Math 832, Assignment #1: Solutions

1. The proof is by induction on n. When $n = 0$, the formula gives $\mu(0)$, so that case holds. For all n we have

$$\begin{aligned}
P_\mu(X_{n+1} = 0) &= P_\mu(X_n = 0,\ X_{n+1} = 0) + P_\mu(X_n = 1,\ X_{n+1} = 0) \\
&= P_\mu(X_n = 0)p(0,0) + P_\mu(X_n = 1)p(1,0) \\
&\quad \text{(by the Markov property)} \\
&= p(1,0) - P_\mu(X_n = 0)(p(0,0) - p(1,0)) \\
&\quad \text{(since } P_\mu(X_n = 1) = 1 - P_\mu(X_n = 0)\text{)} \\
&= \beta + \left(\frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^n \left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\}\right)(1-\alpha-\beta) \\
&\quad \text{(if we assume the formula holds for } n\text{)} \\
&= \frac{\beta}{\alpha+\beta} + (1-\alpha-\beta)^{n+1}\left\{\mu(0) - \frac{\beta}{\alpha+\beta}\right\},
\end{aligned}$$

so the formula holds for all n by induction.

2. If $X_n = N$, then $X_{n+1} = N$ with probability 1. If $X_n = k < N$, then X_{n+1} is either k or $k+1$. It is k precisely if ξ_{n+1} belongs to the set $\{\xi_1, ..., \xi_n\}$; since this set has cardinality k, the chance of this event is k/N, regardless of the values of $X_0, X_1, ..., X_{n-1}$. Note that this formula also applies when $k = N$. Hence

$$p(i,j) = \begin{cases} i/N & \text{if } j = i, \\ 1 - i/N & \text{if } j = i+1, \\ 0 & \text{otherwise.} \end{cases}$$

3. (i) Let $i_1, ..., i_n \in \{-1, 1\}$, $a = |\{m \le n : i_m = 1\}|$, and $b = n - a$.

$$P(X_1 = i_1, ..., X_n = i_n) = \int_0^1 \theta^a (1-\theta)^b \, d\theta$$

$$P(X_1 = i_1, ..., X_n = i_n, X_{n+1} = 1) = \int_0^1 \theta^{a+1}(1-\theta)^b \, d\theta$$

Now $\int_0^1 x^m (1-x)^k \, dx = m!k!/(m+k+1)!$ (proof: use repeated integration by parts, or just look it up in a table of integrals), so

$$P(X_{n+1} = 1 | X_1 = i_1, ..., X_n = i_n) = \frac{(a+1)!b!/(a+b+2)!}{a!b!/(a+b+1)!} = \frac{a+1}{a+b+2} = \frac{\frac{1}{2}(n+S_n)+1}{n+2}$$

(ii) Note that $S_{n+1} = S_n + X_{n+1}$ and $\sigma(S_1, S_2, ..., S_n) = \sigma(X_1, X_2, ..., X_n)$, so $E(S_{n+1}|S_1, S_2, ..., S_n) = S_n + E(X_{n+1}|X_1, X_2, ..., X_n)$. Since both terms depend only on S_n, the sequence $S_1, S_2, ...$ is a Markov chain. But since the sum varies with n, the Markov chain is temporally inhomogeneous.

4. If $x \notin A \cup B$ then $1_{(\tau_A < \tau_B)} \circ \theta_1 = 1_{(\tau_A < \tau_B)}$. Taking expected value we have

$$\begin{aligned} h(x) &= P_x(\tau_A < \tau_B) \\ &= E_x(1_{(\tau_b < \tau_B)} \circ \theta_1) \\ &= E_x(E_x(1_{(\tau_b < \tau_B)} \circ \theta_1 | \mathcal{F}_1)) \\ &= E_x h(X_1) \\ &= \sum_y p(x,y) h(y). \end{aligned}$$

In the particular case specified in the problem, we have $h(1) = 1$, $h(2) = \frac{1}{2}(h(1) + h(3))$, $h(3) = \frac{1}{2}(h(2) + h(4))$, $h(4) = \frac{1}{2}(h(3) + h(5))$, and $h(5) = 1$. Therefore $h(x)$ is a linear function of x, with $h(i) = (5 - i)/4$.

5. If $x \notin A$, $\tau_A \circ \theta_1 = \tau_A - 1$. Taking expected value gives

$$\begin{aligned} g(x) - 1 &= E_x(\tau_A - 1) \\ &= E_x(\tau_A \circ \theta_1) \\ &= E_x(E_x(\tau_A \circ \theta_1 | \mathcal{F}_1)) \\ &= E_x g(X_1) \\ &= \sum_y p(x,y) g(y). \end{aligned}$$

For the particular case, we have $h(1) = 0$, $h(2) = 1 + \frac{1}{2}(h(1) + h(3))$, $h(3) = 1 + \frac{1}{2}(h(2) + h(4))$, $h(4) = 1 + \frac{1}{2}(h(3) + h(5))$, and $h(5) = 1 + h(4)$. Solving this linear system we get $h(1) = 0$, $h(2) = 7$, $h(3) = 12$, $h(4) = 15$, and $h(5) = 16$.

