

# Value of Information: Lecture XXXII

Charles B. Moss

September 23, 2026

- 1 Decision Making and Bayesian Probabilities
- 2 Dynamic Value of Information

# Decision Making and Bayesian Probabilities

- Traditionally, Bayesian analysis involves a procedure whereby new information is integrated into a prior distribution to generate an updated or posterior distribution.
- At times the concept of Bayesian probability theory is confused with the subjective probability theory where an individual has an intuition regarding the probability of an event instead of a frequency view of probability theory which strives for an objective version of probability.

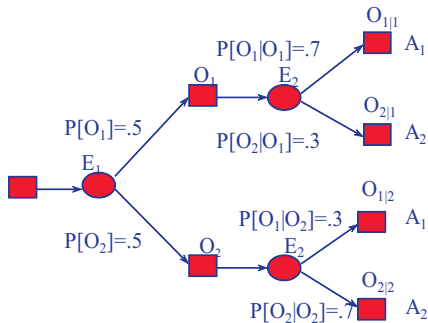
- At the base of Bayesian inference is Bayes's equation

$$P[a|b] = \frac{P[a, b]}{P[b]} \quad (1)$$

where  $P[a|b]$  is the probability of the event  $a$  occurring such that event  $a$  has already occurred,  $P[a, b]$  is the joint probability of both event  $a$  and event  $b$  occurring, and  $P[b]$  is the marginal probability of event  $b$  occurring.

- As depicted in Figure 1 the potential outcomes of two random events each of which have two potential outcomes.

Figure: Decision Tree



- The first event yields outcomes  $O_1$  and  $O_2$  each of which occurs with probability 0.5.
- The second event results in  $O_{1|1}$  and  $O_{2|1}$  if  $O_1$  occurred in the first event and  $O_{1|2}$  and  $O_{2|2}$  given that  $O_2$  occurred in the first event.

- Intuitively, we can picture  $O_{1|1}$  and  $O_{1|2}$  as the same event with a different intervening event. Similarly,  $O_{2|1}$  and  $O_{2|2}$  are the same event.
- What does change with the intervening event is the relative probability that each outcome will occur.
  - Given  $O_1$  the probability of outcome 1 in the second stage ( $O_{1|1}$ ) is 0.7 compared with a probability of outcome 2 in the second stage ( $O_{1|2}$ ) of 0.3.
  - The difference in the probability of the payoffs given the outcome in the first stage gives rise to the value of information.

- Next, we want to introduce two alternatives  $A_1$  which pays \$10 in state 1 and \$0 in state 2. Similarly,  $A_2$  pays \$5 in state 1 and \$4 in state 2.
  - To determine the expected value of the investment, we must first determine the probabilities that state 1 and state 2 will occur. For states 1 and 2 the total probabilities are

$$\begin{aligned}
 P[S_1] &= P[O_1] P[O_{1|1}] + P[O_2] P[O_{1|2}] \\
 &= 0.5(0.7) + 0.5(0.3) = 0.5 \\
 P[S_2] &= P[O_1] P[O_{2|1}] + P[O_2] P[O_{2|2}] \\
 &= 0.5(0.3) + 0.5(0.7) = 0.5
 \end{aligned} \tag{2}$$

The expected value of each action can then be derived as

$$\begin{aligned}
 E[A_1] &= P[S_1] V[A_1|S_1] + P[S_2] V[A_1|S_2] \\
 &= 0.5(10) + 0.5(0) = \$5.00 \\
 E[A_2] &= P[S_1] V[A_2|S_1] + P[S_2] V[A_2|S_2] \\
 &= 0.5(5) + 0.5(4) = \$4.50
 \end{aligned} \tag{3}$$

- Continued

- Thus, in the absence of risk aversion, the decision maker would choose  $A_1$  with an expected value of \$5.00.
- The next question is: What is the initial signal worth?
  - Starting from the last node and working backward, assume that  $O_1$  has occurred, what is the optimum decision?

$$\begin{aligned}
 E[A_1|O_1] &= P[O_{1|1}] V[A_1|S_1] + P[O_{2|1}] V[A_1|S_2] \\
 &= 0.7(10) + 0.3(0) = \$7.00 \\
 E[A_2|O_1] &= P[O_{1|1}] V[A_2|S_1] + P[O_{2|1}] V[A_2|S_2] \\
 &= 0.7(5) + 0.3(4) = \$4.70
 \end{aligned} \tag{4}$$

Thus, just like the scenario without the intervening event, we choose  $A_1$ .

- Continued

- Continued

- However, the result is somewhat different given that  $O_2$  occurs. Specifically

$$\begin{aligned} E[A_1|O_2] &= P[O_{1|2}] V[A_1|S_1] + P[O_{2|2}] V[A_1|S_2] \\ &= 0.3(10) + 0.7(0) = \$3.00 \\ E[A_2|O_2] &= P[O_{1|2}] V[A_2|S_1] + P[O_{2|2}] V[A_2|S_2] \\ &= 0.3(5) + 0.7(4) = \$4.30 \end{aligned} \quad (5)$$

- Under this scenario, we would choose  $A_2$  over  $A_1$ . The decision rule is then to choose  $A_1$  if event  $O_1$  occurs and action  $A_2$  if event  $O_2$  occurs. The expected value of this strategy is

$$\begin{aligned} V[\text{Strategy}] &= P[O_1] E[A_1|O_1] + P[O_2] E[A_2|O_2] \\ &= 0.5(7.00) + 0.5(4.30) = \$5.65 \end{aligned} \quad (6)$$

