Homework 1

1. A player is about to play a two-game chess with a computer opponent, and wants to maximize his winning chances. Each game has one of two outcomes.

(a) A win by one of the players (1 point for the winner and 0 for the loser).

(b) A draw (0.5 point for each player).

If the score is tied at 1 − 1 at the end of two games, the match goes into sudden-death mode, whereby the players continue to play until the first time one of them wins a game (and the match).

The player has two playing styles and he can choose one of the two at each game.

(a) Timid play with which he draws with probability $p_d$ and loses with probability $1 - p_d$.

(b) Bold play with which he wins with probability $p_w$ and loses with probability $1 - p_w$.

Thus in any given game, a timid play never wins, and a bold play never draws.

Solve the following questions.

(a) If the score is tied at 1 − 1 at the end of two games, what style should the player choose for the sudden-death mode? What is the probability that the player will win the match conditional on that the score is tied at 1 − 1 at the end of two games?

(b) Use the idea of DP to solve for the best strategy for the player in the two games. (Hint: Let the state be the current score. Write down and solve the DPE. Utilize the result from (a) to find the terminal condition.)

2. *(Deterministic Optimal Growth Model).* The dynamics of the system is defined by

$$X_{n+1} = F(X_n) - c_n, \quad n = 0, 1, \ldots$$
with the initial condition $X_0 = x$ and the control sequence $\{c_n\}$ satisfying the constraint

$$0 \leq c_n \leq F(X_n).$$

Suppose the objective is to maximize the following quantity

$$\sum_{n=0}^{\infty} \beta^n U(c_n)$$

with a discount factor $\beta \in (0, 1)$.

(a) Write down the DPE associated with this control problem, and formally justify it.

(b) Consider the special case where

$$U(c) \doteq \log(c), \quad F(x) = Ax^\alpha$$

for some constants $A > 0$ and $\alpha \in (0, 1)$. Find an explicit solution to the corresponding DPE. (*Hint:* consider a solution of form $a + b \log(x)$ for some constants $a, b$). Also identify the corresponding control policy.

3. Consider a deterministic optimal control whose dynamics is defined by

$$\frac{dX(t)}{dt} = u(t), \quad 0 \leq t \leq 1$$

with initial condition $X(0) = 0$. The control $u = \{u(t)\}$ can take arbitrary values in $\mathbb{R}$. The objective is to minimize the quantity

$$\int_0^1 \left[ 1 + X^2(t) \right] \cdot \left[ 1 + (u^2(t) - 1)^2 \right] \, dt.$$

Show that the value of this control problem is 1 by constructing appropriate controls. Argue that there does not exists an optimal control. (*Hint:* Consider controls taking value $\pm 1$ alternatively).

4. Let $\{X_0, X_1, \ldots, X_N\}$ be a sequence of independent, non-negative, integrable random variables. Define the following sequence of constants $\{A_0, A_1, \ldots, A_N, A_{N+1}\}$ recursively:

$$A_{N+1} \doteq 0 \quad A_N \doteq E [A_{N+1} \lor X_N]$$
\[
\begin{align*}
    A_n &= E[ A_{n+1} \lor X_n ] \\
    &\vdots \\
    A_0 &= E[ A_1 \lor X_0 ]
\end{align*}
\]

Show that
\[
A_0 = \sup_\tau E[ X_\tau ]
\]
where the supremum is taken over all stopping times \( \tau \) taking values in \( \{0, 1, \ldots, N\} \). Also show that an optimal stopping time is given by
\[
\tau^* = \inf \{ n \geq 0 : X_n \geq A_{n+1} \}.
\]

(Hint: Consider the process \( X_n \lor A_{n+1} \), and show it is a supermartingale with respect to the filtration (i.e., information) generated by the sequence \( \{X_n\} \).)

5. Assume that a certain quantity of raw material is required by time \( N \). Denote by \( X_n \) the price of the raw material at time \( n = 0, 1, \ldots, N \). One must decide, given the price at any time, whether to purchase at that price or wait a further period, during which the price may go up or down. Assume that the price dynamics are
\[
X_{n+1} = \lambda X_n + \xi_{n+1}, \quad n = 0, 1, \ldots, N-1,
\]
where \( \{\xi_n\} \) is a sequence of iid non-negative random variables with mean \( \mu = E[\xi_n] > 0 \), and \( \lambda \in [0,1) \) is a constant. The goal is to find a stopping time \( \tau \) taking values in \( \{0, 1, \ldots, N\} \) so as to minimize
\[
E[ X_\tau ]
\]
(a) Write the DPE for this problem.
(b) Show that the optimal policy is as follows: there exist a sequence of positive numbers \( \alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_{N-1} \) such that it is optimal to purchase the raw material the first time when the price \( X_n \) is below \( \alpha_n \).

