EE365 Homework 1

1. *Optimal disposition of a stock.* You must sell a total amount $B > 0$ of a stock in two rounds. In each round you can sell any nonnegative amount of the stock; by the second round all of the initial stock amount B must be sold. The (positive) prices in the two rounds are p_0 and p_1, respectively. These are independent log-normal variables:

$$\log p_0 \sim \mathcal{N}(\mu_0, \sigma_0^2), \quad \log p_1 \sim \mathcal{N}(\mu_1, \sigma_1^2).$$

The goal is to maximize the total expected revenue from the sales in the two rounds.

We consider three different information patterns.

- *Prescient.* You know p_0 and p_1 before you decide the amounts to sell in each period.
- *No knowledge.* You do not know the prices.
- *Partial knowledge.* You are told the price p_0 before you decide how much to sell in period 0, and you are told the price p_1 before you decide how much to sell in period 1.

(a) Find the optimal policies for each of the three different information patterns. The amount sold in each period can depend on the problem data $(B, \mu_0, \mu_1, \sigma_0, \sigma_1)$ and of course the additional information available, which depends on the information pattern.

(b) *Numerical example.* Consider the specific case with

$$B = 10, \quad \mu_0 = 0, \quad \mu_1 = 0.1, \quad \sigma_0 = \sigma_1 = 0.4.$$

Plot the distribution of total revenue for the stochastic control problems for the three different information patterns, using Monte Carlo. Give the expected values of total revenue in each case (again, computed by Monte Carlo).

Hints.

- If $\log x \sim \mathcal{N}(\mu, \sigma^2)$, we have $\mathbf{E}\, x = \exp(\mu + \sigma^2/2)$.
- In matlab you can plot the histogram of a vector `v` in `n` bins using the command `hist(v,n)`.

You don't need to know a general method for solving stochastic control problems to solve this problem. You can solve it directly using basic and simple arguments.

2. *Optimal strategy in a series of chess games.* Alice and Bob have decided to play a series of chess games. They initially play two games in which a player receives two points for each win, one point for each draw, and no points for each loss. If one player has more points after the first two games, then that player is declared the winner of the series; if

the players have the same number of points after the first two games, then the series enters sudden death, and whoever is the next to win a game is declared the winner of the series.

Alice has two distinct styles of play: aggressive and defensive. If Alice plays aggressively, she wins with probability $w_A = 0.5$, loses with probability $l_A = 0.4$, and draws with probability $d_A = 1-(w_A+l_A) = 0.1$; if she plays defensively, she wins, loses and draws with probabilities $w_D = 0.1$, $l_D = 0.2$ and $d_D = 1-(w_D+l_D) = 0.7$, respectively.

(a) Argue that Alice should always play aggressively if the series enter sudden death. What is the conditional probability that Alice wins the series given the series enters sudden death, and she always plays aggressively?

(b) Suppose Alice decides her strategy for each game of the series before the first game. In other words, she uses an *open-loop* or *static* policy.

i. If Alice plays aggressively in all of the games, what is the probability that she wins the series?

ii. If Alice plays defensively in the first two games, and aggressively in any sudden-death games, what is the probability that she wins the series?

iii. If Alice plays defensively in the first game, aggressively in the second game, and aggressively in any sudden-death games, what is the probability that she wins the series?

iv. If Alice plays aggressively in the first game, defensively in the second game, and aggressively in any sudden-death games, what is the probability that she wins the series?

Which of these four strategies should Alice use?

(c) Now suppose Alice is willing to change her strategy as the series progresses, *i.e.*, she uses *recourse*, making her decisions based on he current situation.

i. Suppose Alice wins the first game. What is the probability that she wins the series if she plays aggressively in the second game and any sudden-death games? What is the probability that she wins the series if she plays defensively in the second game, and aggressively in any sudden-death games? How should Alice play in the second game if she wins the first game?

ii. Suppose Alice loses the first game. What is the probability that she wins the series if she plays aggressively in the second game and any sudden-death games? What is the probability that she wins the series if she plays defensively in the second game, and aggressively in any sudden-death games? How should Alice play in the second game if she loses the first game?

iii. Suppose the first game is a draw. What is the probability that Alice wins the series if she plays aggressively in the second game and any sudden-death games? What is the probability that she wins the series if she plays defensively in the second game, and aggressively in any sudden-death games? How should Alice play in the second game if the first game is a draw?

iv. How should Alice play in the first game? What is the probability that she wins the series if she uses the optimal recourse policy? How does this compare to the probability that she wins the series if she uses the optimal static policy?

