BAYESIAN DECISION THEORY AND MONTE CARLO METHODS

ARNE BANG HUSEBY

Abstract. In this note we show how Monte Carlo methods can be used efficiently to find Bayes decision rules when the set of possible actions consists of only two possible actions. The method is demonstrated on an offshore development project. The Monte Carlo methods can also be used in sequential decision problems. As an example we consider the problem of pricing American options. An American option is an option which can be exercised at any point of time within a given interval $[0, \tau]$. Before demonstrating the Monte Carlo methods on this problem, we study under which conditions one can gain anything by exercising the option before the expiration time τ.

1. Introduction

Bayesian decision theory provides a natural framework for solving the fundamental problems of option pricing. To see this we start out by considering the following standard setup. Let θ denote some *uncertain parameter* which is known to belong to some set Θ, and let $\pi(\theta)$ denote the prior distribution of θ. Information about θ is available in terms of *data* denoted by X with a conditional distribution given θ denoted by $f(x|\theta)$ defined for all $x \in \mathcal{X}$. Based on the data X we want to make a *decision*, that is we want to select an action a from some space of possible actions $\mathcal{A}$. The consequence of this decision depends on both the chosen action a and the uncertain variable θ, and is described in terms of a *utility function* $U = U(\theta, a)$, a real valued function defined for all $a \in \mathcal{A}$ and $\theta \in \Theta$. A *decision rule*, $\delta = \delta(x)$, is a function defined from $\mathcal{X}$ into $\mathcal{A}$, representing the decision made if the data $X = x$ is observed.

The expected utility (prior to observing X) given that a decision rule δ is used, can be written as:

$$E[U; \delta] = \int_{\Theta} [\int_{\mathcal{X}} U(\theta, \delta(x)) f(x|\theta) dx] \pi(\theta) d\theta. \tag{1.1}$$

The inner integral of (1.1) is often referred to as the *risk function* of the chosen decision rule, denoted by $R(\theta; \delta)$:

$$R(\theta; \delta) = \int_{\mathcal{X}} U(\theta, \delta(x)) f(x|\theta) dx. \tag{1.2}$$

In non-bayesian decision theory decision rules are compared by considering their respective risk functions. This is difficult as one in general cannot find a *uniformly best* decision rule, i.e., a rule, δ^* such that $R(\theta; \delta^*) \geq R(\theta; \delta)$ for all $\theta \in \Theta$ and for all other decision rule δ. In bayesian decision theory, however, the decision rules are compared by considering their respective expected utilities. A decision rule, δ^B is said to be a *Bayes rule* if $E[U; \delta^B] \geq E[U; \delta]$ for all other decision rule δ. In many cases this leads to an essentially *unique* solution of the decision problem.

In order to find the Bayes rule, one usually rewrites (1.1) using Fubini's theorem as:

$$E[U;\delta] = \int_{\mathcal{X}} [\int_{\Theta} U(\theta, \delta(x))\pi(\theta|x)d\theta]f(x)dx, \tag{1.3}$$

where $\pi(\theta|x)$ denotes the posterior distribution of θ given $X = x$, and $f(x)$ denotes the unconditional distribution of X. The inner integral of (1.3) is referred to as the *posterior expected utility* given that $X = x$ and that the decision rule δ is used:

$$E[U|X = x;\delta] = \int_{\Theta} U(\theta, \delta(x))\pi(\theta|x)d\theta. \tag{1.4}$$

The Bayes rule, δ^B, is then found by letting $\delta^B(x)$ be the action that for each x maximizes the posterior expected utility given that $X = x$.

2. Decision Problems with Binary Action Space

In this paper we will focus on decision problems where the action space $\mathcal{A}$ consists of just to possible actions, i.e., $\mathcal{A} = \{a_0, a_1\}$. In this case the optimal decision for a given x is found by comparing:

$$\begin{aligned} E[U|X = x; a_0] &= \int_{\Theta} U(\theta, a_0)\pi(\theta|x)d\theta, \\ E[U|X = x; a_1] &= \int_{\Theta} U(\theta, a_1)\pi(\theta|x)d\theta, \end{aligned} \tag{2.1}$$

and choosing the action with the highest expected utility. Thus, the Bayes rule, δ^B takes the following form:

$$\delta^B(x) = \begin{cases} a_0 & \text{if } E[U|X = x; a_0] \geq E[U|X = x; a_1] \\ a_1 & \text{if } E[U|X = x; a_0] < E[U|X = x; a_1] \end{cases}. \tag{2.2}$$

