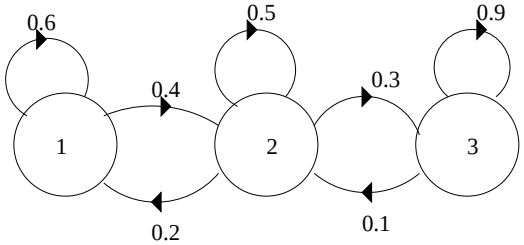
6.041/6.431 Fall 2010 Final Exam Solutions
Wednesday, December 15, 9:00AM - 12:00noon.

Problem 1. (32 points) Consider a Markov chain $\{X_n; n = 0, 1, \ldots\}$, specified by the following transition diagram.



1. **(4 points)** Given that the chain starts with $X_0 = 1$, find the probability that $X_2 = 2$.

   Solution: The two-step transition probability is:

$$\begin{aligned} r_{12}(2) &= p_{11} \cdot p_{12} + p_{12} \cdot p_{22} \\ &= 0.6 \cdot 0.4 + 0.4 \cdot 0.5 \\ &= 0.44. \end{aligned}$$

2. **(4 points)** Find the steady-state probabilities π_1, π_2, π_3 of the different states.

   Solution: We set up the balance equations of a birth-death process and the normalization equation as such:

$$\begin{aligned} \pi_1 p_{12} &= \pi_2 p_{21} \\ \pi_2 p_{23} &= \pi_3 p_{32} \\ \pi_1 + \pi_2 + \pi_3 &= 1. \end{aligned}$$

   Solving the system of equations yields the following steady state probabilities:

$$\begin{aligned} \pi_1 &= 1/9 \\ \pi_2 &= 2/9 \\ \pi_3 &= 6/9. \end{aligned}$$

   In case you did not do part 2 correctly, in **all** *subsequent parts of this problem you can just use the symbols π_i: you do not need to plug in actual numbers.*

3. **(4 points)** Let $Y_n = X_n - X_{n-1}$. Thus, $Y_n = 1$ indicates that the nth transition was to the right, $Y_n = 0$ indicates it was a self-transition, and $Y_n = -1$ indicates it was a transition to the left. Find $\lim_{n\to\infty} \mathbf{P}(Y_n = 1)$.

   Solution: Using the total probability theorem and steady state probabilities,

$$\begin{aligned} \lim_{n\to\infty} \mathbf{P}(Y_n = 1) &= \sum_{i=1}^{3} \pi_i \cdot \mathbf{P}(Y_n = 1 \mid X_{n-1} = i) \\ &= \pi_1 p_{12} + \pi_2 p_{23} \\ &= 1/9. \end{aligned}$$

4. **(4 points)** Is the sequence Y_n a Markov chain? Justify your answer.

   Solution: No. Assume the Markov process is in steady state. To satisfy the Markov property,

   $$\mathbf{P}(Y_n = 1 \mid Y_{n-1} = 1, Y_{n-2} = 1) = \mathbf{P}(Y_n = 1 \mid Y_{n-1} = 1).$$

   For large n,

   $$\mathbf{P}(Y_n = 1 \mid Y_{n-1} = 1, Y_{n-2} = 1) = 0,$$

   since it is not possible to move upwards 3 times in a row. However in steady state,

   $$\begin{aligned} \mathbf{P}(Y_n = 1 \mid Y_{n-1} = 1) &= \frac{\mathbf{P}(\{Y_n = 1\} \cap \{Y_{n-1} = 1\})}{\mathbf{P}(Y_{n-1} = 1)} \\ &= \frac{\pi_1 p_{12} p_{23}}{\pi_1 p_{12} + \pi_2 p_{23}} \\ &\neq 0. \end{aligned}$$

   Therefore, the sequence Y_n is not a Markov chain.

5. **(4 points)** Given that the nth transition was a transition to the right ($Y_n = 1$), find the probability that the previous state was state 1. (You can assume that n is large.)

