6.041/6.431 Fall 2011 Final Exam Solutions

Problem 1: (10 points)

(a) **(5 points)** Let $Y = X_1^4$.

$$
\begin{aligned}
F_Y(y) &= \mathbf{P}(Y \le y) \\
&= \mathbf{P}(X_1^4 \le y) \\
&= \mathbf{P}(-y^{\frac{1}{4}} \le X_1 \le y^{\frac{1}{4}}) \\
&= \mathbf{P}(X_1 \le y^{\frac{1}{4}}) \qquad \text{for } y \ge 0 \\
&= \begin{cases} 1 - e^{-8y^{\frac{1}{4}}}, & y \ge 0 \\ 0, & y < 0. \end{cases}
\end{aligned}
$$

Differentiating the CDF of Y, we get the PDF of Y:

$$
f_Y(y) = \begin{cases} 2y^{-\frac{3}{4}} e^{-8y^{\frac{1}{4}}}, & y \ge 0 \\ 0, & y < 0. \end{cases}
$$

(b) **(5 points)** Let $Z = X_1 + X_2 + X_3$, which is recognized as Erlang of order 3 with $\lambda = 8$. Therefore,

$$
\begin{aligned}
f_Z(z) &= \frac{8^3 z^2 e^{-8z}}{2!} \\
&= \begin{cases} 256 z^2 e^{-8z}, & z \ge 0 \\ 0, & z < 0. \end{cases}
\end{aligned}
$$

Problem 2: (25 points)

(a) **(4 points)** The only path that leads to $X_2 = 3$ is $1 \to 2 \to 3$. Therefore $\mathbf{P}(X_2 = 3) = \frac{1}{4} \cdot \frac{3}{8} = \frac{3}{32}$.

(b) **(4 points)** Since the chain starts at 1, the first change of state will be from state 1 to 2. In order for second change of state to lead to state 3, state 2 must lead to state 3 when it changes state. State 2 is equally likely to change state to 1 or 3. Therefore, the probability of interest is $\frac{1}{2}$.

(c) **(5 points)**

$$
\begin{aligned}
\mathbf{P}(X_{1000} = 2 \mid X_{1000} = X_{1001}) &= \frac{\mathbf{P}(X_{1000} = X_{1001} = 2)}{\mathbf{P}(X_{1000} = X_{1001})} \\
&= \frac{\mathbf{P}(X_{1000} = 2) p_{22}}{\mathbf{P}(X_{1000} = 1) p_{11} + \mathbf{P}(X_{1000} = 2) p_{22} + \mathbf{P}(X_{1000} = 3) p_{33}} \\
&\approx \frac{\pi_2 \cdot p_{22}}{\pi_1 \cdot p_{11} + \pi_2 \cdot p_{22} + \pi_3 \cdot p_{33}},
\end{aligned}
$$

where the $\pi_i' s$ are the steady-state probabilities.

To compute these probabilities, we set up the balance equations noting that the Markov chain represents a birth-death process and noting the symmetry of states 1 and 3.

$$\begin{aligned}
&\pi_1 \cdot \frac{1}{4} = \pi_2 \cdot \frac{3}{8} \\
&\pi_1 = \pi_3 \\
&\pi_1 + \pi_2 + \pi_3 = 1.
\end{aligned}$$

Solving this system of equations yields

$$\begin{aligned}
\pi_1 &= \frac{3}{8}, \\
\pi_2 &= \frac{1}{4}, \\
\pi_3 &= \frac{3}{8}.
\end{aligned}$$

Therefore,

$$\mathbf{P}(X_{1000} = 2 \mid X_{1000} = X_{1001}) = \frac{1}{10}.$$

(d) **(6 points)** We set up the system of equations to find the mean first passage time to reach state 3 starting from 1.

$$\begin{aligned}
t_1 &= 1 + \frac{3}{4}t_1 + \frac{1}{4}t_2 \\
t_2 &= 1 + \frac{3}{8}t_1 + \frac{1}{4}t_2
\end{aligned}$$