If the expected utilities needed in (2.2) can be calculated analytically, it is easy to find an explicit form of the Bayes rule. In other cases, however, one may need to estimate the expectations by using Monte Carlo simulation. For a given value of x, this involves sampling from the posterior distribution of θ given x. Sometimes, however, one may wish to derive the complete decision rule. That is, one may want to determine $\delta = \delta(x)$ by using Monte Carlo simulation. More specifically, since the action space, $\mathcal{A}$ consists of only two elements, one wants to determine the following sets:

$$\begin{aligned} \mathcal{X}_0 &= \{x \in \mathcal{X} : E[U|X = x; a_0] \geq E[U|X = x; a_1]\}, \\ \mathcal{X}_1 &= \{x \in \mathcal{X} : E[U|X = x; a_0] < E[U|X = x; a_1]\}. \end{aligned} \tag{2.3}$$

In general the data X may be a vector of observations. In such cases it may be difficult to determine the sets $\mathcal{X}_0$ and $\mathcal{X}_1$ by using Monte Carlo simulation. Often, however, it is possible to reduce the dimension of the data by using the concept of *sufficiency*. If $S = S(X)$ is some sufficient statistic, then by definition the posterior distribution of θ given X depends on X only through S. Thus, we may express the decision rule as $\delta = \delta(S(X))$, or simply $\delta = \delta(S)$. Since the dimension of S can be significantly smaller than the dimension of X, this leads to a much simpler problem. In the following we avoid introducing a new quantity S by assuming that X is itself a sufficient statistic with the smallest possible dimension.

We then introduce the following two functions:

$$\begin{aligned} \psi_0(x) &= E[U|X=x;a_0], \\ \psi_1(x) &= E[U|X=x;a_1]. \end{aligned} \tag{2.4}$$

In order to determine the sets $\mathcal{X}_0$ and $\mathcal{X}_1$, we would like to estimate ψ_0 and ψ_1 by using Monte Carlo simulation. The naive approach to this would be to choose a suitable large but finite set of x-values, say $\mathcal{X}_{MC} = \{x_1, \ldots, x_n\} \subseteq \mathcal{X}$. Then for each $x_i \in \mathcal{X}_{MC}$ we estimate $\psi_0(x_i)$ and $\psi_1(x_i)$ by sampling from the posterior distribution of θ given $X = x_i$. From this we obtain the following partition of the set $\mathcal{X}_{MC}$:

$$\begin{aligned} \mathcal{X}_{MC,0} &= \{x_i : \hat{\psi}_0(x_i) \geq \hat{\psi}_1(x_i)\}, \\ \mathcal{X}_{MC,1} &= \{x_i : \hat{\psi}_0(x_i) < \hat{\psi}_1(x_i)\}, \end{aligned} \tag{2.5}$$

where $\hat{\psi}_0$ and $\hat{\psi}_1$ are the Monte Carlo estimates of ψ_0 and ψ_1. From the sets $\mathcal{X}_{MC,0}$ and $\mathcal{X}_{MC,1}$ we try to get estimates for the "true" sets $\mathcal{X}_0$ and $\mathcal{X}_1$, e.g., by considering convex closures or similar techniques. The problem with this approach, however, is that due to sampling errors we may not get a nice and sharp border between the sets $\mathcal{X}_{MC,0}$ and $\mathcal{X}_{MC,1}$. That is, for x_i-values close to the border between $\mathcal{X}_0$ and $\mathcal{X}_1$, where $\psi_0(x_i)$ and $\psi_1(x_i)$ typically are close to each other, which one of $\hat{\psi}_0$ and $\hat{\psi}_1$ which is the larger one, may fluctuate a lot.

