Bayesian Decision Theory

IE 5545: Decision Analysis
Prof. Krishnamurthy Iyer*

May 2, 2026

1 Bayesian modeling of uncertainty

To incorporate epistemic uncertainty into our framework, we will adopt the Bayesian framework. This framework represents epistemic uncertainty in the form of "beliefs" of the decision maker. These beliefs are probability distributions over the set of possibilities, and capture how likely it is for each possibility to occur, as believed by the decision maker.

The first step of the Bayesian framework involves identifying *all* the different ways in which the underlying uncertainty can get resolved. Any particular way of resolving the uncertainty is referred to as a *state*. Each state resolves the uncertainty in a different way, and specifying the set of all states then captures all possible ways in which the uncertainty could be resolved. We let Θ denote the set of all possible states, and denote the true (but unknown) state as $\bar{\theta}$. The set of states Θ can be finite or infinite.

Example (Pricing)**.** Consider a seller who must decide the price for an item. The seller incurs a unit cost $c = 0.25$ for the item, and wants to set a price to maximize the profit, given by $(p - c) \cdot D(p)$, where $D(p)$ is the demand at price p. However, the seller is uncertain about the demand for the item. The decision maker knows that either the item will be a hit, or it will be a flop. If the item is a hit, the demand for the item at price p is given by $D_{\mathsf{hit}}(p) = \max\{100 - 25p, 0\}$. If the item is a flop, the demand for the item is $D_{\mathsf{flop}}(p) = \max\{80 - 40\sqrt{p}, 0\}$.

In this example, the set of states is finite and is given by $\Theta = \{\mathsf{hit}, \mathsf{flop}\}$. Note that in either state, the demand is zero if $p > 4$. $\square$

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 $p \in [0, 1]$, then with the discount, she will purchase it with probability $H(p) = \frac{1+p}{2} \geq p$. The retailer wants to maximize her expected revenue R, where $R = 10p$ if no discount is offered, and $R = 8H(p) = 4(1 + p)$ if a discount is offered. However, the retailer does not know the probability p with which the user will purchase an item without a discount.

In this example, the set of states is all the possible values of $p \in [0, 1]$. If the retailer has no information about p, then the set of states is infinite and is given by $\Theta = [0, 1]$. $\square$

Once the set of states Θ is identified, the second step requires the decision maker to determine how likely each state is. To do this, the decision maker uses all the information at her disposal, and comes up with a probability distribution μ over Θ. This distribution is referred to as her *belief*. When Θ is finite, i.e., $\Theta = \{\theta_1, \theta_2, \ldots, \theta_K\}$, we represent the belief μ as a probability mass function (pmf), with $\mu(\theta_k)$ denoting the probability that the decision maker assigns to the state θ_k. Note that, in this case, we have $\mu(\theta_k) \geq 0$ and $\sum_{k=1}^{K} \mu(\theta_k) = 1$.

Note: An alternative representation for μ, when Θ is a set of real numbers is as the cumulative distribution function (cdf). When Θ is an interval of real numbers, we may also represent it using probability density functions (pdf).

We denote the set of all beliefs, i.e., the set of all distributions over Θ, as $\Delta(\Theta)$.

Example (Pricing (contd.))**.** We can represent the seller's belief about the state by a distribution μ over $\Theta = \{\mathsf{hit}, \mathsf{flop}\}$. For example, if the seller believes that that the item is twice as likely to be a hit as it is likely to be a flop, then we have the belief μ as

$$\mu(\mathsf{hit}) = \mathbf{P}(\bar{\theta} = \mathsf{hit}) = \frac{2}{3} \qquad\qquad \mu(\mathsf{flop}) = \mathbf{P}(\bar{\theta} = \mathsf{flop}) = \frac{1}{3}.$$

□

Example (Retail promotion (contd.))**.** The retailer's belief about the probability p is specified by a distribution μ over the set $[0, 1]$. In this case, we can represent this distribution as a cumulative distribution function (cdf) F, where $F(x) = \mathbf{P}(p \leq x)$. If the cdf F is differentiable, we can equally represent μ using the probability density function (pdf) f, which is the derivative of the cdf F, i.e., $f(x) = F'(x)$ for all x. Note that, if the pdf f is known, then cdf F can be recovered using the formula $F(x) = \int_{-\infty}^{x} f(u) du$.