3. *Counting the number of policies.* Consider a stochastic optimal control problem:

$$\text{minimize} \quad J = \mathbf{E}\left(\sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T, w_T)\right).$$

We assume that each state x_t is an element of a finite set $\mathcal{X}$, each input u_t is an element of a finite set $\mathcal{U}$, and each disturbance is an element of a finite set $\mathcal{W}$. For each of the following questions, give an answer in terms of the cardinalities of these sets.

(a) An open-loop policy is just a sequence of inputs: $(u_0, \ldots, u_{T-1})$. How many open-loop policies are there?

(b) A prescient policy is a map from the set of sequences $(w_0, \ldots, w_{T-1})$ to the set of input sequences $(u_0, \ldots, u_{T-1})$. How many prescient policies are there?

(c) A closed-loop policy where we do not observe w_t before choosing u_t is a sequence of functions $(\phi_0, \ldots, \phi_{T-1})$, where $\phi_t : \mathcal{X} \to \mathcal{U}$. How many such closed-loop policies are there?

(d) A closed-loop policy where we observe w_t before choosing u_t is a sequence of functions $(\phi_0, \ldots, \phi_{T-1})$, where $\phi_t : \mathcal{X} \times \mathcal{W} \to \mathcal{U}$. How many such closed-loop policies are there?

4. *Brute-force approach to a simple stochastic optimal control problem.* Consider a simple stochastic optimal control problem with state space $\mathcal{X} = \{1, 2\}$, action space $\mathcal{U} = \{1, 2\}$, $\mathcal{W} = \{1, 2\}$, state-transition function

$$f_t(x, u, w) = u,$$

(so that $x_{t+1} = u_t$), and cost functions:

$$g_t(x, u, w) = \begin{cases} -a & x = u = 1, w = 1, \\ (t+1)^2 & x = 1, u = 2, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \quad g_T(x, w) = \begin{cases} 0 & x = 2, \\ \infty & \text{otherwise.} \end{cases}$$

We assume that $x_0 = 1$, $T = 4$, and the w_t are iid random variables with

$$\mathbf{Prob}(w_t = 1) = p = 0.3 \qquad \text{and} \qquad \mathbf{Prob}(w_t = 2) = 1 - p = 0.7.$$

Since this problem is so small, it is actually possible to solve using a brute-force approach. Executing the function `simple_system_data` defines the following variables.

- `T`, the time horizon $T = 4$
- `p`, the probability $p = 0.3$ that $w_t = 1$
- `a`, the parameter $a = 10$ in the stage-cost function

- x0, the initial state $x_0 = 1$
- xf, the final state $x_f = 2$
- ct, the vector $(c_0, \ldots, c_{T-1})$ of costs associated with moving from state 1 to state 2 (that is, ct(t+1) gives the cost of moving from state 1 to state 2 at time t; note that the offset, which is necessary because time starts at $t = 0$, but MATLAB indexing starts at 1)
- U_seq, a matrix containing every possible input sequence; each column is an input sequence
- W_seq, a matrix containing every possible w_t sequence; each column is a w_t sequence
- pw, a vector containing the probability of each w_t sequence; pw(k) is the probability of the sequence W_seq(:,k)
- phi_cl, a cell array containing the possible closed-loop policies; in particular, phi_cl{k} is a 3-dimensional matrix, where phi_cl{k}(x,w,t+1) gives the value of $\phi_t(x, w)$ (note that we need to add one to the time argument because time starts at $t = 0$, but MATLAB indexing starts at 1); we only include the policies that send the state to x_f at time T; closed-loop policies are explained in part (c) below

You are welcome to look in simple_system_data if you want to see how it works, but this should not be necessary.

(a) An open-loop policy is just an input sequence $(u_0, \ldots, u_{T-1})$. Evaluate the cost

$$\sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T, w_t)$$

for every pair of input and w_t sequences. Find the open-loop policy that achieves the minimum expected cost. Report this minimum expected cost, the index of the optimal policy in the matrix U_seq, and the optimal policy itself.

(b) A prescient policy associates an input sequence with each w_t sequence. You have already evaluated the cost for every pair of input and w_t sequences. Use this information to find the optimal prescient policy for each w_t sequence. Report the expected cost of the optimal prescient policy.

(c) A closed-loop policy is a sequence of functions $(\phi_0, \ldots, \phi_{T-1})$, where each ϕ_t is a map from state/disturbance pairs (x_t, w_t) to actions u_t. Compute the cost of each closed-loop policy under each w_t sequence. Find the closed-loop policy that achieves the minimum expected cost. Report this minimum expected cost, and the index of the optimal policy in the cell array phi_cl.

(d) Plot the probability mass function of the total cost for each of the optimal policies. You can use the command subplot to include all of the plots on one figure.