A better approach, yielding more stable results, is to use what is known as the *common random number* (or CRN) method. (See e.g., Glasserman and Yao [1].) To explain this method, we start out by considering the problem of sampling θ from the posterior distribution $\pi(\theta|x)$. Generally this is done by generating a suitable number of uniformly distributed random variables (using some sort of random number generator), and then transform these variables to variables having the correct distribution. Note that depending on the distribution and the algorithm being used we may need more than one uniformly distributed variable per θ. Assume e.g., that we want to generate $\theta_1, \ldots, \theta_N$ from $\pi(\theta|x)$, and let R_i denote the vector of uniformly distributed random numbers needed to produce θ_i, $i = 1, \ldots, N$. The values $\theta_1, \ldots, \theta_N$ are then obtained by applying the appropriate transformation, say:

$$\theta_i = G(R_i, x),\ i = 1, \ldots, N. \tag{2.6}$$

Note that since θ depends on X the value of X is typically included in the transformation as indicated above. In general the number of uniformly distributed random numbers needed, i.e., the dimension of R_i, may depend on the value of X. This may happen if e.g., a rejection sampling method is used, where the rejection criterion may depend on the value of X. In this setting, however, we ignore this minor technical difficulty.

Having generated $\theta_1, \ldots, \theta_N$, ψ_0 and ψ_1 are estimated by:

$$\begin{aligned} \hat{\psi}_0(x) &= \frac{1}{N}\sum_{i=1}^{N} U(\theta_i, a_0) = \frac{1}{N}\sum_{i=1}^{N} U(G(R_i, x), a_0) \\ \hat{\psi}_1(x) &= \frac{1}{N}\sum_{i=1}^{N} U(\theta_i, a_1) = \frac{1}{N}\sum_{i=1}^{N} U(G(R_i, x), a_1) \end{aligned} \tag{2.7}$$

Now, the crucial idea is to use the *same random numbers* $R_1, \ldots, R_N$ for *all* values of X. Thus, $\hat{\psi}_0$ and $\hat{\psi}_1$ can be evaluated for any value of X by entering this value into (2.7). The border between the sets $\mathcal{X}_0$ and $\mathcal{X}_1$ can then be estimated by finding the roots of the equation:

$$\sum_{i=1}^{N} U(G(R_i, x), a_0) = \sum_{i=1}^{N} U(G(R_i, x), a_1). \tag{2.8}$$

If N is large, the calculations can be slow. However, if the dimension of X is small, it is possible to find satisfactory results.

We close this section by considering a simple example where we use the above simulation method. Assume that we are considering an offshore field development project. One of the main uncertainties in the project is the recoverable oil volume. We denote this volume by θ. Before we make the final development decision, we have the option of drilling a test well. The relevant data from the test well is summarized in a sufficient one-dimensional statistic X. The posterior distribution of θ (in mill.Sm^3) given X is a lognormal distribution with mean value X and standard deviation 2.5.

Given that the results of the test well are satisfactory, the field is developed. The present value of the total cost of a full field development (in mill. NOK), including operational costs and the cost of the test well, is denoted C_1. We assume that $E[C_1] = 2900$. The present value of the resulting sales (in mill.NOK) is denoted S. This quantity can be a very complex function involving θ as well as many other uncertain quantities. Thus, to estimate $E[S]$, it is typically necessary to use Monte Carlo simulation. In order to demonstrate the ideas, however, we use the following very simple model:

$$S = \theta \cdot D, \tag{2.9}$$

where $E[D] = 350$.

If the results of the test well are unsatisfactory, all further development is stopped. In this case the only cost is the cost of the test well (in mill. NOK) which we denote C_0. We assume that $E[C_0] = 250$.

The two possible actions after the value of X is observed, are "*No further development*", and "*Full development*", denoted by a_0 and a_1 respectively. For simplicity we assume a linear utility function, denoted by U. If a_0 is chosen, $U = -C_0$, while is a_1 is chosen, $U = S - C_1$.

