mas651 exercises

Jaemin Oh

April 14, 2021

Problem (5.1.1)**.**

Let $(S, \mathcal{S})$ be a state space of X_n where $S = \{1, 2, \cdots N\}$ and $\mathcal{S} = 2^S$. Note that N is an absorbing state. And $X_1 = 1$ with probability 1. For fixed k such that $1 \leq k < N$, $k \leq n$,

$$P(X_{n+1} = k+1 | X_n = k) = \frac{N-k}{N}$$

and

$$P(X_{n+1} = k | X_n = k) = \frac{k}{N}.$$

If $k > n$, then the above are all 0. So it is a temporally inhomogeneous. The Markov property is trivial since the very next state only depends on the current state.

□

Problem (5.1.2)**.**

$$P(X_4 = 2 | X_3 = 1, X_2 = 1, X_1 = 1, X_0 = 0) = (1/16)/(1/4) = 1/4$$

but

$$P(X_4 = 2 | X_3 = 1, X_2 = 0, X_1 = 0, X_0 = 0) = (1/16)/(1/8) = 1/2.$$

Thus X_n is not a Markov chain.

□

Problem (5.1.5)**.**

$$P(X_{n+1} = k+1 | X_n = k) = \frac{m-k}{m}\frac{b-k}{m}$$

because we must choose a white ball in the left urn and a black ball in the right urn.

$$P(X_{n+1} = k | X_n = k) = \frac{k}{m}\frac{b-k}{m} + \frac{m-k}{m}\frac{m+k-b}{m}$$

since there are two cases, choosing both black or both white.

$$P(X_{n+1} = k-1 | X_n = k) = \frac{k}{m}\frac{m+k-b}{m}$$

since we must choose a black ball in the left urn and a white ball in the right urn. Note that the sum of the above is 1, so there is no other transition probability.

□

Problem (5.1.6)**.**

$$P(S_{n+1} = k+1 | S_n = k) = \frac{P(X_{n+1} = 1, S_n = k)}{P(S_n = k)}$$

where the denominator is

$$\int_{\theta\in(0,1)} P(S_n = k|\theta)dP = \binom{n}{x}\frac{x!y!}{(n+1)!} = \frac{1}{n+1}$$

for $x =$*the number of* i *such that* $U_i \le \theta$ *and* $y = n - x$. *Note that* $x = (n+k)/2$ *and* $y = (n-k)/2$ *since* $x + y = n$ *and* $x - y = k$. *The numerator is*

$$\int_{\theta\in(0,1)} P(X_{n+1} = 1, S_n = k|\theta)dP = \binom{n}{x}\frac{(x+1)!y!}{(n+2)!}$$

These are because $P(S_n = k|\theta) = \theta^x(1-\theta)^y\binom{n}{x}$ *and* $P(X_{n+1} = 1, S_n = k|\theta) = P(X_{n+1} = 1|\theta)P(S_n = k|\theta) = \binom{n}{x}\theta^{x+1}(1-\theta)^y$ *and using the kernel of beta distribution.*

Thus, the probability what we want is $(n+k+2)/(2n+4)$ *which depends on* n. *So* X_n *is temporally inhomogeneous.*

$$P(S_{n+1} = k+1|S_1 = t_1, \cdots, S_n = k) = P(X_{n+1} = 1|X_1 = t_1, \cdots, X_n = t_n)$$

where $\sum_{i=1}^n t_i = k$. *We can show the above is equal to* $P(S_{n+1} = k+1|S_n = k) = (n+k+2)/(2n+4)$ *similarly, by omitting the* $\binom{n}{x}$ *term of both denominator and numerator.*

$\square$

Problem (5.2.1)**.**

By the given hint,

$$E(1_A 1_B|\mathcal{F}_n) = E(1_A E(1_B|\mathcal{F}_n)|X_n)$$

so it suffices to show that $E(1_B|\mathcal{F}_n) = E(1_B|X_n)$.

Let $Y = 1_{B_n}(\omega_0)\cdots 1_{B_{n+k}}(\omega_k)$. *Then* $Y \circ \theta_n =$ *the indicator function of* $\{X_n \in B_n, \cdots, X_{n+k} \in B_{n+k}\} = B$. *By the markov property,*

$$P(B|\mathcal{F}_n) = E_{X_n}Y.$$

Let $\varphi(x) = E_x Y$ *then* $\varphi(X_n)$ *is* $\sigma(X_n)$*-measurable mapping. Thus, when* B *has a form of* $\{X_n \in B_n, \cdots, X_{n+k} \in B_{n+k}\}$ *for some nonnegative integer* k,

$$P(B|\mathcal{F}_n) = P(B|X_n).$$

Note that a collection of such B *generates* $\sigma(X_n, X_{n+1}, \cdots)$.