- Continued
  - Continued
    - The value of the information is then the difference between the conditioned and unconditioned decision

$$V[\text{Information}] = 5.65 - 5.00 = \$0.65 \quad (7)$$

# Dynamic Value of Information

- Chavas, Jean-Paul and Rulan Pope. 1984. Information: Its Measurement and Valuation *American Journal of Agricultural Economics* 66, 705-10.
  - The objective of the paper is to discuss the measurement and economic valuation of information.
- Concepts of Information
  - One way to define information is to focus on the entropy of the signal. Following Shannon and Weaver the entropy of a signal is defined as

$$H = - \sum_{i=1}^n p_i \ln(p_i) \quad (8)$$

- Continued

- Intuitively, as  $H$  increases the value of the signal decreases. Further  $H$  decreases as one event becomes increasing likely.
  - From the example in the previous lecture, the unconditioned probability has equally likely events. The value of the information in that distribution function is presented in Table 1.
  - Alternatively, the value of either conditioned distribution function is depicted in Table 2.

Table: Entropy of Information

| State              | Probability | $-p_i \ln(p_i)$ |
|--------------------|-------------|-----------------|
| $S_1$              | 0.50        | 0.3466          |
| $S_2$              | 0.50        | 0.3466          |
| $H =$              |             | 0.6931          |
| Note $\exp(H) = 2$ |             |                 |

Table: Entropoy of Information

| State                     | Probability | $-p_i \ln(p_i)$ |
|---------------------------|-------------|-----------------|
| $S_1$                     | 0.30        | 0.3612          |
| $S_2$                     | 0.70        | 0.2497          |
| $H =$                     |             | 0.6109          |
| Note $\exp(H) = 1.842023$ |             |                 |

- The next level of complexity is added by the concepts of prior and posterior probabilities.
  - Assume that we have two signals,  $p$  and  $q$ , each of which with two potential outcomes.
  - One question is: what is the value of the information in signal  $q$  given the information already observed in signal  $p$ . Their gives the value of such as signal as depicted in Table 3.

$$I(p, q) = \sum_{i=1}^n p_i \ln \left( \frac{p_i}{q_i} \right) \quad (9)$$

- If the intervening event would have been uninformative the results are depicted in Table 4.

Table: Information Content of Signal

| State | Probability $p$ | Probability $q$ | $p_i \ln(p_i/q_i)$ |
|-------|-----------------|-----------------|--------------------|
| $S_1$ | 0.50            | 0.70            | -0.1682            |
| $S_2$ | 0.50            | 0.30            | 0.2554             |
|       |                 | $I =$           | 0.0872             |

Table: Information Content of Uninformative Signal

| State | Probability $p$ | Probability $q$ | $p_i \ln(p_i/q_i)$ |
|-------|-----------------|-----------------|--------------------|
| $S_1$ | 0.50            | 0.50            | 0.0000             |
| $S_2$ | 0.50            | 0.50            | 0.0000             |
|       |                 | $I =$           | 0.0000             |

- Continued
  - "Unfortunately, there is in general no relationship between the entropy measure for information and the value of information in a particular decision-making process."
- Another statistical based definition of information comes from the estimation process via the likelihood function.
  - One estimation procedure involves choosing the value of parameters which maximizes the likelihood of a particular sample.
  - Under normality, the natural log of the likelihood for a linear regression can be written as

$$L = -\frac{T}{2} \ln \left( \frac{\left[ \sum_{t=1}^T \{ y_t - \alpha_0 - \alpha_1 x_t \} \right]^2}{T} \right) + \text{Constant} \quad (10)$$

- Continued

- The matrix of second moments is referred to as the information matrix. This matrix is the basis for statistical inference and yields such things as the standard error of the parameter estimates.
- "Information can also be defined as a message which alters tastes or perceptions which are certain."
- "Finally, information can be defined as a message which alters probabilistic perceptions of random events."

## A Model of Information

- Back to Equations: First, assume that economic agents possess a concave utility function  $U(y, x_1, x_2, e_1, e_2)$ . There goal is to choose the level of  $x_1$  and  $x_2$  which maximizes that level of utility

$$\max_{x_1, x_2} E_1 U(y, x_1, x_2, e_1, e_2) \quad (11)$$

- Making all decisions at time 1 corresponds to the "open loop" solution where no learning takes place.
- Given that  $e_1$  can be observed before decisions on  $x_2$  have to be made, the second phase of decision process can be rewritten as

$$V(x_1, y, e_1) = \max_{x_2} E_2 U(y, x_1, x_2, e_1, e_2) \quad (12)$$

- In other words, this process represents the optimum selection of  $x_2$  conditioned on the new information, or  $e_1$ . Given this formulation,  $x_1$  can then be selected to maximize the expectation of this value function

$$\max_{x_1} E_1 V(x_1, y, e_1) \quad (13)$$

- Based on these optimal actions

$$E_1 \max_{x_1} V(x_1, y, e_1) \geq \max_{x_1} E_1 V(x_1, y, e_1) \quad (14)$$

- Assuming that information is costly, the most that can be paid for the information is  $B$  defined by

$$B_1 = \left\{ -\frac{1}{2} \left[ \frac{\partial^2 V}{\partial E_1 \partial x_1} \left( \frac{\partial^2 V}{\partial x_1^2} \right)^{-1} \frac{\partial^2 V}{\partial x_1 \partial e_1} \right] / \frac{\partial V}{\partial y} \right\} V(e_1) \quad (15)$$