For example, if the retailer believes that all values of p in $[0, 1]$ are equally likely, then the cdf F of her belief μ is given by $F(x) = x$ for all $x \in [0, 1]$. In this case, the pdf is given by $f(x) = 1$ for all x. Alternatively, suppose the retailer believes that larger values of p are more likely than smaller values. One particular distribution that captures this belief has the pdf $f(x) = 2x$, and consequently, the cdf $F(x) = x^2$. □

With these two steps, we can now extend our decision making framework to include settings where the decision maker has incomplete information. In this extension, we let the decision-maker's utility function u and the rules $\bar{p}$ relating actions to outcomes both depend on the state $\theta \in \Theta$. Formally, a decision problem requires specifying

1. the set of states Θ representing the uncertainty faced by the decision maker,
2. a probability distribution $\mu \in \Delta(\Theta)$ that represents the decision maker's *belief* about the true state $\bar{\theta}$,
3. a set of actions $\mathcal{S} = \{s_1, s_2, \ldots, s_n\}$,
4. a set of outcomes $\mathcal{O} = \{o_1, o_2, \ldots, o_m\}$,
5. the rules relating actions to outcomes: a function $\bar{p}_\theta : \mathcal{S} \to \Delta(\mathcal{O})$ for each $\theta \in \Theta$. In other words, if the state is $\theta \in \Theta$, then an action $s \in \mathcal{S}$ results in an outcome that is drawn from the distribution $\bar{p}_\theta(s) \in \Delta(\mathcal{O})$, and
6. a utility function $u : \Theta \times \mathcal{O} \to \mathbb{R}$ representing the decision maker's preferences over $\Delta(\mathcal{O})$. In other words, if the state is Θ, then the decision maker's utility in outcome $o \in \mathcal{O}$ is given by $u(\theta, o)$. (We will sometimes denote the utility function in state θ as u_θ.)

How does the decision maker evaluate her preferences in the presence of incomplete information? Note that, if the true state θ were known, the decision maker knows her preferences exactly. In particular, as per the expected utility framework, the decision maker's preference in a state θ for an outcome that is drawn from a distribution $\bar{p} \in \Delta(\mathcal{O})$ is captured by the expected utility $\mathbf{E}[u(\theta, X)|X \sim \bar{p}]$. However, the main issue that we are trying to address is that the decision maker does not know the true state θ, and hence does not know her true preferences.

To resolve this conundrum, we invoke one of the central assumptions of Bayesian decison making. This assumption states that the decision maker's preferences are represented by her expected utility, where the expectation is taken over *both* the random outcomes as well as the unknown state. The latter expectation is taken with respect to her *beliefs*. In other words, for each action $s \in \mathcal{S}$, the decision maker evaluates her expected utility $\mathbf{E}[u(\theta, X)|X \sim \bar{p}_\theta(s), \theta \sim \mu]$ for the action, where the expectation is both

1. over the outcomes $X \in \mathcal{O}$, drawn from $\bar{p}_\theta(s)$ in state θ; and
2. over the states $\theta \in \Theta$, drawn according to her belief μ.

We refer to such a decision maker as *Bayesian*.