Solving this system of equation yields

$$\mathbf{E}[T] = t_1 = \frac{32}{3}.$$

(e) **(6 points)** We will condition on X_1 using the total probability theorem to compute $\mathbf{E}[S]$.

$$\begin{aligned}
\mathbf{E}[S] &= (1 + \mathbf{E}[S \mid X = 1])\mathbf{P}(X = 1) + (1 + \mathbf{E}[S \mid X = 2])\mathbf{P}(X = 2) \\
&\quad + (1 + \mathbf{E}[S \mid X = 3])\mathbf{P}(X = 3) \\
&= (1 + \mathbf{E}[T]) \cdot \frac{3}{8} + (1 + \mathbf{E}[S]) \cdot \frac{1}{4} + (1 + \mathbf{E}[T]) \cdot \frac{3}{8} \\
&= \frac{4}{3}\left(1 + \frac{3}{4}\mathbf{E}[T]\right) \\
&= 12.
\end{aligned}$$

An alternative solution is the following. Let G be the number of transitions until the chain leaves state 2 initially, which is a geometric random variable with $p = 3/4$. Once the chain leaves state 2, it is either in 1 or 3. The time it then takes to reach state 3 from 1 or state 1 from 3 is T. So,

$$\mathbf{E}[S] = \mathbf{E}[G] + \mathbf{E}[T] = \frac{4}{3} + \frac{32}{3} = 12.$$

Problem 3: (34 points)

(a) **(4 points)**

This is a Poisson probability with $k = 3$, $\lambda = 3$ and $\tau = 0.5$.

$$P(3, 0.5) = e^{-\frac{3}{2}} \frac{(\frac{3}{2})^3}{3!} = \frac{9}{16} e^{-\frac{3}{2}}.$$

(b) **(4 points)** The kth student will stay with the professor until another student arrives. This can be considered a merged process of the two Poisson processes with $\lambda = 8$. Therefore, the expected length of time is $\frac{1}{\lambda} = \frac{1}{8}$ or 7.5 minutes.

(c) **(5 points)** Given an arrival, the probability that the student will be a graduate student is $\frac{5}{8}$. The expected number of new arrivals is then the expected value of geometric random variable with $p = \frac{5}{8}$, which is $\frac{1}{p} = \frac{8}{5}$.

(d) **(5 points)** This is a binomial probability with $n = 7$, $k = 5$ and $p = \frac{3}{8}$. The probability is then $\binom{7}{5} \left(\frac{3}{8}\right)^5 \left(\frac{5}{8}\right)^2$.

(e) **(6 points)** There are two methods to approximate the probability. Let M be the event that the professor meets more than 205 students in 24 hours.

Method 1 We can approximate the Poisson probability $P(k, \tau)$ using the CLT and the De Moivre-Laplace approximation. Let N_{24} be the number of students the professor meets in 24 hours so that $\mathbf{E}[N_{24}] = \text{var}(N_{24}) = 192$.

$$\begin{aligned}
\mathbf{P}(M) &= \mathbf{P}(N_{24} > 205) \\
&\approx \mathbf{P}\left(\frac{N_{24} - 192}{\sqrt{192}} > \frac{205.5 - 192}{\sqrt{192}}\right) \\
&\approx 1 - \Phi\left(\frac{13.5}{8\sqrt{3}}\right) \\
&\approx 1 - \Phi\left(\frac{13.5}{8\sqrt{3}}\right) \\
&\approx .1660.
\end{aligned}$$

Method 2 Let Y_k be the length of duration of student k in the professor's office hours, which is exponential with $\lambda = 8$. Also, let $S_k = Y_1 + Y_2 + \cdots + Y_k$ so that $\mathbf{E}[S_k] = k/8$ and $\text{var}(S_k) = k/64$. The probability of interest is $\mathbf{P}(S_{206} \leq 24)$. Using the CLT and the De Moivre-Laplace approximation,

