Bayesian Decision Theory
Sequential Decision-Making with an Asymmetric Cost Function

Jonas Petersen

October 28, 2024

Abstract

This study explores Bayesian decision theory within a sequential decision-making framework, focusing on deriving optimal decision rules from asymmetric cost functions.

Bayesian decision theory provides a probabilistic approach for making decisions under uncertainty, allowing decision-makers to account for varying costs associated with different outcomes. By incorporating asymmetric cost functions, this approach enables a more nuanced treatment of decision costs, which is crucial in fields requiring precise risk management where different actions may lead to consequences with diverse cost implications.

This study explores Bayesian decision theory within a sequential decision-making framework, focusing on deriving optimal decision rules from asymmetric cost functions. Bayesian decision theory provides a probabilistic approach for making decisions under uncertainty, allowing decision-makers to account for varying costs associated with different outcomes. By incorporating asymmetric cost functions, this approach enables a more nuanced treatment of decision costs, which is crucial in fields requiring precise risk management where different actions may lead to consequences with diverse cost implications.

To illustrate these concepts, consider a specific setting in which decision-making unfolds as an interactive game between two entities: Nature, which represents the randomness in the environment, and a Robot, which operates as a decision-maker tasked with managing a finite stock of units over time. At each time iteration t, the Robot makes decisions based on its action space Ω_U, while Nature selects outcomes from its action space Ω_S. For example, if the Robot begins with an initial stock of N_0 units, each time step involves a certain number of units being removed by Nature ($s_t \in \Omega_S$), while the Robot has the option to reorder stock, subject to an unspecified lead time.

This stock management scenario serves as an example of sequential decision-making under uncertainty, where costs are associated with both overstocking (when $N > 0$) and understocking (when $N < 0$), which influence the Robot's decisions in each iteration. To support its decision-making, the Robot is provided with historical data D, containing records of past unit removals and relevant features for demand forecasting. The stock level at time t is given by

$$N_t \equiv N_0 + \sum_{t'=1}^{t}(U_{t'-L} - s_{t'}), \tag{1}$$

and the Robot's decisions are informed by a probabilistic forecast based on D in the form

$$p(s_1, s_2, \ldots | D, I), \tag{2}$$

where I denotes any additional background information (Sivia and Skilling, 2006). In order to construct the asymmetric cost functions considered in this study, the swish function

$$\text{swish}(N_t, \beta) = \frac{N_t}{1 + e^{-N_t \beta}}, \tag{3}$$

with $\beta \gg 1$, is used as a continuous imitation of Relu, meaning

$$\text{swish}(\pm N_t, \beta) \simeq N_t^{(\pm)}, \tag{4}$$

such that

$$N_t^{(\pm)} \subset C(u, s). \tag{5}$$

Given the observations D, the Robot's objective is to formulate a set of decision rules, $\xi \equiv \{U_0(D), U_1(D), \ldots\}$ with $U_j(D) = u_j$, that minimize the expected cost associated with its decisions

$$\xi^* \equiv \underset{\xi}{\text{argmin}}(\mathbb{E}[C|D, I]), \tag{6}$$

where C is the cost infinitely far into the future. In this study, the cost will be assumed to be discounted, meaning

$$C = \sum_{t=1}^{\infty} \gamma^t C_t \tag{7}$$

where $C_t \equiv C(U_t(D), s_t)$ and $\gamma \in [0, 1]$. The optimal decisions are then defined by

$$\frac{d}{dU_m}\mathbb{E}[C|D, I]\Big|_{U_m = U_m^*} \overset{!}{=} 0 \quad \forall m. \tag{8}$$

2

Example 1. ————————————————————————————————

The cost function

$$C = \sum_{t=1}^{\infty} \gamma^t (k_t^{\mathrm{sc}} N_t^{(+)} + k_t^{\mathrm{uv}} N_t^{(-)})$$

(9)

where k^{sc} and k^{uv} represent the "storage cost" and the "unit value" (lost value if under stocking), leads to the decision rules (see appendix A)

$$p(N_t^* \geq 0 | D, I) = \frac{k_t^{\mathrm{uv}}}{k_t^{\mathrm{uv}} + k_t^{\mathrm{sc}}}$$