   Solution: Using Bayes' Rule,

   $$\begin{aligned} \mathbf{P}(X_{n-1} = 1 \mid Y_n = 1) &= \frac{\mathbf{P}(X_{n-1} = 1)\mathbf{P}(Y_n = 1 \mid X_{n-1} = 1)}{\sum_{i=1}^{3} \mathbf{P}(X_{n-1} = i)\mathbf{P}(Y_n = 1 \mid X_{n-1} = i)} \\ &= \frac{\pi_1 p_{12}}{\pi_1 p_{12} + \pi_2 p_{23}} \\ &= 2/5. \end{aligned}$$

6. **(4 points)** Suppose that $X_0 = 1$. Let T be defined as the first *positive time* at which the state is again equal to 1. Show how to find $\mathbf{E}[T]$. (It is enough to write down whatever equation(s) needs to be solved; you do not have to actually solve it/them or to produce a numerical answer.)

   Solution: In order to find the the mean recurrence time of state 1, the mean first passage times to state 1 are first calculated by solving the following system of equations:

   $$\begin{aligned} t_2 &= 1 + p_{22} t_2 + p_{23} t_3 \\ t_3 &= 1 + p_{32} t_2 + p_{33} t_3. \end{aligned}$$

   The mean recurrence time of state 1 is then $t_1^* = 1 + p_{12} t_2$.

   Solving the system of equations yields $t_2 = 20$ and $t_3 = 30$ and $t_1^* = 9$.

7. **(4 points)** Does the sequence $X_1, X_2, X_3, \ldots$ converge in probability? If yes, to what? If not, just say "no" without explanation.

   Solution: No.

8. **(4 points)** Let $Z_n = \max\{X_1, \ldots, X_n\}$. Does the sequence $Z_1, Z_2, Z_3, \ldots$ converge in probability? If yes, to what? If not, just say "no" without explanation.

   Solution: Yes. The sequence converges to 3 in probability.

   For the original markov chain, states $\{1, 2, 3\}$ form one single recurrent class. Therefore, the Markov process will eventually visit each state with probability 1. In this case, the sequence Z_n will, with probability 1, converge to 3 once X_n visits 3 for the first time.

Problem 2. (68 points) Alice shows up at an Athena* cluster at time zero and spends her time exclusively in typing emails. The times that her emails are sent are a Poisson process with rate λ_A per hour.

1. **(3 points)** What is the probability that Alice sent exactly three emails during the time interval $[1, 2]$?

   Solution: The number of emails Alice sends in the interval $[1, 2]$ is a Poisson random variable with parameter λ_A. So we have:
   $$\mathbf{P}(3, 1) = \frac{{\lambda_A}^3 e^{-\lambda_A}}{3!}.$$

2. Let Y_1 and Y_2 be the times at which Alice's first and second emails were sent.

   (a) **(3 points)** Find $\mathbf{E}[Y_2 \mid Y_1]$.
   Solution: Define T_2 as the second inter-arrival time in Alice's Poisson process. Then:
   $$Y_2 = Y_1 + T_2$$
   $$\mathbf{E}[Y_2 \mid Y_1] = \mathbf{E}[Y_1 + T_2 \mid Y_1] = Y_1 + \mathbf{E}[T_2] = Y_1 + 1/\lambda_A.$$

   (b) **(3 points)** Find the PDF of Y_1^2.
   Solution: Let $Z = Y_1^2$. Then we first find the CDF of Z and differentiate to find the PDF of Z:
   $$F_Z(z) = \mathbf{P}(Y_1^2 \le z) = \mathbf{P}(-\sqrt{z} \le Y_1 \le \sqrt{z}) = \begin{cases} 1 - e^{-\lambda_A \sqrt{z}} & z \ge 0 \\ 0 & z < 0. \end{cases}$$

   $$f_Z(z) = \frac{dF_Z(z)}{dz} = \lambda_A e^{-\lambda_A \sqrt{z}} \left(\frac{1}{2} z^{-1/2}\right) \qquad (z \ge 0)$$
   $$f_Z(z) = \begin{cases} \frac{\lambda_A}{2\sqrt{z}} e^{-\lambda_A \sqrt{z}} & z \ge 0 \\ 0 & z < 0. \end{cases}$$

   (c) **(3 points)** Find the joint PDF of Y_1 and Y_2.
   Solution:
   $$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= f_{Y_1}(y_1) f_{Y_2|Y_1}(y_2|y_1) \\ &= f_{Y_1}(y_1) f_{T_2}(y_2 - y_1) \\ &= \lambda_A e^{-\lambda_A y_1} \lambda_A e^{-\lambda_A (y_2 - y_1)} \qquad y_2 \ge y_1 \ge 0 \\ &= \begin{cases} \lambda_A^2 e^{-\lambda_A y_2} & y_2 \ge y_1 \ge 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

3. You show up at time 1 and you are told that Alice has sent exactly one email so far. (Only give answers here, no need to justify them.)