Now let $\mathcal{G} = \{C : P(C|\mathcal{F}_n) = P(C|X_n)\}$. *By putting* $B_{n+i} = S$ *for* $0 \le i \le k$, *we earn* $\Omega_0 \in \mathcal{G}$. *If* $C, D \in \mathcal{G}$ *and* $C \subset D$, *then by properties of conditional expectation,* $D \setminus C \in \mathcal{G}$. *If* $C_i \in \mathcal{G}$ *and* $C_i \uparrow C$ *then by monotone convergence theorem for conditional expectation,* $C \in \mathcal{G}$. *Thus* $\mathcal{G}$ *is a lambda system containing a collection of* B*'s which generates* $\sigma(X_n, \cdots)$. *Therefore, by Dynkin's theorem, the third equation is satisfied by any* $B \in \sigma(X_n, \cdots)$. *By the first equation, we can derive the conclusion.*

$\square$

Problem (5.2.4)**.**

First, claim that

$$P_x\left(X_n = y | T_y = m\right) = P_y\left(X_{n-m} = y\right).$$

This is because

$$\begin{aligned}
P_x\left(X_n = y | T_y = m\right) &= \frac{P_x\left(X_n = y, T_y = m\right)}{P_x\left(T_y = m\right)} \\
&= \frac{\int_{T_y = m} 1_{(X_n = y)} dP_x}{P_x\left(T_y = m\right)} \\
&= \frac{\int_{T_y = m} E\left(1_{(X_n = y)}|\mathcal{F}_m\right) dP_x}{P_x\left(T_y = m\right)} \\
&= \frac{\int_{T_y = m} E_{X_m} 1_{(X_{n-m} = y)} dP_x}{P_x\left(T_y = m\right)} \\
&= \frac{P_x\left(T_y = m\right) P_y\left(X_{n-m} = y\right)}{P_x\left(T_y = m\right)}.
\end{aligned}$$

Now, note that $P_x\left(X_n = y\right) = \sum_{m=1}^{n} P_x\left(X_n = y, T_y = m\right)$. *From this and* *the above discussion,*

$$\begin{aligned}
p^n(x,y) = P_x\left(X_n = y\right) &= \sum_{m=1}^{n} P_x\left(X_n = y | T_y = m\right) P_x\left(T_y = m\right) \\
&= \sum_{m=1}^{n} P_y\left(X_{n-m} = y\right) P_x\left(T_y = m\right) = \sum_{m=1}^{n} P_x\left(T_y = m\right) p^{n-m}(y,y).
\end{aligned}$$

□

Problem (5.2.6).

Fix $x \in S \setminus C$. Since $P_x(T_C = \infty) = \lim_{M\to\infty} P_x(T_C > M) < 1$, we can choose N_x and ε so that

$$P_x(T_C > M) \leq 1 - \varepsilon$$

whenever $M \geq N_x$. Note that we can choose N_x as an integer. Put $N = \max_{x\in S\setminus C} N_x$. Now we get

$$\begin{aligned}
P_y(T_C > 2N) &= \sum_{x\in S\setminus C} P_y(T_C > 2N, T_C > N, X_N = x) \\
&= \sum_{x\in S\setminus C} P_y\left(T_C > 2N | X_N = x, T_C > N\right) P_y(X_n = x, T_C > N) \\
&\leq \sum_{x\in S\setminus C} P_x(T_C > N) P_y(X_N = x, T_C > N) \\
&\leq (1-\varepsilon) \sum_{x\in S\setminus C} P_y(X_N = x, T_C > N) \\
&\leq (1-\varepsilon)^2.
\end{aligned}$$

By induction, the result follows.

Remark 1. By $k \to \infty$, we can say that $P_y\left(T_C = \infty\right) = 0$. That is, $P_y\left(T_C < \infty\right) = 1$.

□

Problem (5.2.7).

1. *It is similar to the manipulation of problem 5.2.4:*

$$\begin{aligned}
P_x\left(V_A < V_B\right) &= \sum_y P_x\left(V_A < V_B, X_1 = y\right) = \sum_y P_x\left(V_A < V_B | X_1 = y\right) P_x\left(X_1 = y\right) \\
&= \sum_y p(x,y) P_x\left(V_A < V_B | X_1 = y\right) = \sum_y p(x,y) P_y\left(V_A < V_B\right)
\end{aligned}$$

where the first term is $h(x)$ and the last term is $\sum_y p(x,y)h(y)$.

2. *I think we must further assume that h is bounded and measurable. For convenience, let $\tau = V_A \wedge V_B = V_{A\cup B}$. By the equation (5.2.2) of our textbook, we get*

$$\begin{aligned} E_x\left(h(X_{n+1})|\mathcal{F}_n\right) &= \sum_y p(X_n, y)h(y) \\ &= h(X_n) \end{aligned}$$

for $X_n \notin A \cup B$.

