STAT201A – Sec. 102
Homework #9

Steven Pollack
24112977

#1.

Proof. Let (I, II) be the random vector that represents the color of the ball drawn from box I and box II. So, $(I, II) = (B, W)$ means a blue ball was drawn from I and a white from II. Furthermore, set X_n to be the count of white balls in box I, and note there exist three cases

$$\begin{aligned}
X_n &= X_{n-1} - 1 &&\iff (I, II) = (W, B) \\
X_n &= X_{n-1} &&\iff (I, II) = (B, B) \text{ or } (W, W) \\
X_n &= X_{n-1} + 1 &&\iff (I, II) = (B, W)
\end{aligned}$$

Also, we note that $X_0 \sim \text{HypGeo}(N = 10, G = 4, n = 5)$. Now, it's clear that $P(X_n = i_n \mid X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}) = P(X_n = i_n \mid X_{n-1} = i_{n-1})$ thus $\{X_n\}$ is an irreducible, aperiodic Markov chain with a finite state-space. Hence, our big theorem says that there exists a unique distribution π such that $\pi = \pi\mathbb{P}$. So, let's try calculating $\mathbb{P}$.

First, for $i, j \in \{0, 1, 2, 3, 4\}$, if $j \neq i, i \pm 1$, then $p_{i,j} = 0$. Next, from our analysis above,

$$\begin{aligned}
p_{i,i} &= P((I, II) = (B, B) \text{ or } (W, W) \mid \text{box I has } i \text{ white balls}) &&= \frac{5-i}{5}\frac{1+i}{5} + \frac{i}{5}\frac{4-i}{5} \\
p_{i,i+1} &= P((I, II) = (B, W) \mid \text{box I has } i \text{ white balls}) &&= \frac{5-i}{5}\frac{4-i}{5} \\
p_{i,i-1} &= P((I, II) = (W, B) \mid \text{box I has } i \text{ white balls}) &&= \frac{i}{5}\frac{1+i}{5}
\end{aligned}$$

Which yields

$$\mathbb{P} = \frac{1}{25}\begin{pmatrix} 5 & 20 & 0 & 0 & 0 \\ 2 & 11 & 12 & 0 & 0 \\ 0 & 6 & 13 & 6 & 0 \\ 0 & 0 & 12 & 11 & 2 \\ 0 & 0 & 0 & 20 & 5 \end{pmatrix}$$

and if we let $\pi_i = P(X_0 = i) = \binom{4}{i}\binom{6}{5-i}/\binom{10}{5}$, then we see that $\pi = (\pi_0, \pi_1, \pi_2, \pi_3, \pi_4)$ solves the balance equations, $\pi\mathbb{P} = \pi$.

Hence, π is our stationary distribution and thus $\pi_0 = 1/42$ is the amount of time we expect box I to be empty. $\square$

#2.

Proof. 1. Let T_j be the random variable that counts the number of steps a chain $\{X_n\}$

takes to reach state j, and write $m_{ij} = E(T_j \mid X_0 = i)$. Then,

$$\begin{aligned}
E(T_j \mid X_0 = i) &= \sum_{k=1}^{\infty} kP(T_j = k \mid X_0 = i) \\
&= P(T_j = 1 \mid X_0 = i) + 2P(T_j = 2 \mid X_0 = i) + 3P(T_j = 3 \mid X_0 = i) + \cdots \\
&= p_{ij} + 2\sum_{k_1 \neq j} p_{ik_1}p_{k_1 j} + 3\sum_{k_1,k_2 \neq j} p_{ik_1}p_{k_1k_2}p_{k_2 j} + \cdots \\
&= \left(p_{ij} + \sum_{k_1 \neq j} p_{ik_1}p_{k_1 j} + \sum_{k_1,k_2 \neq j} p_{ik_1}p_{k_1k_2}p_{k_2 j} + \cdots \right) \\
&+ \left(\sum_{k_1 \neq j} p_{ik_1}p_{k_1 j} + 2\sum_{k_1,k_2 \neq j} p_{ik_1}p_{k_1k_2}p_{k_2 j} + 3\sum_{k_1,k_2,k_3 \neq j} p_{ik_1}p_{k_1k_2}p_{k_2k_3}p_{k_3 j} + \cdots \right) \\
&= S_1 + S_2
\end{aligned}$$