   (a) **(3 points)** What is the conditional expectation of Y_2 given this information?
   Solution: Let A be the event {exactly one arrival in the interval [0,1]}. Looking forward from time $t = 1$, the time until the next arrival is simply an exponential random variable (T). So,

   $$\mathbf{E}[Y_2 \mid A] = 1 + \mathbf{E}[T] = 1 + 1/\lambda_A.$$

   (b) **(3 points)** What is the conditional expectation of Y_1 given this information?
   Solution: Given A, the times in this interval are equally likely for the arrival Y_1. Thus,

   $$\mathbf{E}[Y_1 \mid A] = 1/2.$$

4. Bob just finished exercising (without email access) and sits next to Alice at time 1. He starts typing emails at time 1, and fires them according to an independent Poisson process with rate λ_B.

   (a) **(5 points)** What is the PMF of the total number of emails sent by the two of them together during the interval $[0, 2]$?
   Solution: Let K be the total number of emails sent in $[0, 2]$. Let K_1 be the total number of emails sent in $[0, 1)$, and let K_2 be the total number of emails sent in $[1, 2]$. Then $K = K_1 + K_2$ where K_1 is a Poisson random variable with parameter λ_A and K_2 is a Poisson random variable with parameter $\lambda_A + \lambda_B$ (since the emails sent by both Alice and Bob after time $t = 1$ arrive according to the merged Poisson process of Alice's emails and Bob's emails). Since K is the sum of independent Poisson random variables, K is a Poisson random variable with parameter $2\lambda_A + \lambda_B$. So K has the distribution:

   $$p_K(k) = \frac{(2\lambda_A + \lambda_B)^k e^{-(2\lambda_A+\lambda_B)}}{k!} \qquad k = 0, 1, \ldots.$$

   (b) **(5 points)** What is the expected value of the total typing time associated with the email that Alice is typing at the time that Bob shows up? (Here, "total typing time" includes the time that Alice spent on that email both before and after Bob's arrival.)
   Solution: The total typing time Q associated with the email that Alice is typing at the time Bob shows up is the sum of S_0, the length of time between Alice's last email or time 0 (whichever is later) and time 1, and T_1, the length of time from 1 to the time at which Alice sends her current email. T_1 is exponential with parameter λ_A. and $S_0 = \min\{T_0, 1\}$, where T_0 is exponential with parameter λ_A.
   Then,

   $$Q = S_0 + T_1 = \min\{T_0, 1\} + T_1$$

   and

   $$\mathbf{E}[Q] = \mathbf{E}[S_0] + \mathbf{E}[T_1].$$

   We have: $\mathbf{E}[T_1] = 1/\lambda_A$.

We can find $\mathbf{E}[S_0]$ via the law of total expectations:

$$\begin{aligned}
\mathbf{E}[S_0] = \mathbf{E}[\min\{T_0, 1\}] &= \mathbf{P}(T_0 \leq 1)\mathbf{E}[T_0 \mid T_0 \leq 1] + \mathbf{P}(T_0 > 1)\mathbf{E}[1|T_0 > 1] \\
&= \left(1 - e^{-\lambda_A}\right) \int_0^1 t f_{T|T_0 \leq 1}(t)\, dt + e^{-\lambda_A} \\
&= \left(1 - e^{-\lambda_A}\right) \int_0^1 t \frac{\lambda_A e^{-\lambda_A t}}{(1 - e^{-\lambda_A})}\, dt + e^{-\lambda_A} \\
&= \int_0^1 t \lambda_A e^{-\lambda_A t}\, dt + e^{-\lambda_A} \\
&= \frac{1}{\lambda_A} \int_0^1 t \lambda_A^2 e^{-\lambda_A t}\, dt + e^{-\lambda_A} \\
&= \frac{1}{\lambda_A} \left(1 - e^{-\lambda_A} - \lambda_A e^{-\lambda_A}\right) + e^{-\lambda_A} \\
&= \frac{1}{\lambda_A} \left(1 - e^{-\lambda_A}\right)
\end{aligned}$$

