Problems to be handed in

1. (Markov chains) Recall that a discrete-time random process \( \{X_t\}_{t=0}^{\infty} \), where each \( X_t \) is a random element of some finite set \( \mathcal{X} \), is a Markov chain if, for each \( t \) and all tuples \( x_0^t = (x_0, x_1, \ldots, x_t) \in \mathcal{X}_0^t \),

\[
P[X_t = x_t | X_0^{t-1} = x_0^{t-1}] = P[X_t = x_t | X_{t-1} = x_{t-1}].
\]

In other words, the process distribution of a Markov chain is completely specified by the probability law \( \mu_0 \) of the initial state \( X_0 \) and by the one-step transition probability matrices \( P^{(t)} \), \( t = 0, 1, 2, \ldots \), where both the rows and the columns of each \( P^{(t)} \) are indexed by the elements of the state space \( \mathcal{X} \), and the entries are given by

\[
P^{(t)}(x, x') = P[X_{t+1} = x' | X_t = x], \quad x, x' \in \mathcal{X}.
\]

(i) Let \( \mu_t \) denote the probability distribution of \( X_t \), the state at time \( t \). We represent \( \mu_t \) as a row vector, whose coordinates are indexed by the elements of \( \mathcal{X} \). Prove that, for each \( t \geq 1 \),

\[
\mu_t = \mu_0 P^{(0)} P^{(1)} \cdots P^{(t-1)}.
\]

Let \( g : \mathcal{X} \to \mathbb{R} \) be some real-valued function on the state space \( \mathcal{X} \), which we represent as a column vector with coordinates \( g(x) \), \( x \in \mathcal{X} \). Prove that, for each \( t \) and each \( x \in \mathcal{X} \), the conditional expectation \( E[g(X_t) | X_0 = x] \) is given by

\[
E[g(X_t) | X_0 = x] = \left( P^{(0)} P^{(1)} \cdots P^{(t-1)} g \right)(x)
\]

and that

\[
E[g(X_t)] = \mu_0 P^{(0)} P^{(1)} \cdots P^{(t-1)} g.
\]

(ii) Consider a random process \( \{X_t\}_{t=0}^{\infty} \), where the \( X_t \)'s are generated recursively as follows: \( X_0 \) is drawn from some distribution \( \mu_0 \) over \( \mathcal{X} \), and for each \( t = 0, 1, \ldots \)

\[
X_{t+1} = f_t(X_t, W_t),
\]

for some functions \( f_t \), where \( W_0, W_1, \ldots \) is a sequence of independent random variables that are also independent of the initial state \( X_0 \). Prove that \( \{X_t\} \) is a Markov chain and write down the expression for the entries of the one-step transition probability matrix \( P^{(t)} \) for each \( t \).

(iii) Let \( \{X_t\} \) be a Markov chain with state space \( \mathcal{X} \), specified in terms of the initial state distribution \( \mu_0 \) and the one-step transition probability matrices \( P^{(t)} \), \( t \geq 0 \). Prove that there exists a sequence of i.i.d. random variables \( W_0, W_1, \ldots \) independent of \( X_0 \) and a sequence of functions \( f_t \), such that \( X_{t+1} = f_t(X_t, W_t) \) for each \( t \geq 0 \).

Hint: Let \( W_0, W_1, \ldots \) be i.i.d. Uniform\([0,1]\) random variables.
2 (Controlled Markov chains) Consider the setting of stochastic control with finite state space $X$ and finite action space $U$, with the state update rule

$$x_{t+1} = f_t(x_t, u_t, W_t), \quad t = 0, 1, \ldots$$

(1)

where $W_0, W_1, \ldots$ is a sequence of independent stochastic disturbances.

(i) Let $u \in U$ be a fixed element of the action space and consider a constant policy $g = \{g_t\}_{t \geq 0}$, where $g_t(\cdot) = u$ for each $t$. In other words, at each time step $t$ we ignore all available information and just take the action $u$. Let $X_0$ be a random initial state with distribution $\mu_0$, which is independent of the disturbance process $\{W_t\}$. Prove that the random process $\{X_t\}_{t \geq 0}$ is a Markov chain with initial state distribution $\mu_0$ and with one-step transition probability matrices $P_u^{(t)}$ that depend on $u$. Write down an expression for the entries of $P_u^{(t)}$ for each $t$.

(ii) Let a deterministic sequence $u_0, u_1, \ldots$ of actions be given, and consider a policy $g = \{g_t\}$, where $g_t(\cdot) = u_t$ for each $t$. In other words, at each time $t$, we take a fixed action $u_t$ that may depend on time, but does not depend on the history of states and actions up to time $t$. Policies of this sort are referred to as open-loop policies. Prove that the random process $\{X_t\}_{t \geq 0}$, where $X_0 \sim \mu_0$ is independent of $\{W_t\}$, is a Markov chain and write down the expression for the entries of the one-step transition probability matrix at each time $t$.

(iii) Now consider a Markov policy $g = \{g_t\}$, where, for each $t$, $g_t$ is a function $X \rightarrow U$ that selects an action $u_t$ as a function of the current state $x_t$. Prove that $\{X_t\}$ is a Markov chain and write down the expression for the entries of the one-step transition probability matrix at each time step $t$.

(iv) In the setting of part (iii), prove that the state-action process $\{(X_t, U_t)\}_{t \geq 0}$ is also a Markov chain.

3 (Optimality of Markov strategies) Fill in all the details in the proof of Theorem 1.1 from the first week of lectures.

4 (Stochastic dominance) Recall the definition of stochastic dominance for real-valued random variables: $X$ stochastically dominates $Y$ (which we write as $X \preceq Y$) if $P[X \leq x] \leq P[Y \leq x]$ for all $x \in \mathbb{R}$.

(i) Let $W$ be a zero-mean real-valued random variable. Let $X = m + W$ and $Y = m' + W$ for some $m > m'$. Prove that $X \preceq Y$.

(ii) Consider the following stochastic controlled system with continuous state space $X = \mathbb{R}$, arbitrary action space $U$, and the state update rule

$$x_{t+1} = m(x_t, u_t) + W_t,$$
where $W_0, W_1, \ldots$ are i.i.d. copies of $W$, and where $m : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ is nondecreasing in $x$ for any $u \in \mathcal{U}$,

$$x > x' \implies m(x, u) \geq m(x', u).$$

Prove that this controlled system has the stochastic monotonicity property (the definition is exactly the same as in the finite-state case).

(iii) Give a concrete example to illustrate part (ii).