: No Structure** | 0.01 | 0.03 | 0.05 | 0.212 | 0.30 = *E*3 |

Under each state of nature, we have the joint probabilities *e.g.*, probability of finding a closed structure followed by the discovery of a 500K field is .07: P(*S*1)\*P(*E*1|*S*1) = .12 \* 7/12 = .07, etc. The far right column, the *marginal* probabilities lists the unconditional probability of the seismic test ending in *E*1, *E*2 or *E*3, *i.*e., regardless of the type of field found. In 36% of all cases, the seismic test results in a closed geological structure (*E*1), 34% of the cases some structure, *etc*. The marginal probabilities are found by simply summing the joint probabilities row-wise. In 7% of cases *E*1 has occurred and a 500k field has been found; in additional 9% of the cases *E*1 occurred and a 200K field is found; in 11% of cases *E*1 has occurred and a 50K field has been found; finally in 9% cases *E*1 has occurred and field was dry. Thus, the sum of these mutually exclusive joint events gives us the probability of *E*1 regardless of what type of a field it was followed by.

Using the multiplication rule again, this time as P(*S*1 and *E*1) = P(*E*1)\* P(*S*1|*E*1), we obtain the posterior probabilities. Here we know P(*S*1 and *E*1) = .07 and P(*E*1) = 0.36; therefore the posterior probability (of the state *S*1conditional on event *E*1) is simply P(*S*1|*E*1) = P(*S*1 and *E*1) /P(*E*1) = .07/.36 = .1944. To see this intuitively reason this way: if we have 1000 cases; the test is expected to result in *E*1 360 times; of these 360 times in 70 cases the field is expected to yield 500K. Thus the probability of a 500K field after observing *E*1 a closed structure is 70 out of 360 or 19.44%. The other posterior probabilities are obtained in the same way yielding the table below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **.12** | **,16** | **.24** | **.48** |
|  | ***S*1:500K** | ***S*2:200K** | ***S*3:50K** | ***S*4:dry** |
| ***E*1: Closed Structure** | 0.1944 | 0.2500 | 0.3056 | 0.2500 |
| ***E*2: Some Structure** | 0.1176 | 0.1176 | 0.2353 | 0.5294 |
| ***E*3: No Structure** | 0.0333 | 0.1000 | 0.1667 | 0.7000 |

Remember the prior probability of a 500K field was a mere .12. Bayes law revises this upward to a .1944 if the test reveals a closed structure; also if no structure is discovered (*E*3) probability of a 500K field goes down to a mere .0333.

How much is this test worth? Namely what is the EVSI, the expected value of this imperfect information that can be obtained from the seismic test? We answer the question the same way we answered it for perfect information—as the difference between expected value with the information from, the seismic test and the expected value without the seismic test (which we already know is $0.68M). To compute the expected value with this information, we reason that there is 36% chance of discovering a closed-structure (*E*1) in which case the decision problem is the same—same alternatives, same states of nature and same pay-offs, except the probabilities are now the posteriors. In other words the problem we would encounter in 36% of the cases is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **State Probs.** | **0.1944** | **0.2500** | **0.3056** | **.2500** | **EMV** |
| **States** | **500K** | **200K** | **50K** | **Dry** |  |
| **Drill** | $6.5M | $2.0M | $-0.25M | $-0.75M | **$1.4997M** |
| **Unc. Lease** | $0.45M | $0.45M | $0.45M | $0.45M | $0.45M |
| **Cond. Lease** | $0.25M | $1.0M | 0 | 0 | $0.299M |

The best option is to *drill* netting a $1.4997M.

There is also 34 % chance of discovering *E*2 in which case the problem to be solved is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **State Probs.** | **0.1176** | **0.1176** | **0.2353** | **0.5294** | **EMV** |
| **States** | **500K** | **200K** | **50K** | **Dry** |  |
| Drill | $6.5M | $2.0M | $-0.25M | $-0.75M | **$0.5437M** |
| Unc. Lease | $0.45M | $0.45M | $0.45M | $0.45M | 0.45M |
| Cond. Lease | $0.25M | $1.0M | 0 | 0 | 0.147M |

The best option is still to *drill* with an expected pay-off of $0.5437M.