Now, put $Y_n = h(X_n)$. Then

$$Y_{n\wedge\tau} - Y_0 = h\left(X_{n\wedge\tau}\right) - h(X_0) = \sum_{k=1}^{n} 1_{(\tau\geq k)}\left(Y_k - Y_{k-1}\right).$$

By using the above,

$$\begin{aligned} E_x\left(Y_{n+1\wedge\tau} - Y_0|\mathcal{F}_n\right) &= \sum_{k=1}^{n+1} 1_{(\tau\geq k)} E_x\left(Y_k - Y_{k-1}|\mathcal{F}_n\right) \\ &= 1_{(\tau\geq n+1)}\left(Y_n - Y_0\right) + 1_{(\tau<n+1)}\left(Y_\tau - Y_0\right) \\ &= 1_{(\tau>n)}(Y_n - Y_0) + 1_{(\tau\leq n)}(Y_\tau - Y_0) \\ &= Y_{n\wedge\tau} - Y_0. \end{aligned}$$

So $h(X_{n\wedge\tau})$ is a martingale. Note that the first equality is due to (5.2.2), and the last is due to optional stopping.

3. *We assumed that h is bounded. Thus, our martingale is uniformly bounded, so the optional stopping theorem can be applied:*

$$x = E_x h(X_0) = E_x h(X_\tau) = E_x\left[E_x\left(h(X_\tau)|\mathcal{F}_\tau\right)\right]$$

where the last term is equal to

$$E_x\left[E_{X_\tau} h(X_0)\right] = E_x\left[1_{(X_\tau\in A)} + 0\cdot 1_{(X_\tau\in B)}\right] = E_x\left[1_{(X_\tau\in A)}\right].$$

The above is because $P_x(\tau < \infty) = 1$ and h is 1 on A and 0 on B. Note that this implies the result, since $X_\tau \in A$ is equivalent to $V_A < V_B$.

□

Problem (5.2.8)**.**

Let $\tau = V_0 \wedge V_N$. Then $X_{n\wedge\tau}$ is an uniformly bounded martingale since the state space of X_n is finite. By the optional stopping theorem, we get

$$E_x X_0 = E_x X_\tau$$

where the LHS is equal to x. Note that, by the remark 1, we can say that $P_x\left(\tau < \infty\right) = 1$. Then the RHS of the above eqn is equal to $0P_x\left(V_0 < V_N\right) + NP_x\left(V_N < V_0\right)$. Thus,

$$x = NP_x\left(V_N < V_0\right).$$

□

Problem (5.2.11)**.**

1. *It is similar to the manipulation of problem 5.2.7:*

$$\begin{aligned}
E_x V_A &= \sum_{k\geq 1} P_x\left(V_A \geq k\right) \\
&= P_x\left(V_A \geq 1\right) + \sum_{k\geq 2} P_x\left(V_A \geq k\right) \\
&= 1 + \sum_{k\geq 2} P_x\left(V_A \geq k\right) \\
&= 1 + \sum_{k\geq 2}\sum_{y} P_x\left(V_A \geq k | X_1 = y\right) P_x\left(X_1 = y\right) \\
&= 1 + \sum_{y} p(x,y) \sum_{k\geq 2} P_y\left(V_A \geq k-1\right) = 1 + \sum_y p(x,y) E_y V_A
\end{aligned}$$

   where $P_x\left(V_A \geq 1\right) = 1$ *since* x *lies outside of* A.

   Also, $E_x V_A < \infty$ *because*

$$\begin{aligned}
E_x \frac{V_A}{N} &= \sum_{k\geq 1} P_x\left(V_A \geq kN\right) \\
&\leq \sum_{k\geq 1} \left(1-\varepsilon\right)^k < \infty.
\end{aligned}$$

2. *I think we should assume the measurability, and boundedness of* g. *By the manipulation used in problem 5.2.7, we get:*

$$\begin{aligned}
E_x\left(g(X_{n+1}) + n + 1 | \mathcal{F}_n\right) &= n + 1 + \sum_y p(X_n, y) g(y) \\
&= n + g(X_n)
\end{aligned}$$

   for $X_n \notin A$.

   Now put $Y_n = g(X_n) + n$ *and* $\tau = V_A$ *for convenience. Then*

$$\begin{aligned}
E_x\left(Y_{n+1\wedge\tau} - Y_0 | \mathcal{F}_n\right) &= \sum_{k=1}^{n+1} 1_{(\tau\geq k)} E_x\left(Y_k - Y_{k-1} | \mathcal{F}_n\right) \\
&= 1_{(\tau\geq n+1)}(Y_n - Y_0) + 1_{(\tau< n+1)}(Y_\tau - Y_0) \\
&= Y_{n\wedge\tau} - Y_0.
\end{aligned}$$

   So $X_{n\wedge V_A} + n \wedge V_A$ *is a martingale.*

3. *From the boundedness of* g *and the fact that* V_A *is* L^1 *function, our martingale is uniformly integrable. Thus we can apply optional stopping theorem:*

$$E_x g(X_0) = E_x\left[V_A + g(X_{V_A})\right]$$

where the first term is $g(x)$ and the second term is $E_x V_A + E_x g(X_{V_A})$. But X_{V_A} lies in A and g is 0 on A. Thus the second term of the equation is $E_x V_A$.