With these assumptions it is easy to see that:

$$\begin{aligned} \psi_0(x) &= E[U|X = x; a_0] = -250, \\ \psi_1(x) &= E[U|X = x; a_1] = 350 \cdot x - 2900. \end{aligned} \tag{2.10}$$

From this it follows that the Bayes rule δ^B is given by:

$$\delta^B(x) = \begin{cases} a_0 & \text{if } x \leq 7.57 \\ a_1 & \text{if } x > 7.57 \end{cases}. \tag{2.11}$$

The same result can also be found by using Monte Carlo simulation to estimate $\psi_0(x)$ and $\psi_1(x)$. In Figure 1 we have plotted the estimated functions $\hat{\psi}_0(x)$ (the curve marked with triangles) and $\hat{\psi}_1(x)$ (the curve marked with squares). In the simulations we used x-values $5, 6, \ldots, 10$. By interpolating between these values, we find that the two curves intersect when x is approximately 7.5 which is close to the optimal value. When x is above this value, $\hat{\psi}_1(x) > \hat{\psi}_0(x)$, so for such x-values a_1

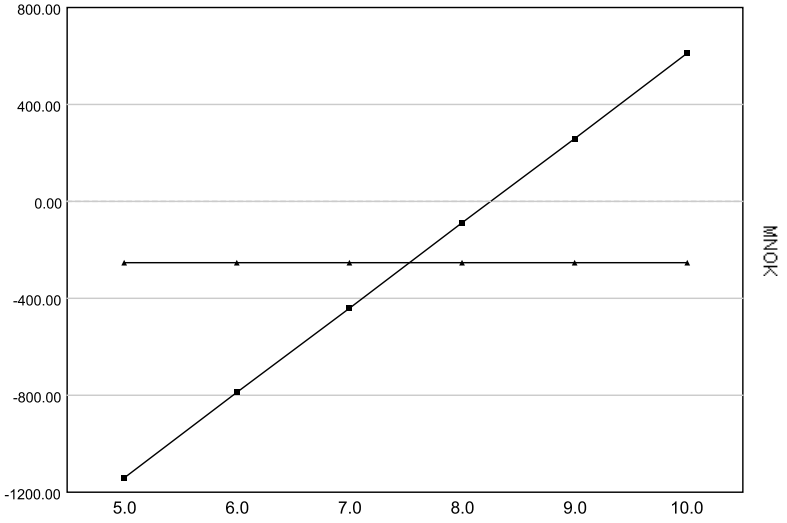

FIGURE 1. *Estimated expected utilities as functions of x.*

should be chosen. On the other hand, when x is below this value, $\hat{\psi}_1(x) < \hat{\psi}_0(x)$, implying that a_0 is the best choice.

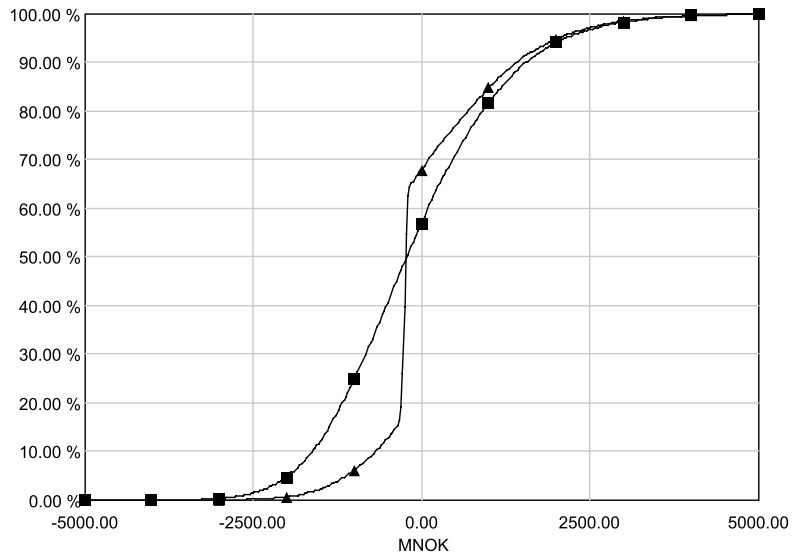

FIGURE 2. *Estimated Cumulative Utility Distributions.*

In Figure 2 we have plotted the estimated the cumulative distribution functions for the project utility given that the optimal decision rule is applied (the curve marked with triangles) versus the corresponding curve if a_1 is chosen regardless of the value of x (the curve marked with squares). By comparing these two curves we see how optimal use of the data from the test well improves the situation.