However, notice that $S_1 = \sum_{k=1}^{\infty} P(T_j = k \mid X_0 = i)$, hence $S_1 = 1$. Furthermore,

$$\begin{aligned}
S_2 &= \sum_{k_1 \neq j} p_{ik_1} \left(p_{k_1 j} + \sum_{k_2 \neq j} 2p_{k_1k_2}p_{k_2 j} + \sum_{k_2,k_3 \neq j} 3p_{k_1k_2}p_{k_2k_3}p_{k_3 j} + \cdots \right) \\
&= \sum_{k_1 \neq j} p_{ik_1} E(T_j \mid X_0 = k_1) \\
&= \sum_{k_1 \neq j} p_{ik_1} m_{k_1 j}
\end{aligned}$$

Thus,

$$m_{ij} = 1 + \sum_{k \neq j} p_{ik} m_{kj}$$

2. Using the result in part 1.

$$\begin{aligned}
\sum_i \pi_i m_{ij} &= \sum_i \left(\pi_i + \sum_{k \neq j} \pi_i p_{ik} m_{kj} \right) \\
\iff E_{X_0}\left[m_{ij}\right] &= E_{X_0}\left[1 + \sum_{k \neq j} p_{ik} m_{kj} \right] \\
\iff E_{X_0}\left[m_{ij}\right] &= 1 + \sum_{k \neq j} E_{X_0}\left[p_{ik} m_{kj}\right] \\
\iff E_{X_0}\left[m_{ij}\right] &= 1 + \sum_{k \neq j} m_{kj} E_{X_0}\left[p_{ik}\right] \\
\iff E_{X_0}\left[m_{ij}\right] &= 1 + \sum_{k \neq j} m_{kj} \pi_k
\end{aligned}$$

Where the second to last equality stems from the fact that $\sum_k \pi_i p_{ik} = \pi_k$ since π is the stationary distribution of $\{X_n\}$. Finally, we observe that

$$E_{X_0}[m_{ij}] = \sum_{k=1}^{n} m_{kj}\pi_k = m_{jj}\pi_j + \sum_{k\neq j} m_{kj}\pi_k$$

Thus,

$$E_{X_0}\left[m_{ij}\right] = 1 + \sum_{k\neq j} m_{kj}\pi_k \iff m_{jj}\pi_j = 1 \iff m_{jj} = \frac{1}{p_j}$$

□

#3.

Proof. Set $D_1 = Y - \hat{Y}$, and $D_2 = X_2 - \hat{X}$, and by the hint, we note that $\operatorname{cov}(D_1, X_i) = 0$ for $i = 1, 2$ and $\operatorname{cov}(D_2, X_1) = 0$. Also, note that we may write $\hat{X} = X_2 - D_2$, hence

$$\begin{aligned}
\tilde{Y} &= \mu_Y + c_{*1}(X_1 - \mu_{X_1}) + c_{*2}(\hat{X} - \mu_{X_2}) \\
&= \mu_Y + c_{*1}(X_1 - \mu_{X_1}) + c_{*2}(X_2 - \mu_{X_2} - D_2) \\
&= \hat{Y} - c_{*2}D_2
\end{aligned}$$

Thus, using our first observations

$$\begin{aligned}
\operatorname{cov}(Y - \tilde{Y}, X_1) &= \operatorname{cov}((Y - \hat{Y}) + c_{*2}D_2, X_1) \\
&= \operatorname{cov}(D_1, X_1) + c_{*2}\operatorname{cov}(D_2, X_1) \\
&= 0
\end{aligned}$$

which shows that $\tilde{Y}$ is the best linear predictor of Y based on X_1. □