Value of Information

IE 5545: Decision Analysis
Prof. Krishnamurthy Iyer*

May 2, 2026

Often a decision maker must decide whether to acquire new information and at what cost. To do this, one must assign a value to the information. To do this, we first observe that, typically, we value information because it allows us to take better actions. This in turn suggests that the value of any information must depend on the decision problem for which it will be used.

Recall how a Bayesian decision maker incorporates new information while making decisions: the decision maker uses the signal structure, as well as the realized signal, to update their beliefs, and then takes the action that maximizes her expected utility, where the expectation is with respect to her updated posterior belief.

From this, we can infer that the value of an information, given a decision problem, is the increase in the decision maker's expected utility from receiving the information.

$$\text{Value of information} = \begin{pmatrix} \text{Expected payoff from acting} \\ \text{optimally } \textbf{with} \text{ the information} \end{pmatrix} - \begin{pmatrix} \text{Expected payoff from acting} \\ \text{optimally } \textbf{without} \text{ the information} \end{pmatrix}$$

We will formalize this definition next.

1 Value of information

Suppose the decision maker faces a decision problem with incomplete information. Let Θ denote the set of states, and μ be the decision maker's prior belief. Let $\mathcal{S}$ denote the set of actions, and let $\pi(\theta, s)$ denote the (state-dependent) payoff at state θ when the decision maker takes action s.

The decision maker's optimal expected payoff before receiving any additional information is given by

$$\Pi(\mathsf{without\ info}) = \max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)], \tag{1}$$

where $\mathbf{E}_\mu[\pi(\bar{\theta}, s)] = \mathbf{E}[\pi(\bar{\theta}, s) | \bar{\theta} \sim \mu]$.

Now suppose the decision maker can choose to purchase a new source of information at a cost. Let this information be represented by the signal structure $(\mathcal{X}, \mathbb{P})$. How much should the decision maker be willing to pay for this information?

To answer this question, suppose the decision maker decides to purchase this information, and the realized signal is $\bar{x} = x_\ell \in \mathcal{X}$. Then, the decision maker's belief updates to

$$\mu_{\mathsf{new}}(\theta) = \mathbf{P}(\bar{\theta} = \theta | \bar{x} = x_\ell).$$

(This right-hand-side can be found using Bayes' rule.) With this belief, the expected payoff of the decision maker for choosing an action s is given by $\mathbf{E}_\mu[\pi(\bar{\theta}, s)|\bar{x} = x_\ell]$. Hence, the decision maker's optimal expected payoff on seeing any signal $\bar{x} = x_\ell$ is given by

$$\Pi(x_\ell) = \max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)|\bar{x} = x_\ell].$$

Of course, the decision maker does not get to see the signal $\bar{x} = x_\ell$ before purchasing the information. So, the decision maker does not know beforehand what signal she will see. However, given that the signal structure $(\mathcal{X}, \mathbb{P})$ is known to the decision maker, she believes that any signal $\bar{x} = x_\ell$ is realized with probability

$$\mathbf{P}(\bar{x} = x_\ell) = \sum_{\theta \in \Theta} \mu(\theta)\mathbb{P}(\bar{x} = x_\ell | \bar{\theta} = \theta).$$

(The above is an application of the law of total probability.)

Taken together, the decision maker's optimal expected payoff on having access to the information is given by

$$\begin{aligned}
\Pi(\mathsf{info}) &= \mathbf{E}[\Pi(\bar{x})] \\
&= \sum_{x_\ell \in \mathcal{X}} \mathbf{P}(\bar{x} = x_\ell)\Pi(x_\ell) \\
&= \sum_{x_\ell \in \mathcal{X}} \mathbf{P}(\bar{x} = x_\ell) \left(\max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)|\bar{x} = x_\ell] \right) \\
&= \mathbf{E}\left[\max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)|\bar{x}] \right]
\end{aligned}$$

Now, we can define the value of the information $\mathcal{V}(\mathcal{X}, \mathbb{P})$ as follows.