Note that, in the case of finite $\Theta = \{\theta_1, \ldots, \theta_K\}$ and finite $\mathcal{O} = \{o_1, \ldots, o_m\}$, we have

$$\begin{aligned}\mathbf{E}[u(\theta, X)|X \sim \bar{p}_\theta(s), \theta \sim \mu] &= \sum_{k=1}^{K} \mu(\theta_k)\mathbf{E}[u(\theta_k, X)|X \sim \bar{p}_{\theta_k}(s)] \\ &= \sum_{k=1}^{K} \mu(\theta_k) \sum_{i=1}^{m} \bar{p}_{\theta_k}(s, o_i) \cdot u(\theta_k, o_i), \end{aligned} \tag{1}$$

where $\bar{p}_{\theta_k}(s, o_i)$ is the probability with which the action s leads to outcome o_i in state θ_k.

To explain this Bayesian model further, we can consider the setting in two steps. First, suppose the decision maker thinks they are in a particular state $\theta \in \Theta$. Then, analogous to the case without incomplete information, we can define a *state-dependent* payoff function $\pi(\theta, s)$, which denotes the decision maker's expected utility in state $\theta \in \Theta$ when action $s \in \mathcal{S}$ is chosen:

$$\begin{aligned}\pi(\theta, s) &= \mathbf{E}[u(\theta, X)|X \sim \bar{p}_\theta(s)] \\ &= \sum_{i=1}^{m} \bar{p}_\theta(s, o_i)u(\theta, o_i).\end{aligned}$$

Note that the above payoff function $\pi : \Theta \times \mathcal{S} \to \mathbb{R}$ depends on both the unknown state θ and the action chosen. To compute her overall expected utility, a Bayesian decision maker takes an expectation of the state-dependent payoff over the unknown state:

$$\mathbf{E}[\pi(s, \theta)|\theta \sim \mu] = \sum_{k=1}^{K} \mu(\theta_k)\pi(s, \theta_k). \tag{2}$$

It is not hard to verify that the right-hand-side of the two equations (1) and (2) are equal.

The left-hand-side of Eqn. (2) is the expected payoff of the decision maker, with respect to her beliefs μ, when she takes the action $s \in \mathcal{S}$. We will sometimes denote this by $\Pi(s)$ (or $\Pi_\mu(s)$ if we want to show the dependence on the belief).

How should the decision maker decide which action to choose? Under the Bayesian framework, since the decision maker's preferences are represented by her expected utility over both outcomes and states, a rational Bayesian decision maker must choose the action s that maximizes her expected utility $\mathbf{E}[u(\theta, X)|X \sim \bar{p}_\theta(s), \theta \sim \mu]$. Given our definition of the payoff function, we can state this simply as the rational Bayesian decision maker chooses the action s that maximizes her expected payoff $\Pi(s)$.

Example (Pricing (contd).)**.** Consider again the seller's pricing problem in the above framework.

1. We have already stated the set of states as $\Theta = \{\mathsf{hit}, \mathsf{flop}\}$, and suppose the seller's belief μ is given by $\mu(\mathsf{hit}) = \frac{2}{3}$ (and hence $\mu(\mathsf{flop}) = \frac{1}{3}$).

2. The set of actions can be taken as $\mathcal{S} = [0, 4]$, since the price can only take non-negative values, and the seller will never set a price strictly greater than 4, since the demand is then zero regardless of the state.

3. Since this is a deterministic problem, we can let the price to be the outcome itself. Thus $\mathcal{O} = [0, \infty)$.

4. Due to the same reason, the distribution $\bar{p}_\theta(s)$ is deterministically equal to s, i.e., when the price is s and the state is θ, the outcome is s with probability 1.