$\square$

Problem (5.3.1)**.**

Abbreviation of notation: $P(X_1 \le x_1, \cdots, X_n \le x_n)$ *as* $P(X \le x)$*, which is a distribution function of a random vector.*

First, let's see that they are identically distributed. Since y *is recurrent, the strong Markov property always holds when we are considering* $\mathcal{F}_{R_k}$*.*

$$P_y(\nu_k \le x) = E_y P_y(\nu_k \le x | \mathcal{F}_{R_{k-1}}) = E_y E_{X_{R_{k-1}}} 1_{(\nu_1 \le x)} = P_y(\nu_1 \le x).$$

So ν_k *are identically distributed.*

Now, let's see that they are independent.

$$\begin{aligned} P_y(\nu_1 \le x_1, \cdots, \nu_n \le x_n) &= E_y P_y(\nu_1 \le x_1, \cdots, \nu_n \le x_n | \mathcal{F}_{R_{n-1}}) \\ &= E_y \left[1_{(\nu_1 \le x_1)} \cdots 1_{(\nu_{n-1} \le x_{n-1})} P_y(\nu_1 \le x_n) \right] \\ &= P_y(\nu_1 \le x_n) P_y(\nu_1 \le x_1, \cdots, \nu_{n-1} \le x_{n-1}) \\ &= \cdots \\ &= P_y(\nu_n \le x_n) \cdots P_y(\nu_1 \le x_1) \end{aligned}$$

since they are identically distributed. Note that $\nu_1, \cdots \nu_{n-1}$ *are* $\mathcal{F}_{R_{n-1}}$ *measurable. This is because* $\left\{X_{R_{n-2}+i} \in B\right\} \cap \{R_{n-1} = k\} \in \mathcal{F}_k$*.*

□

Problem (5.3.2)**.**

On $\{T_y < \infty\}$*, by the strong Markov property,*

$$\rho_{xy} = P_y(T_z < \infty) = E_x \left(1_{(T_y + T_z < \infty)} | \mathcal{F}_{T_y} \right).$$

Thus,

$$\begin{aligned} \rho_{xy}\rho_{yz} &= E_x 1_{(T_y < \infty)} E_x \left(1_{(T_y + T_z < \infty)} | \mathcal{F}_{T_y} \right) \\ &= E_x 1_{(T_y < \infty)} 1_{(T_y + T_z < \infty)} \\ &= P_x \left(T_y + T_z < \infty \right) \le P_x \left(T_z < \infty \right) = \rho_{xz}. \end{aligned}$$

Problem (5.3.5)**.**

As in the proof of theorem 5.3.8, $\varphi(X_{n\wedge\tau})$ *is a nonnegative supermartingale. So the supmartingale converges to* Y *a.s. From the modified condition* $\varphi \to 0$*, we know that* $\{x : \varphi(x) > M\}$ *is a finite set. So* X_n *visits* $\{x : \varphi(x) > M\}$ *only finitely many times for all* $M > 0$*. Thus* $\varphi(X_n) \to 0$ *as* $n \to \infty$*. If* $\tau < \infty$ *almost surely, then* $\varphi(X_{n\wedge\tau}) \to \varphi(X_\tau) = 0$*. But we have the other condition:* $\varphi > 0$ *on* F*. So* $\varphi(X_\tau) = 0$ *cannot happen; thus* $P_x(\tau = \infty) > 0$*.*

If the chain is recurrent, then $\tau < \infty$ *a.s. But our case is not the case, so the chain must be transient.*

□

Problem (5.3.7)**.**

First assume the recurrence. Let f *be a superharmonic function, so* $f(X_n)$ *is a nonnegative supermartingale. By the martingale convergence theorem,* $f(X_n) \to$

Y a.s. By the recurrence, $P(X_n = x$ i.o.$) = 1$ for all $x \in S$. So $P(f(X_n) = f(x)$ i.o.$) = 1$, which says $f(X_n) \to f(x) = Y$ a.s. But $x \in S$ is arbitrary, f must be a constant.

Now, assume the transience. Fix $z \in S$. Let $V = \inf\{n \geq 0 : X_n = z\}$. Let $f(x) = P_x(V < \infty)$. We will show that f is a nonconstant superharmonic function. For $x \neq z$,

$$f(x) = P_x(V < \infty) = \sum_y p(x,y) P_y(V < \infty) = \sum_y p(x,y) f(y).$$

For $x = z$,

$$\sum_y p(z,y) f(y) \leq \sum_y p(z,y) = 1 = f(z)$$

since $f \leq 1$. Thus, f is superharmonic.

Now claim that there is $y \in S$ such that $f(y) < 1$. If no such y exists, then $f(y) = 1$ for all $y \in S$. This says that $P_y(V < \infty) = 1$ for all y which is equivalent to the recurrence of z. But we assumed the recurrence of our chain. So there is $y \in S : f(y) < 1$. And this says f is nonconstant.