6. $X'_1, ..., X'_n$ is a Markov chain. To see this, note that for all $k < n$ and for all $i_1, ..., i_{k+1}$ we have

$$\begin{aligned} & P(X'_{k+1} = i_{k+1} | X'_1 = i_1, ..., X'_k = i_k) \\ = & \frac{P(X_1 = i_1, ..., X_k = i_k, X_{k+1} = i_{k+1}, X_n \in A)}{P(X_1 = i_1, ..., X_k = i_k, X_n \in A)} \end{aligned}$$

$$
\begin{aligned}
&= \frac{P(X_1 = i_1, ..., X_k = i_k)p(i_k, i_{k+1})p^{n-k-1}(i_{k+1}, A)}{P(X_1 = i_1, ..., X_k = i_k)p^{n-k}(i_k, A)} \\
&= \frac{p(i_k, i_{k+1})p^{n-k-1}(i_{k+1}, A)}{p^{n-k}(i_k, A)}
\end{aligned}
$$

which is independent of $i_1, ..., i_{k-1}$.

However, the conditioned chain need not be temporally homogeneous, even if the original chain was. To see this, consider the case in which the X_n's are i.i.d. and uniform on $\{1, 2\}$, and let $A = \{1\}$. Then for all $1 < k < n$ we have $P(X_k = 1|X_{k-1} = 1) = \frac{1}{2}$, but $P(X_n = 1|X_{n-1} = 1) = 1$. Less trivially, we could take random walk on $\mathbf{Z}$ started at location 0 at time 0 and run for n steps (with n even), and take $A = \{0\}$. With n large we have $P(X'_2 = 0|X'_1 = 1) \approx \frac{1}{2}$ but $P(X'_n = 0|X'_{n-1} = 1) = 1$.

7. No. Since $P_x(N(y) = k) = \rho_{xy}(1 - \rho_{yy})\rho_{yy}^{k-1}$ we have $E_x([N(y)]^2) = \rho_{xy}(1 - \rho_{yy})(1 + \rho_{yy})/(1 - \rho_{yy})^3$, so the second moment is finite.

8. (i)

$$
\begin{aligned}
u_n &= P_a(X_n = a) \\
&= \sum_{m=1}^{n} P_a(T_a = m, X_n = a) \\
&= \sum_{m=1}^{n} P_a(T_a = m)P(X_n = a|T_a = m) \\
&= \sum_{m=1}^{n} f_m u_{n-m}
\end{aligned}
$$

where the equation $P(X_n = a|T_a = m) = P_a(X_{n-m} = a) = u_{n-m}$ is a consequence of the strong Markov property. (Compare with the formula in exercise 2.2.)

(ii) Using (i) and Fubini's theorem, and then changing variables in the inner sum, we get

$$
\begin{aligned}
u(s) - 1 &= \sum_{n=1}^{\infty} u_n s^n \\
&= \sum_{n=1}^{\infty} \sum_{m=1}^{n} f_m u_{n-m} s^n
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{m=1}^{\infty} \sum_{n=m}^{\infty} f_m s^m u_{n-m} s^{n-m} \\
&= \sum_{m=1}^{\infty} f_m s^m \sum_{k=0} u_k s^k \\
&= f(s)u(s)
\end{aligned}
$$

9. (i) Differentiating $h(x) = (1-x)^{-1/2}$ gives

$$
\begin{aligned}
h'(x) &= \frac{1}{2} \cdot (1-x)^{-3/2} \\
h''(x) &= \frac{1}{2} \cdot \frac{3}{2} \cdot (1-x)^{-5/2} \\
&\vdots \\
h^{(m)}(x) &= \frac{(2m)!}{m! 4^m} (1-x)^{-(2m+1)/2}
\end{aligned}
$$

Hence the coefficient of x^m in $(1-x)^{-1/2}$ is $\frac{(2m)!}{(m!)^2 4^m}$, whence the coefficient of s^{2m} in $(1-4pqs^2)^{-1/2}$ is $\frac{(2m)!}{(m!)^2} p^m q^m$, which is also the coefficient of s^{2m} in $u(s)$. Hence $u(s) = (1-4pqs^2)^{-1/2}$, so using Exercise 3.2 we have $f(s) = 1 - 1/u(s) = 1 - (1-4pqs^2)^{1/2}$. Note: The series expansion is convergent as long as $4pqs^2$ is strictly less than 1 in absolute value.

(ii) Setting $s = 1$, we have $P_0(T_0 < \infty) = 1 - (1-4pq)^{1/2} = 1 - |1-2p|$, valid as long as $pq < \frac{1}{4}$, i.e., as long as $p \neq \frac{1}{2}$. (The formula is actually valid for $p = \frac{1}{2}$ as well, but the method of proof does not apply directly, since the series expansion is divergent.)