Finally, there is 30% chance of discovering no structure whence the problem is

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **State Probs.** | **0.0333** | **0.1000** | **0.1667** | **0.2500** | **EMV** |
| **States** | **500K** | **200K** | **50K** | **Dry** |  |
| **Drill** | $6.5M | $2.0M | $-0.25M | $-0.75M | $-0.1502M |
| **Unc. Lease** | $0.45M | $0.45M | $0.45M | $0.45M | **0.45M** |
| **Cond. Lease** | $0.25M | $1.0M | 0 | 0 | 0.1083M |

The best option now is to *unconditionally lease* with an expected pay-off of $0.45M.

So what is the expected value if the seismic test is used? We have 36% chance for a pay-off of 1.4997M, plus 34% chance to a pay-off of $5437M and 30% chance to a pay-off of $0.45M. Adding these all up we have expected value with the seismic information: .36\*1.4997 + .34\* .5437 + .30\*0.45 = $.8596M. So the seismic test is worth $.8596M – 0.68M = $.1796M. Since the test costs only $10K (=.01M), it ought to be used. If the test shows closed structure (*E*1) or some structure (*E*2) the optimal decision is to drill, if no structure is found (*E*3) it should be unconditionally leased. This decision rule has an optimal expected gain of $.8496M, net of the test’s cost

**Decision trees:**

Up to now we have considered single stage decisions where there is a choice to be made under some uncertainty. You make your choice and the problem ends. In many business situations, decisions sometimes have to be made in succession where the outcome of the decision made at some point, among other things, depends on the previous decisions (or on the decision history). As an example consider employment decisions that a manufacturing company makes each month. The payroll cost this month depends in part on how many employees are already on the payroll which occurred as a result of hiring/firing decisions up to now. These are called sequential decisions. Decision trees are very useful for analyzing sequential decisions. We can regard the above oil field problem as a simple sequential decision consisting of stage 1) deciding whether to use the seismic test (when there is a cost, the answer is not automatically yes) and stage 2) which of the three actions to take based on the result of the test, if taken.

[Decision trees](file:///%5C%5Cacfiles%5Cakinc%5Cwww-home%5CFIN203%5Coil_tree.bmp) consist of decision nodes (represented by small squares), chance nodes (represented by small circles) and lines (arcs) drawn between them to display the chronology of what happens. Here is the decision tree for the oil problem. Pay special attention to the order of things happening. The first thing to be done is to decide if the seismic test is to be taken or not. This is represented by a square with two branches. The lower branch pictures “no test”. In this case, the decision maker is immediately confronted with the three choices: *drill (A*1)*, unconditionally* *lease* (*A*2) *or conditionally lease* (*A*3). At the end of *A*1 we have four branches representing the four states of nature. Since *A*2 leads to a pay-off of $.45M, no matter what the state is, we need not explicitly show the four states; and since the payoffs with *A*3 for states *S*3 and *S*4 are both zero we combined them.

If the test branch is taken, decision maker must observe the uncertain result: *E*1, *E*2 or *E*3, each can occur with the marginal probabilities computed. Each branch is then followed by the three decision alternatives (*A*1, *A*2, *A*3) which in turn are followed by the disposition of the field (states of nature) to be discovered. We have shown the terminal pay-offs at the end of the branches (terminals) or associated with arcs. For example the immediate cost of taking the ‘test’ branch is $.01M (a $-0.01M pay-off), and is so noted. We also have written in red the probabilities associated with the branches of the chance nodes. Creating the chronology of the decision situation like this from left to right is called “folding out.”

To **solve** the decision situation, we “fold back” by collapsing the branches from right to left. For chance nodes we collapse the branches by computing expected value. For instance, the expected value of *A*1 (the ‘drill’ option) following the *E*1 branch is .1944\*6.5 + .2500 \*2.0 + .3056\*(-0.25) + .2500\*(-0.75) = 1.4997. The decision nodes are collapsed by selecting the branch with the best expected outcome. For instance, if we ever came to the decision node following *E*2, we would choose *A*1 branch and ignore the other branches (shown as pruned branches). The decision problem is solved when we reach the far left decision node where the optimal value is $0.8496M. The optimal decision sequence is to take the seismic test, if *E*1 or *E*2 is observed “drill” otherwise ‘unconditionally lease.’

We can also see that the value of the seismic test (EVSI) is the difference between the expected values at the decision node following the ‘no test’ branch and the ‘test’ branch which, as before, 0.8596 -0.68 = 0.1796, greater than the .01M cost of it.