$\square$

Problem (5.5.2)**.**

Let N_y^x be the number of hittings to y before returning to x. Then $\mu_x(y) = E_x N_y^x$. Also, we can write

$$N_y^x = \sum_{k=1}^{\infty} 1_{\left(T_y,\cdots,T_y^k < T_x\right)} = \sum_{k=1}^{\infty} 1_{\left(T_y^k < T_x\right)}.$$

By using the strong Markov property and the induction, we get

$$\begin{aligned}
E_x N_y^x &= \sum_{k=1}^{\infty} P_x\left(T_y^k < T_x\right) \\
&= \sum_{k=1}^{\infty} E_x\left[P_x\left(T_y, T_y^k < T_x | \mathcal{F}_{T_y}\right)\right] \\
&= \sum_{k=1}^{\infty} E_x\left[1_{(T_y < T_x)} P_y\left(T_y^{k-1} < T_x\right)\right] \\
&= \cdots \\
&= \sum_{k=1}^{\infty} P_x(T_y < T_x) P_y(T_y < T_x)^{k-1} \\
&= \frac{P_x(T_y < T_x)}{1 - P_y(T_y < T_x)} = \frac{w_{xy}}{w_{yx}}.
\end{aligned}$$

□

Problem (5.5.3)**.**

Irreducibility and recurrence implies the existence of unique(up to constant multiple) stationary measure. The recurrence implies that $y \mapsto \mu_x(y)$ is the stationary measure. Thus, for some $c > 0$, $\mu_y(z) = \mu_x(z) \cdot c$. So $\mu_y(z)p(z,y) = c\mu_x(z)p(z,y)$. By adding over z,

$$1 = \mu_y(y) = \sum_z \mu_y(z)p(z,y) = c\sum_z \mu_x(z)p(z,y) = c\mu_x(y).$$

Thus $c = 1/\mu_x(y)$ and this leads the result.

□

Problem (5.5.4)**.**

$E_x T_y = \infty$ says that the chain is expected to not reach the state y. So $\mu_x(y) = 0$, contradiction. Also

$$\sum_x E_x T_y p(x,y) + 1 = E_y T_y < \infty$$

by positive recurrence. Therefore $E_x T_y$ should be finite.

□

Problem (5.5.5).

On the contrary, assume that p is positive recurrent. Then, with irreducibility, the existence of stationary distribution π is guaranteed. Also, any stationary measure is constant multiple of π by theorem 5.5.9. Thus $c\pi(x) = \mu(x)$ for some constant $c > 0$. Then,

$$\infty = \sum_x \mu(x) = c \sum_x \pi(x) = c < \infty,$$

which is a contradiction. Hence p cannot be positive recurrent.

□

Problem (5.5.9).

Note that $Y_{n\wedge\tau} - Y_0 = \sum_{k=1}^{n} 1_{(\tau\geq k)} (Y_k - Y_{k-1})$. Using this,

$$\begin{aligned} E_x (Y_{n+1\wedge\tau} - Y_0|\mathcal{F}_n) &= \sum_{k=1}^{n} 1_{(\tau\geq k)} (Y_k - Y_{k-1}) + 1_{(\tau\geq n+1)} [E_x (X_{n+1}|\mathcal{F}_n) - X_n + \varepsilon] \\ &\leq \sum_{k=1}^{n} 1_{(\tau\geq k)} (Y_k - Y_{k-1}) + 1_{(\tau\geq n+1)} [X_n - \varepsilon - X_n + \varepsilon] \\ &= Y_{n\wedge\tau} - Y_0. \end{aligned}$$

Thus $Y_{n\wedge\tau}$ is a nonnegative supermartingale. Now, by theorem 4.8.4, we have

$$x = E_x[Y_0] \geq E_x[Y_\tau] = E_x[X_\tau + \tau\varepsilon] \geq \varepsilon E_x[\tau]$$

since $X_n \geq 0$. By dividing both sides by ε, we earn the result.

□

Problem (5.6.1).

For $n = 1$,

$$P_\mu(X_1 = 0) = \mu(0)p(0,0) + (1-\mu(0))p(1,0) = (1-\alpha-\beta)\mu(0) + \beta.$$

Now,

$$\begin{aligned}
P_\mu(X_{n+1} = 0) &= E_\mu P_\mu\left(X_{n+1} = 0|\mathcal{F}_n\right) \\
&= E_\mu\left[P_0\left(X_1 = 0\right)1_{(X_n=0)} + P_1\left(X_1 = 0\right)1_{(X_n=1)}\right] \\
&= P_\mu(X_n = 0)p(0,0) + P_\mu(X_n = 1)p(1,0) \\
&= (1-\alpha)P_\mu(X_n = 0) + \beta(1 - P_\mu(X_n = 0)) \\
&= (1-\alpha-\beta)P_\mu(X_n = 0) + \beta \\
&= (1-\alpha-\beta)^{n+1}\left[\mu(0) - \frac{\beta}{\alpha+\beta}\right] + \frac{\beta(1-\alpha-\beta)}{\alpha+\beta} + \beta
\end{aligned}$$

where the last term is what we desired.

□

Problem (5.6.2).