$$\begin{aligned}
\mathcal{V}(\mathcal{X}, \mathbb{P}) &= \begin{pmatrix} \text{Expected payoff from acting} \\ \text{optimally } \textbf{with} \text{ the information} \end{pmatrix} - \begin{pmatrix} \text{Expected payoff from acting} \\ \text{optimally } \textbf{without} \text{ the information} \end{pmatrix} \\
&= \Pi(\mathsf{info}) - \Pi(\mathsf{without\ info}) \\
&= \mathbf{E}\left[\max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)|\bar{x}] \right] - \max_{s \in \mathcal{S}} \mathbf{E}_\mu[\pi(\bar{\theta}, s)].
\end{aligned}$$

The above expression describes how much the information (represented by $(\mathcal{X}, \mathbb{P})$) is worth to the decision maker. Consequently, this is the maximum that the decision maker should be willing to pay for obtaining this information.

In the next section, we illustrate this formal setup with an example. We end this section with few remarks:

1. The value of information depends on the underlying decision problem. The same signal structure may have different value for different decision problems.

2. The value of information can never be negative. This is because the decision maker can always ignore the additional information and choose the action that was optimal without the information. Formally, let s^* denote the optimal action without the information:

$$\Pi(\mathsf{without\ info}) = \max_{s \in \mathcal{S}} \mathbf{E}_\mu \left[\pi(\bar{\theta}, s) \right] = \mathbf{E}_\mu \left[\pi(\bar{\theta}, s^*) \right].$$

Now, we have

$$
\begin{aligned}
\Pi(\mathsf{info}) &= \mathbf{E}\left[\max_{s\in\mathcal{S}} \mathbf{E}_\mu[\pi(\bar\theta, s)|\bar x]\right] \\
&\geq \mathbf{E}\left[\mathbf{E}_\mu[\pi(\bar\theta, s^*)|\bar x]\right] \\
&= \sum_{x_\ell\in\mathcal{X}} \mathbf{P}(\bar x = x_\ell)\left(\sum_{\theta\in\theta} \mathbf{P}(\bar\theta=\theta|\bar x = x_\ell)\pi(\theta, s^*)\right) \\
&= \sum_{x_\ell\in\mathcal{X}}\sum_{\theta\in\theta} \mathbf{P}(\bar x = x_\ell)\mathbf{P}(\bar\theta=\theta|\bar x = x_\ell)\pi(\theta, s^*) \\
&= \sum_{\theta\in\theta}\left(\sum_{x_\ell\in\mathcal{X}} \mathbf{P}(\bar x = x_\ell)\mathbf{P}(\bar\theta=\theta|\bar x = x_\ell)\right)\pi(\theta, s^*) \\
&= \sum_{\theta\in\theta} \mathbf{P}(\bar\theta=\theta)\pi(\theta, s^*) \quad \text{(by law of total probability)} \\
&= \sum_{\theta\in\theta} \mu(\theta)\pi(\theta, s^*) \quad \text{(since } \mu \text{ is the prior belief)} \\
&= \mathbf{E}_\mu[\pi(\bar\theta, s^*)] \\
&= \Pi(\mathsf{without\ info}).
\end{aligned}
$$

Thus, $\mathcal{V}(\mathcal{X}, \mathbb{P}) = \Pi(\mathsf{info}) - \Pi(\mathsf{without\ info}) \geq 0$.

3. If the signal structure $(\mathcal{X}, \mathbb{P})$ reveals full information, i.e., $\mathcal{X} = \Theta$ and $\mathbb{P}(\bar x = \theta|\bar\theta = \theta) = 1$ for all $\theta \in \Theta$, then upon receiving the signal $\bar x = \theta$, the decision maker's optimal payoff is given by $\max_{s\in\mathcal{S}} \pi(\theta, s)$. Thus, the value of full information is given by

$$
\mathcal{V}(\mathsf{full\ info}) = \mathbf{E}_\mu\left[\max_{s\in\mathcal{S}} \pi(\bar\theta, s)\right] - \max_{s\in\mathcal{S}} \mathbf{E}_\mu\left[\pi(\bar\theta, s)\right].
$$

   The value of any other signal structure must be less than or equal to the value of full information. (Exercise: Argue why?) In other words, full information has the maximum value.