We observe that by using the optimal decision rule, the potential downside of the project is reduced considerably. The reason for this is that by utilizing the test well results we are able to stop the project when the results are unsatisfactory, and thus avoid disasters. Note, however, that by using the optimal rule, we also lose some of the potential upside, since we sometimes stop the project when we should have proceeded. Still this effect is weaker than the reduction in potential downside, so the expected utility is higher when using the optimal rule.

3. Sequential Decision Problems

One of the nice properties with the Monte Carlo method is that it can be extended to sequential decision problems. Such problems are typically very difficult to solve analytically. In order to illustrate the sequential method we consider an asset value process $S(t)$ and an American option whose pay-off function, h, depends on this process. The option can be exercised at any point of time within a given interval $[0, \tau]$. If the option is exercised at time t, the cash value of the pay-off is $h(S(t))$.

In order to avoid arbitrage we assume that all expectations are calculated with respect to a given risk neutral probability measure. This implies that:

$$E[\exp^{-r(t_2-t_1)} S(t_2)|S(u), 0 \leq u \leq t_1] = S(t_1), \tag{3.1}$$

for all $0 \leq t_1 \leq t_2 \leq \tau$ where r denotes the risk free interest rate.

We assume that we have purchased one such option at time 0, and then consider a point of time t such that $t < \tau$. At this point of time we may choose to exercise the option or wait. If we exercise the option, the value of the option will have a present value relative to time t, denoted $V_0(t)$ which is equal to the cash value of the pay-off. That is,

$$V_0(t) = h(S(t)). \tag{3.2}$$

If on the other hand, we wait until time τ, the expected present value of the option relative to time t, denoted $V_1(t)$ will be:

$$V_1(t) = E[\exp^{-r(\tau-t)} h(S(\tau))|S(u), 0 \leq u \leq t]. \tag{3.3}$$

Obviously, we would exercise the option at time t if and only if $V_0(t) > V_1(t)$. If there is a positive probability that $V_0(t) > V_1(t)$ at some point of time t, then the value of such an option would be greater than a European option with fixed exercise time τ. On the other hand if $V_0(t) \leq V_1(t)$ for all t almost surely, then the value of the American option is equal to the corresponding European option.

The optimal decision at time t depends on the shape of the pay-off function. In the following we assume that the pay-off function h is convex. Under this assumption we can use the well-known Jensen's inequality, and obtain the following:

$$\begin{aligned} V_1(t) &= E[\exp^{-r(\tau-t)} h(S(\tau))|S(u), 0 \leq u \leq t] \\ &\geq \exp^{-r(\tau-t)} h(E[S(\tau)|S(u), 0 \leq u \leq t]) \end{aligned} \tag{3.4}$$

Since h is convex we also get that:

$$(1-\alpha)h(0) + \alpha h(x) \geq h((1-\alpha)0 + \alpha x) = h(\alpha x), \tag{3.5}$$

for all $0 \leq \alpha \leq 1$. Hence, by replacing α by $\exp^{-r(\tau-t)}$, and inserting (3.5) into (3.4), we get that:

$$\begin{aligned} V_1(t) &\geq h(\exp^{-r(\tau-t)} E[S(\tau)|S(u), 0 \leq u \leq t]) - (1-\alpha)h(0) \\ &= h(E[\exp^{-r(\tau-t)} S(\tau)|S(u), 0 \leq u \leq t]) - (1-\alpha)h(0) \\ &= h(S(t)) - (1-\alpha)h(0) \\ &= V_0(t) - (1-\alpha)h(0), \end{aligned} \tag{3.6}$$

where we also have used that the probability measure is assumed to be risk neutral, i.e., (3.1).

From (3.6) the following result is immediate:

Theorem 3.1. *Assume that the pay-off function, h, of an American option is convex and that $h(0) = 0$. Then the option is equivalent to an European option. That is, it is always optimal to exercise the option at the expiration time.*

As an example we consider an American call option with pay-off function:

$$h(s) = (s-K)^+ = \max(s-K, 0), \tag{3.7}$$

where $K > 0$ denotes the strike price. It is easy to see that h is convex and that $h(0) = 0$. Thus, it follows by Theorem 3.1 that this option is equivalent to a European option. That is, one should never exercise an American call option before its expiration time. This result is actually very well-known. See e.g., Ross [3].