(10)

with

$$N_t^* \equiv N_0 + \sum_{t'=1}^{t} (U_{t'-L}^* - s_{t'}).$$

(11)

Example 2. ————————————————————————————————

The cost function

$$C = \sum_{t=1}^{\infty} \gamma^t (k_t^{\mathrm{sc}} N_t^{(+)} + k_t^{\mathrm{uv}} s_t N_t^{(-)})$$

(12)

yielding the decision rule (see appendix B)

$$\frac{p(N_t \geq 0 | D, I)}{\mathbb{E}[s_t | D, I] - \mathbb{E}[s_t | N_t \geq 0, D, I]} = \frac{k_t^{\mathrm{uv}}}{k_t^{\mathrm{sc}}}.$$

(13)

A   Minimization of Expected Cost: First Cost Function

Define the cost function viz

$$C = \sum_{t=1}^{\infty} \gamma^t (k_t^{\text{sc}} N_t^{(+)} + k_t^{\text{uv}} N_t^{(-)}),\tag{14}$$

where k^{sc} and k^{uv} represent the "storage cost" and the "unit value". To determine the decision rules, ξ^*, that minimize the expected cost $\mathbb{E}[C|D, I]$, the derivative of the cost function with respect to ξ is needed

$$\frac{dC}{dU_m} = \sum_{t=1}^{\infty} \gamma^t \left(k_t^{\text{sc}} \frac{dN_t^{(+)}}{dU_m} + k_t^{\text{uv}} \frac{dN_t^{(-)}}{dU_m} \right)\tag{15}$$

where

$$
\begin{aligned}
\frac{dN_t^{(+)}}{dU_m} &= \sum_q \frac{dN_t^{(+)}}{dN_q} \frac{dN_q}{dU_m} \\
&\simeq \sum_q 1_{N_t \geq 0} \delta_{t,q} \frac{dN_q}{dU_m} \\
&= 1_{N_t \geq 0} \sum_{t'=1}^{t} \frac{dU_{t'-L}}{dU_m} \\
&= 1_{N_t \geq 0} \sum_{t'=1}^{t} \delta_{t'-L,m}, \\
\frac{dN_t^{(-)}}{dU_m} &\simeq (1_{N_t \geq 0} - 1) \sum_{t'=1}^{t} \delta_{t'-L,m},
\end{aligned}\tag{16}
$$

and it has been used that

$$\frac{dN_t^{(+)}}{dN_t} = \frac{\beta N_t e^{-\beta N_t}}{(1+e^{-\beta N_t})^2} + \frac{1}{1+e^{-\beta N_t}}$$

$$\simeq \frac{1}{1+e^{-\beta N_t}}$$

$$= \sigma(\beta N_t)$$

$$\simeq 1_{N_t \geq 0},$$

$$\frac{dN_t^{(-)}}{dN_t} = \frac{\beta N_t e^{\beta N_t}}{(1+e^{\beta N_t})^2} - \frac{1}{1+e^{\beta N_t}}$$

$$\simeq -\frac{1}{1+e^{\beta N_t}}$$

$$= \sigma(\beta N_t) - 1$$

$$\simeq 1_{N_t \geq 0} - 1. \tag{17}$$

Using equation (16) in equation (15)

$$\frac{dC}{dU_m} \simeq \sum_{t=1}^{\infty} \gamma^t \left(k_t^{\mathrm{sc}} 1_{N_t \geq 0} + k_t^{\mathrm{uv}}(1_{N_t \geq 0} - 1) \right) \sum_{t'=1}^{t} \delta_{t'-L,m}. \tag{18}$$

For some generic function g_t

$$\sum_{t=1}^{\infty} g_t \sum_{t'=1}^{t} \delta_{t'-L,m} = g_1 \delta_{1-L,m} + g_2(\delta_{1-L,m} + \delta_{2-L,m}) + \dots$$

$$= \sum_{t=L+m}^{\infty} g_t \tag{19}$$

meaning

$$\frac{dC}{dU_m} \simeq \sum_{t=L+m}^{\infty} \gamma^t (k_t^{\mathrm{sc}} 1_{N_t \geq 0} + k_t^{\mathrm{uv}}(1_{N_t \geq 0} - 1)). \tag{20}$$