6. Consider an unemployed worker who is searching for a job. At each time period \( n \), the worker receives an offer \( X_n \). The worker has the option of rejecting the offer, in which case he or she receives \( c \) this period in unemployment compensation. Alternatively, the worker can
accept the offer to work at wage $X_n$, in which case he or she receive a wage $X_n$ per period forever (i.e. the wage is fixed at $X_n$ for each period $j = n, n+1, \ldots$). Neither quitting nor firing is permitted. Let $\beta \in (0, 1)$ be the discounted factor. The goal is to maximizing the total expected discounted income

$$E \left[ \sum_{j=0}^{\tau-1} \beta^j c + \sum_{j=\tau}^{\infty} \beta^j X_j \right] = E \left[ \sum_{j=0}^{\tau-1} \beta^j c + \beta^\tau \frac{X_\tau}{1-\beta} \right]$$

by judiciously choosing a stopping time $\tau$ (which represents the time that he or she accepts the offer).

Assume the offer at time 0 is $X_0$, and that the distribution of the subsequent offers $\{X_n : n = 1, 2, \ldots\}$ are iid, non-negative, bounded random variables with $P\{X_n > c\} > 0$. Write down the DPE and solve it explicitly.

7. A burglar may at any night $n$ choose to retire with his cumulated earnings $X_n$ or enter a house and bring home an amount $\xi_n$ (and thus $X_{n+1} = X_n + \xi_n$). However, in the latter case, he gets caught with probability $p$, and then he is forced to terminate his activities and forfeit his earnings thus far. Assume that $\{\xi_n\}$ are iid, non-negative random variables with mean $\mu > 0$. The goal is to find a policy that maximizes the burglar’s expected earnings over $N$ nights. Write down the DPE and show that the optimal policy is to retire whenever the cumulated wealth $X_n$ exceeds the threshold $(1 - p)\mu/p$. Note this threshold does not depend on $n$ even though it is a finite-horizon problem.

8. In the above problem, consider its infinite horizon counterpart. Write down the DPE. Solve the DPE explicitly under the extra assumption that $\{\xi_n\}$ are iid exponential random variables with rate $\lambda = 1/\mu$.

9. Assume that a certain quantity of raw material is required by time $N$. Denote by $X_n$ the price of the raw material at time $n = 0, 1, \ldots, N$. One must decide, given the price at any time, whether to purchase at that price or wait a further period, during which the price may go up or down. Assume that the price dynamics are

$$X_{n+1} = \lambda X_n + \xi_{n+1}, \quad n = 0, 1, \ldots, N - 1,$$

where $\{\xi_n\}$ is a sequence of iid non-negative random variables with mean $\mu = E[\xi_n] > 0$, and $\lambda \in (0, 1)$ is a constant. The goal is to find
a stopping time $\tau$ taking values in $\{0, 1, \ldots, N\}$ so as to minimize

$$E [X_\tau]$$

(a) Write the DPE for this problem.

(b) Show that the optimal policy is as follows: there exist a sequence of positive numbers $\alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_{N-1}$ such that it is optimal to purchase the raw material the first time when the price $X_n$ is below $\alpha_n$.

10. Consider an unemployed worker who is searching for a job. At each time period $n$, the worker receives an offer $X_n$. The worker has the option of rejecting the offer, in which case he or she receives $c$ this period in unemployment compensation. Alternatively, the worker can accept the offer to work at wage $X_n$, in which case he or she receive a wage $X_n$ per period forever (i.e. the wage is fixed at $X_n$ for each period $j = n, n+1, \ldots$). Neither quitting nor firing is permitted. Let $\beta \in (0, 1)$ be the discounted factor. The goal is to maximizing the total expected discounted income

$$E \left[ \sum_{j=0}^{\tau-1} \beta^j c + \sum_{j=\tau}^{\infty} \beta^j X_\tau \right] = E \left[ \sum_{j=0}^{\tau-1} \beta^j c + \beta^\tau \frac{X_\tau}{1-\beta} \right]$$

by judiciously choosing a stopping time $\tau$ (which represents the time that he or she accepts the offer).

Assume the offer at time 0 is $X_0$, and that the distribution of the subsequent offers $\{X_n : n = 1, 2, \ldots\}$ are iid, non-negative, bounded random variables with $P \{X_n > c\} > 0$. Write down the DPE and solve it explicitly.

11. A burglar may at any night $n$ choose to retire with his cumulated earnings $X_n$ or enter a house and bring home an amount $\xi_n$ (and thus $X_{n+1} = X_n + \xi_n$). However, in the latter case, he gets caught with probability $p$, and then he is forced to terminate his activities and forfeit his earnings thus far. Assume that $\{\xi_n\}$ are iid, non-negative random variables with mean $\mu > 0$. The goal is to find a policy that maximizes the burglar’s expected earnings over $N$ nights. Write down the DPE and show that the optimal policy is to retire whenever the cumulated wealth $X_n$ exceeds the threshold $(1-p)\mu/p$. Note this threshold does not depend on $n$ even though it is a finite-horizon problem.
12. In the above problem, consider its infinite horizon counterpart. Write down the DPE. Solve the DPE explicitly under the extra assumption that \( \{\xi_n\} \) are iid exponential random variables with rate \( \lambda = 1/\mu \).