10. Not necessarily. As a degenerate case, let $X_1, X_2, \ldots$ be the Markov chain on $\{1,2\}$ with $p(1,2) = p(2,1) = 1$, and let $Y_1, Y_2, \ldots$ be the same chain. In the independent product we have $p((1,1),(2,2)) = p((2,2),(1,1)) = p((1,2),(2,1)) = p((2,1),(1,2)) = 1$. Since $p^n((1,1),(1,2)) = 0$ for all n, this new 4-state Markov chain is not irreducible.

11. (a) We can model this with a 5-state Markov chain, with states numbered 0 through 4; the system is in state $\geq i$ ($i =$ 0,1,2, or 3) when the last i tosses coincide with the first i symbols in the sequence THTH and when $i+1$ does not have this property. The system is in state 4 once the sequence THTH has occurred. Then the possible transitions are as follows: from state 0 to

state 0 or state 1; from state 1 to state 2 or state 1; from state 2 to state 0 or state 3; from state 3 to state 4 or state 1; and from state 4 to itself. If we let a, b, c, and d represent the expected time until one hits state 4 starting from states 0, 1, 2, and 3 respectively, then the formula in exercise 2.10 gives us the linear relations $a = 1 + \frac{1}{2}(a + b)$, $b = 1 + \frac{1}{2}(b + c)$, $c = 1 + \frac{1}{2}(a + d)$, and $d = 1 + \frac{1}{2}(b + 0)$, which we can solve to yield $a = 20$, $b = 18$, $c = 16$, and $d = 10$. Thus the expected time until THTH occurs is 20.

Or: Consider a 16-state Markov chain whose states are HHHH, HHHT, ..., TTTT and whose transition probabilities $p(ijkl, jklm)$ are all equal to 1/2 (and all of whose transitions probabilities not of this form are equal to zero). Think of the state of this chain as representing the outcomes of the four most recent tosses of the coin. In the stationary state, each state has measure 1/16. If we start the chain in state THTH, the expected time until the first return to THTH is 16. There are two ways in which this could happen: we could return to TH in 2 steps (with probability $\frac{1}{4}$) or it could take longer (with probability $\frac{3}{4}$). Hence $16 = \frac{1}{4} \times 2 + \frac{3}{4} \times x$, where x is the conditional expectation of the first return time given that the next two tosses did not come out TH. Now consider the return time that we're actually interested in, in which returning to THTH at time 2 "doesn't count"; the expected value of this modified return time can be written as $\frac{1}{4} \times (2 + 16) + \frac{3}{4} \times x$, where the 16 is the extra time we expect it to take before we enter state THTH again. Combining the two equations, we see that the modified return time has expected value $16 + \frac{1}{4} \times 16 = 20$.

(b) As in part (a), we can use a 5-state Markov chain. This time, the possible transitions are as follows: from state 0 to state 1 or state 0; from state 1 to state 1 or state 2; from state 2 to state 3 or state 0; from state 3 to state 4 or state 2; and from state 4 to itself. If we let a, b, c, and d represent the expected time until one hits state 4 starting from states 0, 1, 2, and 3 respectively, then the formula in exercise 2.10 gives us the linear relations $a = 1 + \frac{1}{2}(a + b)$, $b = 1 + \frac{1}{2}(b + c)$, $c = 1 + \frac{1}{2}(a + d)$, and $d = 1 + \frac{1}{2}(c + 0)$, which we can solve to yield $a = 18$, $b = 16$, $c = 14$, and $d = 8$. Thus the expected time until HTHH occurs is 18.

Or: Consider the same 16-state chain as in the second solution to part (a), and imagine starting it in state HTHH. It is possible to return to state HTHH at time 3, if the next three tosses are T, H, and H; this happens with probability $\frac{1}{8}$. Otherwise, the first return to HTHH corresponds to the first occurrence of HTHH that does not involve the fictitious tosses in the start-state of the chain. So, as in part (a), we find that the expected time until the first return to state HTHH at time ≥ 4 is $16 + \frac{1}{8} \times 16 = 18$.

(c) Now we must use a 17-state Markov chain with $4 \times 4 = 16$ transient states, corresponding to ordered pairs of states of the chains considered in parts (a) and (b). However, the states (2,2) and (3,3) can be ignored, since they are not accessible from the starting state (0,0). The formula in exercise 2.6 gives us simultaneous linear relations which we can solve to obtain 9/14 as the probability that absorption in $\{(4,0),(4,1),(4,2),(4,3)\}$ occurs before absorption in $\{(0,4),(1,4),(2,4),(3,4)\}$.

Note: "HTHH occurs sooner than THTH" because the expected time until the HTHH occurs is smaller than the expected time until THTH occurs. But "THTH occurs sooner than HTHH" because the probability that THTH occurs before HTHH is 9/14, which is greater than one-half! This paradox shows yet again the pitfalls of using natural language to describe probabilistic situations.