4. The value of information is zero, if the decision maker's optimal action never changes irrespective of the realized signal. In particular, the value of "no information" is zero.

1.1 Example: Retail promotion

Consider an online retailer who must decide whether to offer a user a discount on an item being sold on their website. The regular price for the item is given by $r = \$10$. If the discount is offered, the price is $d = \$8$. The discount has the effect of increasing the customer's probability of purchasing the item: if the customer purchases the item at a regular price with probability $\theta \in [0, 1]$, then with the discount, she will purchase it with probability $H(\theta) = \frac{1+\theta}{2} \geq \theta$. The retailer wants to maximize her expected revenue R, where $R = r \cdot \theta = 10\theta$ if no discount is offered, and $R = d \cdot H(\theta) = 4(1 + \theta)$ if a discount is offered. However, the retailer does not know the probability θ with which the user will purchase an item without a discount.

We have the set of states as $\Theta = [0, 1]$ denoting all possible values of the probability with which any user will purchase the item when offered the regular price. Suppose the retailer's *prior belief* μ is specified by the distribution with density $f(u) = 6u(1-u)$ for $u \in [0, 1]$. (Note that the prior is a $\mathsf{Beta}(2, 2)$ distribution.)

Now, suppose the retailer obtains some addition information about this specific user in the form of their purchase history. This information reveals the number of times the user purchased the item in the past five times the regular price was offered. What is the value of this information to the retailer? (In other words, how much is the retailer willing to pay for access to this information?)

We begin by stating the signal structure underlying the information provided. We know that of the past $N = 5$ times the user was offered the regular price, they could have bought the item anywhere from 0 times to 5 times. So, we can represent the set of signals as $\mathcal{X} = \{0, 1, 2, 3, 4, 5\}$, where $\bar{x}$ denotes the number of times the user bought the item. What are the likelihoods $\mathbb{P}$? If we knew the state is $\bar{\theta} = \theta$, then the conditional distribution of $\bar{x}$ is just the $\mathsf{Binomial}(N, \theta)$ distribution, with

$$\mathbb{P}(\bar{x} = k | \bar{\theta} = \theta) = \binom{N}{k} \theta^k (1 - \theta)^{N-k},$$

for each $k \in \mathcal{X}$.

To compute the value of information, we must find out the retailer's optimal payoff with and without this information. We begin with the case without the information.

1.1.1 Without information

Without the purchase history information, the retailer's payoff (expected revenue) at each state is given by

$$\pi(\theta, s) = \begin{cases} r \cdot \theta & \text{if } s = r; \\ d \cdot H(\theta) & \text{if } s = d. \end{cases}$$

Then, the retailer computes the expected payoff $\Pi(s)$ by taking expectation over the unknown state θ with respect to her prior belief μ:

$$\begin{aligned} \Pi(s) &= \mathbf{E}_\mu[\pi(\bar{\theta}, s)] \\ &= \begin{cases} r \cdot \mathbf{E}_\mu[\bar{\theta}] & \text{if } s = r; \\ d \cdot \mathbf{E}_\mu[H(\bar{\theta})] & \text{if } s = d. \end{cases} \end{aligned}$$

Since μ has the density $f(x) = 6x(1 - x)$, we can compute the two expectations as

$$\begin{aligned} \mathbf{E}_\mu[\bar{\theta}] &= \int_0^1 x f(x) dx = \frac{1}{2} \quad \text{(by symmetry)} \\ \mathbf{E}_\mu[H(\bar{\theta})] &= \frac{1 + \mathbf{E}_\mu[\bar{\theta}]}{2} = \frac{3}{4}. \end{aligned}$$

Thus, we obtain $\Pi(r) = 10\left(\frac{1}{2}\right) = 5$ and $\Pi(d) = 8\left(\frac{3}{4}\right) = 6$. Since $\Pi(d) > \Pi(r)$, the optimal action for the retailer is to offer the discount. Furthermore, we have $\Pi(\mathsf{without\ info}) = 6$.