5. Since the seller wants to maximize the profit, the utility $u(\theta, p)$ when the state is θ and the price is p is given by

$$u(\theta, p) = (p - c)D_\theta(p) = \begin{cases} (p-c)D_{\mathsf{hit}}(p) & \text{if } \theta = \mathsf{hit}; \\ (p-c)D_{\mathsf{flop}}(p) & \text{if } \theta = \mathsf{flop}. \end{cases}$$

How does a rational Bayesian seller set the price? To compute the optimal price, the seller first computes the (state-dependent) payoff function $\pi(\theta, p)$, which is the expected utility at price p when the state is θ; in this problem, there is no other uncertainty once the state is known, so we have the payoff function same as the utility function: $\pi(\theta, p) = u(\theta, p)$. After this, the seller computes the expected payoff $\Pi(p)$, taking expectation over the unknown state θ:

$$\begin{aligned}
\Pi(p) &= \mathbf{E}[\pi(\theta, p)|\theta \sim \mu] \\
&= \mathbf{E}[u(\theta, p)|\theta \sim \mu] \\
&= \mu(\mathsf{hit})u(\mathsf{hit}, p) + \mu(\mathsf{flop})u(\mathsf{flop}, p) \\
&= \mu(\mathsf{hit}) \cdot (p-c)D_{\mathsf{hit}}(p) + \mu(\mathsf{flop}) \cdot (p-c)D_{\mathsf{flop}}(p) \\
&= (p-c)\left(\mu(\mathsf{hit})D_{\mathsf{hit}}(p) + \mu(\mathsf{flop})D_{\mathsf{flop}}(p)\right) \\
&= (p - 0.25)\left(\frac{2}{3}(100 - 25p) + \frac{1}{3}(80 - 40\sqrt{p})\right) \\
&= \frac{1}{12}(4p-1)(280 - 50p - 40\sqrt{p})
\end{aligned}$$

Optimizing this function over $p \in [0, 4]$, we obtain the optimal price to be $p^* \approx 2.09$.

We end the example with few remarks:

1. The optimal price p^* depends on the seller's belief. If the seller's belief were different, so would be the optimal price. In fact, it is not hard to compute the optimal price as a function of $\mu(\mathsf{hit})$.

2. The optimal price p^* is *neither* the optimal price if the state is $\bar{\theta} = \mathsf{hit}$ *nor* the optimal price if the state is $\bar{\theta} = \mathsf{flop}$. (If $\bar{\theta} = \mathsf{hit}$, the optimal price is $17/8 = 2.125$, and if $\bar{\theta} = \mathsf{flop}$, the optimal price is ≈ 1.94.) Thus, a Bayesian decision maker does not choose the optimal action in the most-likely state, but instead chooses the action that maximizes the expected utility with respect to her belief. □

Example (Retail promotion (contd).). Consider the retailer's promotional pricing problem in the Bayesian decision making framework:

1. We have the set of states as $\Theta = [0, 1]$ denoting all possible values of the probability with which the user will purchase the item when offered the regular price. Suppose the retailer's belief is specified by the distribution with density $f(x) = 6x(1-x)$ for $x \in [0, 1]$. (Under this belief, values close to 0.5 are more likely compared to either extremes.)
2. The set of actions can be taken as the price offered $\mathcal{S} = \{r, d\}$. Recall that $r = \$10$ and $d = \$8$.
3. To determine the seller's revenue, we need to know the price as well as whether the user purchased the item. Thus, we can define an outcome as (a, b) where $a \in \{r, d\}$ and $b \in \{\mathsf{buy}, \mathsf{not\ buy}\}$. Thus, $\mathcal{O} = \{(r, \mathsf{buy}), (d, \mathsf{buy}), (r, \mathsf{not\ buy}), (d, \mathsf{not\ buy})\}$.
4. The rules can be described as follows: $\bar{p}_\theta(s)$ is a distribution over $\mathcal{O}$ when price is s and the state is θ. We can state this distrbution by specifying the probabilities $\bar{p}_\theta(s, o)$ of each outcome $o \in \mathcal{O}$:

$$\bar{p}_\theta(r, o) = \begin{cases} \theta & \text{if } o = (r, \mathsf{buy}) \\ 1-\theta & \text{if } o = (r, \mathsf{not\ buy}) \\ 0 & \text{if } o = (d, \mathsf{buy}) \\ 0 & \text{if } o = (d, \mathsf{not\ buy}) \end{cases}$$

$$\bar{p}_\theta(d, o) = \begin{cases} 0 & \text{if } o = (r, \mathsf{buy}) \\ 0 & \text{if } o = (r, \mathsf{not\ buy}) \\ H(\theta) = \frac{1+\theta}{2} & \text{if } o = (d, \mathsf{buy}) \\ 1 - H(\theta) = \frac{1-\theta}{2} & \text{if } o = (d, \mathsf{not\ buy}) \end{cases}$$