Considering (3.4) once again we see that if the risk free interest rate r is zero, the lower bound on $V_1(t)$ can be written as:

$$V_1(t) \geq h(E[S(\tau)|S(u), 0 \leq u \leq t]). \tag{3.8}$$

Moreover, in this case the risk neutral probability measure satisfies:

$$E[S(t_2)|S(u), 0 \leq u \leq t_1] = S(t_1), \tag{3.9}$$

Hence, by combining (3.8) and (3.9) we get that:

$$V_1(t) \geq h(S(t)) = V_0(t). \tag{3.10}$$

Hence we have shown the following result:

Theorem 3.2. *Assume that the pay-off function, h, of an American option is convex and that the risk free interest rate is zero. Then the option is equivalent to an European option. That is, it is always optimal to exercise the option at the expiration time.*

As an example we consider an American put option with pay-off function:

$$h(s) = (K-s)^+ = \max(K-s, 0), \tag{3.11}$$

where $K > 0$ denotes the strike price. Again it is easy to see that h is convex. Thus, if the risk free interest rate is zero, it follows by Theorem 3.2 that this option is equivalent to a European option. That is, one should never exercise an American put option before its expiration time when the risk free interest rate is zero.

Note, however, that if the risk free interest rate is positive, it may sometimes be advantageous to exercise an American put option before its expiration time. In

Figure 3 we have plotted expected pay-off as functions of current asset price. The strike price is 5, while the risk free interest rate is 4%. The curve marked by squares represents the pay-off if the option is exercised at time $\tau-1$, while the curve marked by triangles represents the pay-off if the option is exercised at its expiration time τ. We see that the two curves intersect when the current asset price is approximately 4.3. If the current price is below this number, it is better to exercise the option at time $\tau - 1$, while if the current price is above this number, it is better to exercise the option at its expiration time.

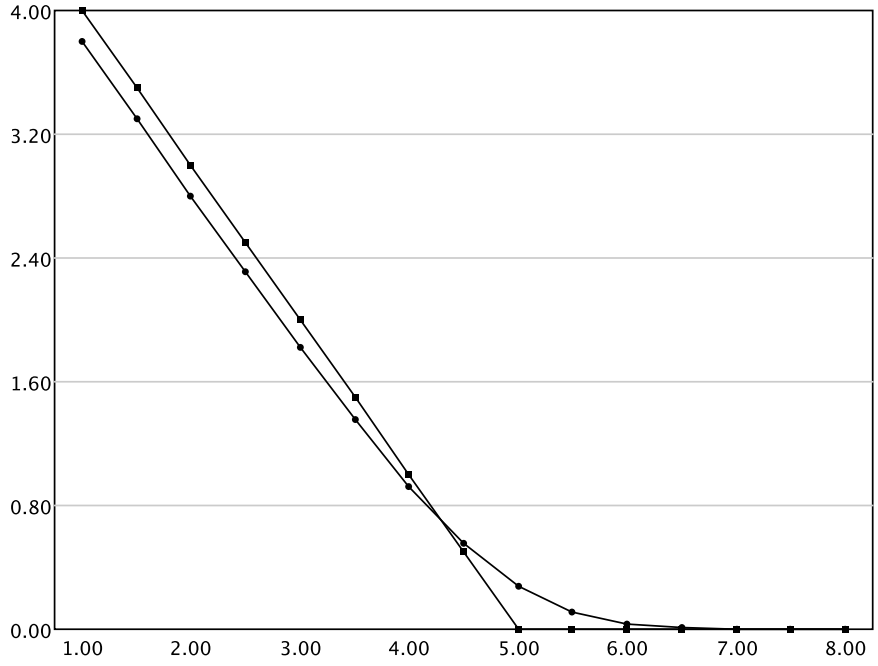

FIGURE 3. *Estimated mean pay-off as functions of current asset price.*

REFERENCES

1. P. Glasserman and D. D. Yao, *Monotone Structure in Discrete-Event Systems* Wiley, (1994).
2. P. Glasserman, *Monte Carlo Methods in Financial Engineering* Springer, (2004).
3. S. M. Ross, *An Elementary Introduction to Mathematical Finance* Cambridge University Press, (2003).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO
E-mail address: arne@math.uio.no