1.1.2 With information

To compute the retailer's optimal expected payoff with the information, we must first find out the retailer's posterior belief for each possible value of the signal realization, i.e., for each possible number of times the user could have purchased the item out of the past five times she was offered the regular price.

To do this, we use the fundamental relationship (or the Bayes' rule). Let $\bar{x} = k \in \mathcal{X} = \{0, 1, 2, 3, 4, 5\}$. Let f_k denote the density of the posterior belief μ_k after seeing $\bar{x} = k$. We have using the fundamental relationship,

$$f_k(\theta) \propto f(\theta) \cdot \mathbb{P}(\bar{x} = k | \bar{\theta} = \theta).$$

Letting C denote the constant of proportionality, we have

$$\begin{aligned} f_k(\theta) &= C \cdot f(\theta) \cdot \mathbb{P}(\bar{x} = k | \bar{\theta} = \theta) \\ &= C \cdot 6\theta(1-\theta) \cdot \binom{N}{k} \theta^k (1-\theta)^{N-k} \\ &= 6C \binom{N}{k} \cdot \theta^{k+1}(1-\theta)^{N-k+1}. \end{aligned}$$

We can see that the posterior density has the form of a $\mathsf{Beta}(\alpha, \beta)$ distribution with parameters $\alpha = k+2$ and $\beta = N - k + 2$. Using this, we obtain the density to be

$$f_k(\theta) = \frac{1}{B(k+2, N-k+2)} \cdot \theta^{k+1}(1-\theta)^{N-k+1}, \quad \text{for all } \theta \in [0, 1].$$

Remark 1 (Probability refresher). *A* $\mathsf{Beta}(\alpha, \beta)$ *distribution has the density over* $[0, 1]$ *given by*

$$f_{\alpha,\beta}(u) = \frac{1}{B(\alpha, \beta)} u^{\alpha-1}(1-u)^{\beta-1},$$

where $B(\alpha, \beta)$ *is a* normalization constant *that ensures* $\int_0^1 f_{\alpha,\beta}(u)du = 1$. *In particular, we have*

$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du.$$

For integer α, β, *the normalization constant* $B(\alpha, \beta)$ *is explicitly given by*

$$B(\alpha, \beta) = \frac{(\alpha-1)!(\beta-1)!}{(\alpha+\beta-1)!}.$$

Finally, the mean of the $\mathsf{Beta}(\alpha, \beta)$ *distribution is given by* $\frac{\alpha}{\alpha+\beta}$. $\square$

Thus, the mean of f_k is given by

$$\mathbf{E}_\mu[\bar{\theta} | \bar{x} = k] = \int_0^1 \theta f_k(\theta) d\theta = \text{mean of } \mathsf{Beta}(k+2, N-k+2) = \frac{k+2}{(k+2) + (N-k+2)} = \frac{k+2}{N+4}.$$

Furthermore, we have

$$\begin{aligned} \mathbf{E}_\mu[H(\bar{\theta}) | \bar{x} = k] &= \frac{1 + \mathbf{E}_\mu[\bar{\theta} | \bar{x} = k]}{2} = \frac{1 + \frac{k+2}{N+4}}{2} = \frac{N+k+6}{2(N+4)} \\ \Pi(r | \bar{x} = k) &= r \cdot \mathbf{E}_\mu[\bar{\theta} | \bar{x} = k] = r \cdot \frac{k+2}{N+4} \\ \Pi(d | \bar{x} = k) &= d \cdot \mathbf{E}_\mu[H(\bar{\theta}) | \bar{x} = k] = d \cdot \frac{N+k+6}{2(N+4)}. \end{aligned}$$

k	$\mathbf{E}_k[\bar{\theta}]$	$\mathbf{E}_k[H(\bar{\theta})]$	$\Pi_k(r)$	$\Pi_k(d)$	Optimal action	Optimal payoff $\Pi(k)$
0	$2/9$	$11/18$	$20/9$	$44/9$	d	$44/9$
1	$3/9$	$12/18$	$30/9$	$48/9$	d	$48/9$
2	$4/9$	$13/18$	$40/9$	$52/9$	d	$52/9$
3	$5/9$	$14/18$	$50/9$	$56/9$	d	$56/9$
4	$6/9$	$15/18$	$60/9$	$60/9$	$\{r,d\}$	$60/9$
5	$7/9$	$16/18$	$70/9$	$64/9$	r	$70/9$