5. The utility function of the revenue-maximizing retailer can be written as

$$u(\theta, o) = \begin{cases} r & \text{if } o = (r, \mathsf{buy}) \\ 0 & \text{if } o = (r, \mathsf{not\ buy}) \\ d & \text{if } o = (d, \mathsf{buy}) \\ 0 & \text{if } o = (d, \mathsf{not\ buy}) \end{cases}$$

The retailer's payoff at each state is given by

$$\begin{aligned} \pi(\theta, s) &= \mathbf{E}[u(\theta, X) | X \sim \bar{p}_\theta(s)] \\ &= \begin{cases} \theta \cdot u(\theta, (r, \mathsf{buy})) + (1-\theta) \cdot u(\theta, (r, \mathsf{not\ buy})) & \text{if } s = r; \\ H(\theta) \cdot u(\theta, (d, \mathsf{buy})) + (1 - H(\theta)) \cdot u(\theta, (d, \mathsf{not\ buy})) & \text{if } s = d \end{cases} \\ &= \begin{cases} \theta \cdot r & \text{if } s = r; \\ H(\theta) \cdot d & \text{if } s = d. \end{cases} \end{aligned}$$

Note that the payoff functions depend on the unknown state θ.

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

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

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

$$\begin{aligned}\mathbf{E}[\theta | \theta \sim \mu] &= \int_0^1 x f(x) dx = \frac{1}{2} \\ \mathbf{E}[H(\theta) | \theta \sim \mu] &= \int_0^1 H(x) f(x) dx = \frac{3}{4}.\end{aligned}$$

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.

Again, it is straightforward to see that the optimal action depends on the retailer's belief. Given the form of $H(\theta)$, one can also see that the optimal action only depends *mean* $\mathbf{E}[\bar{\theta}]$ of the retailer's belief, though this is not true in general. □

We are now ready to summarize our framework incorporating incomplete information. As mentioned at the start of this section, a decision problem is now characterized by a set of states Θ, a belief μ of the decision maker over the set of states, a set of actions $\mathcal{S}$ and a set of outcomes $\mathcal{O}$, state-dependent rules $\bar{p}_\theta$ relating the actions to outcomes, and finally a state-dependent utility function. As before, we assume that the decision maker is *rational* and evaluates distributions over the using the resulting expected utility. In addition, we make the Bayesian assumption that the decision maker evaluates uncertainty over the states by taking expectation with respect to her belief μ. Finally, this entire structure is known to the decision maker; this last assumption is just the consequence of the fact that the decision maker has represented all unknowns in the form of the state θ. A rational Bayesian decision maker then chooses the action s that maximizes her expected payoff $\Pi(s)$.

1.1 A comment on generality of the framework

In the above description of a decision problem, the rules of the setting depend on the unknown state θ. Similarly, the decision maker's preferences, as captured by her utility function, also depends on the unknown state θ. This dependence allows us to capture the decision maker's uncertainty about how actions influence the outcomes and her uncertainty about her preferences. However, neither the set of actions $\mathcal{S}$ nor the set of outcomes $\mathcal{O}$ have explicit dependence on the state θ. Why? First, we claim that the decision maker cannot have incomplete information about the set of actions $\mathcal{S}$. This is because the decision maker must always know what her available options to choose from. While the influence of these actions on the realized outcome may be unknown to the decision maker, the availability of these actions cannot be. Thus, it is not a limitation to assume that the set of actions $\mathcal{S}$ does not depend on the state θ. On the other hand, it is perfectly possible that the decision maker does not have complete information about the set of outcomes. However, we claim that such a setting can still fit into the above model. We illustrate this claim with an example.