Aperiodicy of state x is defined only when the state x is recurrent. So aperiodicy of the chain necessarily contains recurrence of the chain. Note that recurrent chain has stationary measure, $\mu_x(y)$. With finiteness of the state space, by normalizing, we can earn the stationary distribution π. With irreducibility of the chain, by existence of the stationary measure, the chain is positive recurrent. Now, for any $x, y \in S$, by convergence theorem,

$$p^m(x,y) \to \pi(y) > 0.$$

So we can find M_{xy} such that $m \geq M_{xy}$ implies $p^m(x,y) > \pi(y)/2 > 0$. Take $M = \max_{x,y} M_{xy}$. Then $M < \infty$ because of the finiteness. For any $x, y \in S$,

$$p^M(x,y) > \pi(y)/2 > 0.$$

□

Problem (5.6.3).

By the previous problem, there is $m > 0$ such that $p^m(x,y) > 0$ for all $x, y \in S$. Fix $y \in S$. Let $p = \min_{x\in S} p^m(x,y)$. Then

$$P\left(X_{n+m} = Y_{n+m} = y|X_n = x_1, Y_n = x_2\right) \geq p^2$$

for all $x_1, x_2 \in S$ by the definition of $\bar{p}$. So $P\left(X_{n+m} = Y_{n+m}|X_n, Y_n\right) \geq p^2$, *which is equivalent to* $P\left(X_{n+m} \neq Y_{n+m} | X_n, Y_n\right) \leq 1 - p^2$. *Now, consider*

$$
\begin{aligned}
&P(T > km) \\
&= P\left(X_1 \neq Y_1, \cdots, X_{km} \neq Y_{km}\right) \\
&= EP(\,\cdot\,|\mathcal{F}_{(k-1)m}) \\
&= EP_{X_{(k-1)m}, Y_{(k-1)m}}\left(X_{km} \neq Y_{km}, \cdots, X_{(k-1)m+1} \neq Y_{(k-1)m+1}\right) 1_{\left(X_{(k-1)m} \neq Y_{(k-1)m}, \cdots, X_1 \neq Y_1\right)} \\
&\leq (1-p^2) P(X_1 \neq Y_1, \cdots, X_{(k-1)m} \neq Y_{(k-1)m}) \\
&\leq \cdots \\
&\leq (1-p^2)^k.
\end{aligned}
$$

Therefore

$$
P(T > n) \leq P(T > \left[\frac{n}{m}\right] m) \leq (1-p^2)^{\left[\frac{n}{m}\right]} \leq (1-p^2)^{\frac{n}{m}}
$$

where $[\cdot]$ *is the floor function. So the convergence occurs at least exponentially fast.*

□

Problem (5.6.5)**.**

Note that $P_x(T_x^k < \infty) = 1$. *Let* $V_k^f = V_k$, *and* $V_k^{|f|} = V_k'$.

1. *By the strong Markov property,*

$$
\begin{aligned}
P(V_k \leq a) &= EP\left(V_k \leq a | \mathcal{F}_{T_x^k}\right) \\
&= P_x\left(f(X_0) + \cdots + f(X_{T_x^1 - 1}) \leq a\right) = P_x(V_0 \leq a).
\end{aligned}
$$

Thus $\left\{V_k^f\right\}_{k=1}^{\infty}$ *is an identically distributed sequence.*

By the strong Markov property,

$$
\begin{aligned}
P(V_k \leq a_0, \cdots, V_{k+m} \leq a_m) &= EP(\cdot | \mathcal{F}_{T_x^{k+m}}) \\
&= EP_x(V_1 \leq a_m) 1_{(V_k \leq a_0, \cdots, V_{k+m-1} \leq a_{m-1})} \\
&= \cdots \\
&= \prod_{n=0}^{m} P_x\left(V_0 \leq a_n\right) \\
&= P\left(V_k \leq a_0\right) \cdots P\left(V_{k+m} \leq a_m\right)
\end{aligned}
$$

where the last equality is due to the previous result. So they are independent.

Now, consider $E|V_1|$.

$$\begin{aligned}
&\le E\sum_{k\ge 1} 1_{(T_x^1<k\le T_x^2)}|f(X_k)|(=EV_1')\\
&= E\sum_{k\ge 1} 1_{(T_x^1<k\le T_x^2)}\sum_y |f(y)|1_{(X_k=y)}\\
&= \sum_y |f(y)|E\sum_{k\ge 1} 1_{(T_x^1<k\le T_x^2)}1_{(X_k=y)}\\
&= \sum_y |f(y)|\mu_x(y) = \sum_y |f(y)|\pi(y)E_xT_x^1 < \infty.
\end{aligned}$$

Similarly, we can get $EV_1 = \sum_y f(y)\pi(y)E_xT_x^1$.