Combining equations (8) and (20)

$$\sum_{s_1,s_2,\dots} \sum_{t=L+m}^{\infty} \gamma^t (k_t^{\mathrm{sc}} 1_{N_t \geq 0} + k_t^{\mathrm{uv}}(1_{N_t \geq 0} - 1)) p(s_1, s_2, \dots | D, I) \stackrel{!}{=} 0 \quad \forall m \tag{21}$$

The sums can be evaluated viz

$$\sum_{s_1,s_2,\dots} 1_{N_t \geq 0} p(s_1, s_2, \dots | D, I) = p\left(\sum_{t'=1}^{t} s_{t'} \leq N_0 + \sum_{t'=1}^{t} U_{t'-L} | D, I \right)$$

$$= p(N_t \geq 0 | D, I), \tag{22}$$

$$\sum_{s_1,s_2,\dots} p(s_1, s_2, \dots | D, I) = 1.$$

Let

$$\psi_t \equiv (k_t^{\text{sc}} + k_t^{\text{uv}})p(N_t \geq 0|D, I) - k_t^{\text{uv}}, \tag{23}$$

then

$$\frac{d}{dU_m}\mathbb{E}[C|D, I] = \sum_{t=L+m}^{\infty} \gamma^t \psi_t$$
$$\overset{!}{=} 0 \quad \forall m \tag{24}$$

A recursion relation can be derived viz

$$\frac{d}{dU_0}\mathbb{E}[C|D, I] = \sum_{t=L}^{\infty} \gamma^t \psi_t$$
$$= \gamma^L \psi_L + \sum_{t=L+1}^{\infty} \gamma^t \psi_t$$
$$= \gamma^L \psi_L + \frac{d}{dU_1}\mathbb{E}[C|D, I] \tag{25}$$
$$= \gamma^L \psi_L + \gamma^{L+1}\psi_{L+1} + \frac{d}{dU_2}\mathbb{E}[C|D, I]$$
$$= \ldots$$
$$\overset{!}{=} 0$$

Since all derivatives are required to be simultaneously zero,

$$\gamma^j \psi_j \overset{!}{=} 0 \quad \forall j \Rightarrow \psi_j = 0 \tag{26}$$

meaning

$$p(N_t^* \geq 0|D, I) = \frac{k_t^{\text{uv}}}{k_t^{\text{sc}} + k_t^{\text{uv}}}, \tag{27}$$

where

$$N_t^* \equiv N_0 + \sum_{t'=1}^{t}(U_{t'-L}^* - s_{t'}) \tag{28}$$

denote the units on stock given optimal decisions.

B   Minimization of Expected Cost: Second Cost Function

With the cost function

$$C = \sum_{t=1}^{\infty} \gamma^t (k_t^{\text{sc}} N_t^{(+)} + k_t^{\text{uv}} N_t^{(-)}) \tag{29}$$

the derivation proceeds analogously to that of appendix A, resulting again in the criterion $\psi_j = 0$, however, now with

$$\psi_t \equiv k_t^{\text{sc}} p(N_t \geq 0|D, I) + k_t^{\text{uv}}(\mathbb{E}[s_t|N_t \geq 0, D, I] - \mathbb{E}[s_t|D, I]), \qquad (30)$$

and

$$\sum_{s_1, s_2, \ldots} s_t 1_{N_t \geq 0} p(s_1, s_2, \ldots |D, I) = \mathbb{E}[s_t|N_t \geq 0, D, I],$$

$$\sum_{s_1, s_2, \ldots} s_t p(s_1, s_2, \ldots |D, I) = \mathbb{E}[s_t|D, I]. \qquad (31)$$

The resulting optimal decision rules are

$$\frac{p(N_t^* \geq 0|D, I)}{\mathbb{E}[s_t|D, I] - \mathbb{E}[s_t|N_t^* \geq 0, D, I]} = \frac{k_t^{\text{uv}}}{k_t^{\text{sc}}}. \qquad (32)$$

References

D. S. Sivia and J. Skilling. *Data Analysis - A Bayesian Tutorial.* Oxford Science Publications. Oxford University Press, 2nd edition, 2006.