Example. Suppose we have a simple setting with $\mathcal{S} = \{x, y\}$. The decision maker has incomplete information about the set of outcomes. Specifically, suppose the set of states is $\Theta = \{0, 1\}$ and the decision maker knows the set of outcomes is $\mathcal{O}_0 = \{o_0, o_1\}$ in state $\theta = 0$ and $\mathcal{O}_1 = \{o_2, o_3\}$ in state $\theta = 1$. Assume that the decision maker believes that each state is equally likely, i.e., $\mu(0) = \mu(1) = 0.5$. The rules relating actions to realized outcomes are specified by the functions $\bar{p}_\theta : \mathcal{S} \to \Delta(\mathcal{O}_\theta)$ for each $\theta \in \{0, 1\}$. Assume $\bar{p}_0(x) = (0.2o_0, 0.8o_1)$ and $\bar{p}_0(y) = (0.7o_0, 0.3o_1)$. Furthermore, $\bar{p}_1(x) = (0.1o_2, 0.9o_3)$ and $\bar{p}_1(y) = (0.4o_2, 0.6o_3)$. Finally, let the utility function of the decision maker in state $\theta = 0$ be given by $u(0, o_0) = 1, u(0, o_1) = 5$, and in state $\theta = 1$ be given by $u(1, o_2) = 4, u(1, o_3) = -2$. How can we represent this in the framework developed above?

To fit the above setting in our framework, we will define a new outcome set $\widehat{\mathcal{O}}$, new rules $\hat{p}_\theta : \mathcal{S} \to \Delta(\widehat{\mathcal{O}})$ and a new utility function $\hat{u} : \Theta \times \mathcal{O} \to \mathbb{R}$, which will be an equivalent representation of the decision problem.

Start by defining a new outcome set as $\widehat{\mathcal{O}} = \mathcal{O}_0 \cup \mathcal{O}_1 = \{o_0, o_1, o_2, o_3\}$. The new outcome set contains all the possible outcomes across all states. The functions $\hat{p}_\theta : \mathcal{S} \to \Delta(\widehat{\mathcal{O}})$ now represent the rules, defined as

$$\begin{aligned} \hat{p}_0(x) &= (0.2o_0, 0.8o_1, 0.0o_2, 0.0o_3) & \hat{p}_0(y) &= (0.7o_0, 0.3o_1, 0.0o_2, 0.0o_3) \\ \hat{p}_1(x) &= (0.0o_0, 0.0o_1, 0.1o_2, 0.9o_3) & \hat{p}_1(y) &= (0.0o_0, 0.0o_1, 0.4o_2, 0.6o_3). \end{aligned}$$

Finally, we define a new (and equivalent) utility function $\hat{u} : \Theta \times \widehat{\mathcal{O}} \to \mathbb{R}$ as

$$\begin{aligned} \hat{u}(0, o_0) &= 1, & \hat{u}(0, o_1) &= 5, & \hat{u}(0, o_2) &= 4^*, & \hat{u}(0, o_3) &= -2^*, \\ \hat{u}(1, o_0) &= 1^*, & \hat{u}(1, o_1) &= 5^*, & \hat{u}(1, o_2) &= 4, & \hat{u}(1, o_3) &= -2. \end{aligned}$$

(Utility values with superscript * can be replaced with any other value without any effect on decisions.) With these definitions, we see that Θ, the belief μ, $\mathcal{S}$, $\widehat{\mathcal{O}}$, $\hat{p}_\theta$, and the utility functions fit into our earlier framework. Thus, the framework we developed is flexible enough to incorporate incomplete information about the outcomes. □