Table 1: Optimal actions and payoffs after seeing $\bar{x} = k$ for different values of k. Here $\mathbf{E}_k[\bar{\theta}] = \mathbf{E}_\mu[\bar{\theta}|\bar{x} = k]$ and $\mathbf{E}_k[H(\bar{\theta})] = \mathbf{E}_\mu[H(\bar{\theta})|\bar{x} = k]$. Also, $\Pi_k(r) = \Pi(r|\bar{x} = k)$ and $\Pi_k(d) = \Pi(d|\bar{x} = k)$.

Thus, we can write the optimal expected payoff upon seeing the signal $\bar{x} = k$ as

$$\Pi(k) = \max\{\Pi(r|\bar{x} = k), \Pi(d|\bar{x} = k)\} = \max\left\{\frac{r(k+2)}{N+4}, \frac{d(N+k+6)}{2(N+4)}\right\}.$$

Using the above expressions and the fact that $N = 5$, we can now write down the optimal action and the payoff of the retailer for each possible value of k, as in Table 1.

Now that we know the optimal payoff after seeing any signal $\bar{x} = k$, we can compute our optimal expected payoff before seeing the signal. For this, we need the probability $\mathbf{P}(\bar{x} = k)$ of seeing any signal k. We can compute this using the law of total probability as follows:

$$\begin{aligned}
\mathbf{P}(\bar{x} = k) &= \int_0^1 \mathbf{P}(\bar{x} = k|\bar{\theta} = \theta) \cdot f(\theta) d\theta \\
&= \int_0^1 \binom{N}{k} \theta^k (1-\theta)^{N-k} \cdot 6\theta(1-\theta) d\theta \\
&= 6\binom{N}{k} \cdot \int_0^1 \theta^{k+1}(1-\theta)^{N-k+1} d\theta \\
&= 6\binom{N}{k} \cdot B(k+2, N-k+2) \qquad \text{(from the definition of } B(\alpha,\beta)) \\
&= 6\binom{N}{k} \cdot \frac{(k+1)!(N-k+1)!}{(N+3)!} \qquad \text{(using the explicit expression for } B(\alpha,\beta)) \\
&= 6\frac{N!}{k!(N-k)!} \cdot \frac{(k+1)!(N-k+1)!}{(N+3)!} \\
&= 6\frac{(k+1)(N-k+1)}{(N+1)(N+2)(N+3)}.
\end{aligned}$$

Once again, plugging $N = 5$, we obtain the probabilities for seeing each signal k. These are tabulated in Table 2.

Thus, we obtain the retailer's optimal expected payoff with the information as

$$\begin{aligned}
\Pi(\mathsf{info}) &= \sum_{k=0}^{5} \mathbf{P}(\bar{x} = k)\Pi(k) \\
&= 3/28 \cdot 44/9 + 5/28 \cdot 48/9 + 6/28 \cdot 52/9 + 6/28 \cdot 56/9 + 5/28 \cdot 60/9 + 3/28 \cdot 70/9 \\
&= \frac{85}{14}.
\end{aligned}$$

k	$\mathbf{P}(\bar{x} = k)$	Optimal payoff $\Pi(k)$
0	$3/28$	$44/9$
1	$5/28$	$48/9$
2	$6/28$	$52/9$
3	$6/28$	$56/9$
4	$5/28$	$60/9$
5	$3/28$	$70/9$

Table 2: Probability of seeing each signal, along with the corresponding optimal payoff.

1.1.3 Value of information

Thus, the value of the information contained in the purchase history is given by

$$\mathcal{V} = \Pi(\mathsf{info}) - \Pi(\mathsf{without\ info}) = \frac{85}{14} - 6 = \frac{1}{14} > 0.$$

Thus, with access to the information, the retailer can improve her expected payoff by $1/14$ (from this user). In other words, the retailer should be willing to pay at most \$(1/14) for this information.