$$
\begin{aligned}
\mathbf{P}(M) &= \mathbf{P}(S_{206} \leq 24) \\
&= \mathbf{P}\left(\frac{S_{206} - \mathbf{E}[S_K]}{\sqrt{\mathrm{var}(S_{206})}} \leq \frac{24 - \mathbf{E}[S_{206}]}{\sqrt{\mathrm{var}(S_{206})}}\right) \\
&\approx \mathbf{P}\left(\frac{S_{206} - 206/8}{\sqrt{206}/8} \leq \frac{24.5 - 206/8}{\sqrt{206}/8}\right) \\
&\approx \Phi\left(\frac{-1.75}{\sqrt{206}/8}\right) \\
&\approx \Phi\left(-\frac{14}{\sqrt{206}}\right) = \Phi\left(-\frac{7}{\sqrt{103}}\right) \\
&\approx \Phi(-0.69) \\
&\approx 0.2451.
\end{aligned}
$$

(f) **(5 points)** We can think of the arrival process as follows. We have a merged Poisson process at rate $\lambda_U + \lambda + G$. Whenever there is an arrival, we flip a coin, and with probability $\lambda_U/(\lambda_U + \lambda_G)$, the next student is an undergraduate. The outcome of this coin flip has no bearing on the amount of time until the arrival in the merged process occurred. Hence, the answer is $1/(\lambda_U + \lambda_G) = \frac{1}{8}$.

A more formal approach involves the use of Bayes' rule. Let T be the length of time until the next student arrives starting from the time at which the department head arrives. Let A be the random variable that takes on the value 1 if the next student is an undergraduate and 0 otherwise.

Using Bayes' rule,

$$
f_{T|A}(t \mid 1) = \frac{p_{A|T}(1 \mid t) f_T(t)}{p_A(1)}.
$$

The expression $p_{A|T}(1 \mid t) = p_A(1)$ as the time of the arrival is independent of the type of arrival in the merged process. Therefore, $f_{T|A}(t \mid 1) = f_T(t)$, which is the interarrival time of the merged process. The expected value of the distribution is $1/(\lambda_U + \lambda_G)$ as expected.

(g) **(5 points)** Let B be the time until the department heads sees both an undergraduate and a graduate student.

Let $U = 1$ if there is an undergraduate student in the professor's office when the department heads arrives and let $U = 0$ if it is a graduate student. Therefore, $\mathbf{P}(U = 1) = 3/8$ and $\mathbf{P}(U = 0) = 5/8$.

$$
\begin{aligned}
\mathbf{E}[B] &= \mathbf{E}[B \mid U = 1]\mathbf{P}(U = 1) + \mathbf{E}[B \mid U = 0]\mathbf{P}(U = 0) \\
&= \frac{1}{5} \cdot \frac{3}{8} + \frac{1}{3} \cdot \frac{5}{8} = \frac{17}{60}.
\end{aligned}
$$

The second equality holds as $\mathbf{E}[B \mid U = 1]$ is the expected value of an exponential random variable with $\lambda = \lambda_G = 5$ as the department head must wait until a graduate student arrives. The second term follows from the same argument.

Problem 4: (18 points)

(a) **(5 points)** To find $\hat{\theta}$, we find $f_{X_1}(x_1;\theta)$. Given $X_1 = \theta + W_1$, X_1 is a shifted exponential. Therefore,

$$f_{X_1}(x_1;\theta) = \begin{cases} e^{-(x_1-\theta)}, & x_1 \geq \theta, \\ 0, & x_1 < \theta. \end{cases}$$

This quantity is maximized at $\hat{\theta} = x_1$.

(b) **(5 points)** To find $\hat{\theta}$, we find $f_X(x;\theta)$.

$$\begin{aligned} f_X(x;\theta) &= \prod_{i=1}^{n} f_{X_i}(x_i;\theta) \\ &= \begin{cases} \displaystyle\prod_{i=1}^{n} e^{-(x_i-\theta)}, & x_i \geq \theta \quad \forall i, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

This quantity is maximized at $\hat{\theta} = \min_i\{x_i\}$. Any value greater than this would yield zero.