2. *Note that* $N_n(x) = \sup\left\{k : T_x^k \le n\right\}$. *So* $N_n(x) \le K_n$. *If* $N_n(x) < K_n$, *then* $K_n = N_n(x)+1$. *By SLLN and theorem 5.6.1, as* $n \to \infty$,

$$\frac{1}{n}\sum_{m=1}^{K_n} V_m = \frac{N_n(x)}{n}\frac{K_n}{N_n(x)}\frac{1}{K_n}\sum_{m=1}^{K_n} V_m \to \frac{EV_1}{E_xT_x^1} = \sum_y f(y)\pi(y)$$

P_μ *almost surely since* $N_n(x) \to \infty$.

3. *Refer to problem 2.3.17, sufficient(in fact, necessary also) condition for* $\max_{m\le n} V_m'/n \to 0$ *is* $EV_m'^{+} < \infty$ *for all* m. *This is because*

$$\sum_{n\ge 1} P(V_n' \ge n\delta) < \infty$$

for all $\delta > 0$. *Thus, by Borel Cantelli lemma,* $P(V_n' \ge n\delta \text{ i.o.}) = 0$. *So*

$$P(V_n'/n < \delta \text{ all but finitely many } n) = 1.$$

Also,

$$\frac{1}{n}\max_{m\le n} V_m' \le \frac{1}{n}\left(\max_{m\le M} V_m' + \max_{M<m\le n} V_m'\right) \le \frac{\max_{m\le M} V_m'}{n} + \max_{M<m\le n}\frac{V_m'}{m}$$

where the last term $\le \delta$ *as* $n \to \infty$. *Since* δ *is arbitrary, we can get* $\max_{m\le n} V_m'/n \to 0$ *a.s.*
Now, note that $K_n \le n$. *Then*

$$\frac{1}{n}\left|\sum_{m=1}^{n} f(X_m) - \sum_{m=1}^{K_n} V_m\right| \le \frac{1}{n}\max_{m\le n} V_m' \to 0$$

as $n \to \infty$.

□

Problem (6.1.1).

Let φ be a measure preserving map of Ω into Ω. First, $\varphi^{-1}\Omega = \Omega$ since codomain contains the range. Second, if A is an invariant set, then $\varphi^{-1}A^c = \left(\varphi^{-1}A\right)^c = A^c$, so the complement of A is invariant. Third, if A_i is an increasing sequence of invariant set, then $\varphi^{-1}A = \cup_i \varphi^{-1}A_i = \cup_i A_i = A$, so $A = \cup_i A_i$ is also invariant. Therefore $\mathcal{I}$ is a sigma field.

We say two sets A, B are equal a.s. if their corresponding indicator functions are equal a.s.

Let B be a Borel set. Assume that $X \circ \varphi = X$ a.s. Then $\varphi^{-1}X^{-1}B = X^{-1}B$ a.s. so $X^{-1}B$ is invariant, which says X is $\mathcal{I}$ measurable. Let's consider the converse. If X is an indicator function, then the result trivially holds. So we can extend the result to where X is an simplie function, measurable with respect to $\mathcal{I}$. If X is nonnegative function of $\mathcal{I}$, then $s_n \uparrow X$. Since $s_n \circ \varphi = s_n$ a.s, $X \circ \varphi = \lim_n s_n \circ \varphi = \lim_n s_n = X$ a.s. If $X \in \mathcal{I}$ any r.v, then by decomposing it to $X = X^+ - X^-$, we can conclude the result.

□

Problem (6.1.2).

1.
$$\omega \in \varphi^{-1}(B) \Rightarrow \varphi(\omega) \in B \Rightarrow \varphi(\omega) \in \varphi^{-n}(A) \Rightarrow \varphi^{n+1}(\omega) \in A$$
$$\Rightarrow \omega \in \varphi^{-n-1}(A) \Rightarrow \omega \in B$$

   for some $n \geq 0$. Therefore $\varphi^{-1}(B) \subset B$.

2.
$$\omega \in \varphi^{-1}(C) \Rightarrow \varphi(\omega) \in \varphi^{-n}(B) \Rightarrow \varphi^{n+1}(\omega) \in B$$
$$\Rightarrow \omega \in \varphi^{-n-1}(B) \Rightarrow \omega \in \varphi^{-n}(B) \Rightarrow \omega \in C$$

   for all $n \geq 0$. Therefore $\varphi^{-1}(C) \subset C$.

$$\omega \in C \Rightarrow \varphi^n(\omega) \in B \Rightarrow \omega \in (\varphi^{-n+1} \circ \varphi^{-1})(B)$$
$$\Rightarrow \omega \in \varphi^{-n+1}(B) \Rightarrow \omega \in \bigcap_{n \geq 1} \varphi^{-n}(B) = \varphi^{-1}(C)$$

   for all $n \geq 0$. Therefore $C = \varphi^{-1}(C)$.

3. *Let B, C be same as above. Assume that A is almost invariant. Then $A = \varphi^{-n}(A)$ almost surely, so $A = B$ a.s. So $A = C$ a.s, which is equivalent to $P(A\Delta C) = 0$.*

   Conversely, assume that $P(A\Delta D) = 0$ for some strictly invariant D. Since φ is measure preserving, $P(\varphi^{-1}(A\Delta D)) = P(A\Delta D) = 0$. So $\varphi^{-1}(A) = \varphi^{-1}(D)$ a.s. But $A = C = \varphi^{-1}(C)$ a.s. Thus $A = \varphi^{-1}(A)$a.s, equivalent to almost invariance of A.