(c) **(8 points)** We wish to find c such that $\mathbf{P}(\hat{\Theta} - c \leq \theta \leq \hat{\Theta}) \geq 0.95$. Since each of the W_i's is nonnegative, θ must less than or equal to $\hat{\Theta} = \min\limits_i\{X_i\}$. Since, $\mathbf{P}(\theta \leq \hat{\Theta}) = 1$, we need $\mathbf{P}(\hat{\Theta} - c \leq \theta) \geq 0.95$.

$$\begin{aligned} \mathbf{P}(\hat{\Theta} - c \leq \theta) &= \mathbf{P}(\min_i\{X_i\} \leq \theta + c) \\ &= 1 - \mathbf{P}(\min_i\{X_i\} \geq \theta + c) \\ &= 1 - \prod_{i=1}^{1000} \mathbf{P}(X_i \geq \theta + c) \\ &= 1 - e^{-1000c}. \end{aligned}$$

The desired expression is $1 - e^{-1000c} \geq 0.95$ or

$$c \geq \frac{-\log(0.05)}{1000} = \frac{\log(20)}{1000}.$$

Problem 5: (13 points)

(a) **(6 points)** Let I be the instrument used to perform the measurement and $\mathbf{P}(I = 1) = \mathbf{P}(I = 2) = 1/2$.

Also note that,

$$\mathbf{E}[\Theta^2] = \text{var}(\Theta) + \mu_\Theta^2 = 3,$$
$$\mathbf{E}[W^2] = \text{var}(W) + \mu_W^2 = 14.$$

$$\begin{aligned}
\mathbf{E}[X] &= \mathbf{E}[X \mid I = 1]\mathbf{P}(I = 1) + \mathbf{E}[X \mid I = 2]\mathbf{P}(I = 2) \\
&= (\mathbf{E}[\Theta + W] + \mathbf{E}[2\Theta + 3W]) \cdot \frac{1}{2}. \\
&= (3\mu_\Theta + 4\mu_W) \cdot \frac{1}{2} \\
&= \frac{15}{2}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[X^2] &= \mathbf{E}[X^2 \mid I = 1]\mathbf{P}(I = 1) + \mathbf{E}[X^2 \mid I = 2]\mathbf{P}(I = 2) \\
&= (\mathbf{E}[(\Theta + W)^2] + \mathbf{E}[(2\Theta + 3W)^2]) \cdot \frac{1}{2}. \\
&= (5\mathbf{E}[\Theta^2] + 14\mu_\Theta\mu_W + 10\mathbf{E}[W^2]) \cdot \frac{1}{2} \\
&= \frac{197}{2}.
\end{aligned}$$

(b) **(7 points)** The linear estimator of Θ given X is

$$\hat{\Theta} = \mathbf{E}[\Theta] + \frac{\text{cov}(X, \Theta)}{\text{var}(X)}(X - \mathbf{E}[X]).$$

The covariance term is computed by first calculating $\mathbf{E}[X\Theta]$

$$\begin{aligned}
\mathbf{E}[X\Theta] &= \mathbf{E}[X\Theta \mid I = 1]\mathbf{P}(I = 1) + \mathbf{E}[X\Theta \mid I = 2]\mathbf{P}(I = 2) \\
&= \mathbf{E}[(\Theta + W)\Theta]\mathbf{P}(I = 1) + \mathbf{E}[(2\Theta + 3W)\Theta]\mathbf{P}(I = 2) \\
&= (3\mathbf{E}[\Theta^2] + 4\mu_\Theta\mu_W) \cdot \frac{1}{2} \\
&= \frac{21}{2}.
\end{aligned}$$

$$\begin{aligned}
\hat{\Theta} &= \mu_\Theta + \frac{\mathbf{E}[X\Theta] - \mathbf{E}[X]\mu_\Theta}{\mathbf{E}[X^2] - (E[X])^2}(X - \mathbf{E}[X]) \\
&= 1 + \frac{\frac{21}{2} - \frac{15}{2} \cdot 1}{\frac{197}{2} - (\frac{15}{2})^2}(X - \frac{15}{2}) \\
&= \frac{79}{169} + \frac{12}{169}X.
\end{aligned}$$