□

Problem (6.1.4).

Let m be any integer, n be any nonnegative integer. Define

$$\mu_{m,\cdots,m+n}(A_0,\cdots,A_n) = P(X_0 \in A_0,\cdots,X_n \in A_n).$$

Then $\mu_{m,\cdots,m+n}$ is consistent, so the Kolmogorov extension thm applies. Let Y_n be a coordinate map. Then any length $n+1$ distribution of consecutive sequence of Y has same distribution with $X_0,\cdots,X_n$. It means Y_n is a two sided stationary process, and X_n is embeded in Y_n

□.

Problem (6.2.1).

Assume $X \in L^p$. Let $A_n(X'_M) = \sum_{m=0}^{n-1} X'_M \circ \varphi^m/n$. Since $A_n(X'_M) \to E(X'_M|\mathcal{I})$ a.s. and $|A_n(X'_M) - E(X'_M|\mathcal{I})|^p \le (2M)^p \in L^1$, DCT implies L^p convergence of $A_n(X'_M) \to E(X'_M|\mathcal{I})$.

Now consider $\|A_n(X''_M) - E(X''_M|\mathcal{I})\|_p \le \|A_n(X''_M)\|_p + \|X''_M\|_p$.
But $\|A_n(X''_M)\|_p \le \sum_{m=0}^{n-1} \|X''_M \circ \varphi^m\|_p/n = \|X''_M\|_p$. Since $|X''_M|^p \le |X|^p \in L^1$ and $X''_M \to 0$ a.s, DCT implies L^p convergence of $A_n(X''_M) \to E(X''_M|\mathcal{I})$.

□

Problem (6.2.2).

1. *Fix $M > 0$. Let $h_M = \sup_{m \ge M} |g_m - g|$. By our assumption, $h_M \in L^1$ and $h_M \downarrow 0$ a.s. as $M \uparrow \infty$. Since $g \in L^1$,*

$$\frac{1}{n}\sum_{m=0}^{n-1} g_m \circ \varphi^m \to E(g|\mathcal{I}) \ a.s$$

So, it is sufficient to show that

$$\frac{1}{n}\sum_{m=0}^{n-1} (g_m - g) \circ \varphi^m \to 0 \ a.s.$$

Consider

$$\frac{1}{n}\sum_{m=0}^{n-1} |g_m - g| \circ \varphi^m \le \frac{1}{n}\sum_{m=0}^{M-1} |g_m - g| \circ \varphi^m + \frac{1}{n}\sum_{m=M}^{n-1} h_M \circ \varphi^m.$$

By taking $\limsup_{n\to\infty}$ on both sides,

$$\limsup_{n\to\infty} \frac{1}{n}\sum_{m=0}^{n-1} |g_m - g| \circ \varphi^m \le E(h_M|\mathcal{I}).$$

By theorem 4.1.9(Monotone convergence theorem for conditional expectation), as $M \uparrow \infty$, the last term goes to 0 a.s.

2. *Since $g \in L^1$,*

$$\frac{1}{n}\sum_{m=0}^{n-1} g \circ \varphi^m \to E(g|\mathcal{I})$$

a.s. and in L^1 by the Ergodic theorem. Now, it is sufficient to show that

$$\frac{1}{n}\sum_{m=0}^{n-1} (g_m - g) \circ \varphi^m \to 0$$

in L^1 sense. Fix $\varepsilon > 0$. Since $\|g_n - g\|_1 \to 0$, we can choose N such that $\|g_n - g\|_1 < \varepsilon$ whenever $n \geq N$. Then,

$$\left\| \frac{1}{n}\sum_{m=0}^{n-1} (g_m - g) \circ \varphi^m \right\|_1 \leq \frac{1}{n}\sum_{m=0}^{N} \|g_m - g\|_1 + \varepsilon.$$

If n is sufficiently large, then the above is bounded by ε. Since ε is arbitrary, the above goes to 0 in L^1 sense.

$\square$

Problem (6.2.3)**.**
Note that

$$D_k - \alpha > 0 \Leftrightarrow \sup_{i \leq k} \frac{S_i - i\alpha}{i} > 0 \Leftrightarrow \sup_{i \leq k} \frac{\sum_{j=0}^{i-1}(X_j - \alpha)}{i} > 0 \Leftrightarrow \sup_{i \leq k} \sum_{j=0}^{i-1}(X_j - \alpha) > 0.$$

Let the last condition be $M_k > 0$. Then the above says $D_k - \alpha > 0$ is equivalent to $M_k > 0$. Therefore by lemma 6.2.2,

$$0 \leq E\left[(X - \alpha)1_{(M_k > 0)}\right].$$

Thus,

$$\alpha P(D_k > \alpha) \leq EX1_{(M_k > 0)} \leq E|X|.$$

$\square$