where the above integral is evaluated by manipulating the integrand into an Erlang order 2 PDF and equating the integral of this PDF from 0 to 1 to the probability that there are 2 or more arrivals in the first hour (i.e. $\mathbf{P}(Y_2 < 1) = 1 - \mathbf{P}(0, 1) - \mathbf{P}(1, 1)$). Alternatively, one can integrate by parts and arrive at the same result.

Combining the above expectations:

$$\mathbf{E}[Q] = \mathbf{E}[S_0] + \mathbf{E}[T_1] = \frac{1}{\lambda_A} \left(1 - e^{-\lambda_A}\right) + \frac{1}{\lambda_A} = \frac{1}{\lambda_A} \left(2 - e^{-\lambda_A}\right).$$

(c) **(5 points)** What is the expected value of the time until each one of them has sent at least one email? (Note that we count time starting from time 0, and we take into account any emails possibly sent out by Alice during the interval $[0, 1]$.)

Solution: Define U as the time from $t = 0$ until each person has sent at least one email.
Define V as the remaining time from when Bob arrives (time 1) until each person has sent at least one email (so $V = U - 1$).
Define S as the time until Bob sends his first email after time 1.
Define the event $A = \{\text{Alice sends one or more emails in the time interval } [0, 1]\} = \{Y_1 \leq 1\}$, where Y_1 is the time Alice sends her first email.
Define the event $B = \{\text{After time 1, Bob sends his next email before Alice does}\}$, which is equivalent to the event where the next arrival in the merged process from Alice and Bob's orginal processes (starting from time 1) comes from Bob's process.
We have:

$$\mathbf{P}(A) = \mathbf{P}(Y_1 \leq 1) = 1 - e^{-\lambda_A}$$

$$\mathbf{P}(B) = \frac{\lambda_B}{\lambda_A + \lambda_B}.$$

Then,

$$\begin{aligned}
\mathbf{E}[U] &= \mathbf{P}(A)\mathbf{E}[U \mid A] + \mathbf{P}(A^c)\mathbf{E}[U \mid A^c] \\
&= (1-e^{-\lambda_A})(1+\mathbf{E}[V \mid A]) + e^{-\lambda_A}(1+\mathbf{E}[V \mid A^c]) \\
&= (1-e^{-\lambda_A})(1+\mathbf{E}[V \mid A]) + e^{-\lambda_A}(1+\mathbf{P}(B \mid A^c)\mathbf{E}[V \mid B \cap A^c] + \mathbf{P}(B^c \mid A^c)\mathbf{E}[V \mid B^c \cap A^c]) \\
&= (1-e^{-\lambda_A})(1+\mathbf{E}[V \mid A]) + e^{-\lambda_A}(1+\mathbf{P}(B)\mathbf{E}[V \mid B \cap A^c] + \mathbf{P}(B^c)\mathbf{E}[V \mid B^c \cap A^c]) \\
&= (1-e^{-\lambda_A})(1+\mathbf{E}[V \mid A]) + e^{-\lambda_A}\left(1 + \frac{\lambda_B}{\lambda_A+\lambda_B}\mathbf{E}[V \mid B \cap A^c] + \frac{\lambda_A}{\lambda_A+\lambda_B}\mathbf{E}[V \mid B^c \cap A^c]\right).
\end{aligned}$$

Note that $\mathbf{E}[V \mid B^c \cap A^c]$ is the expected value of the time until each of them sends one email after time 1 (since, given A^c, Alice did not send any in the interval $[0,1]$) and given Alice sends an email before Bob. Then this is the expected time until an arrival in the merged process followed by the expected time until an arrival in Bob's process. So, $\mathbf{E}[V \mid B^c \cap A^c] = \frac{1}{\lambda_A+\lambda_B} + \frac{1}{\lambda_B}$.
Similarly, $\mathbf{E}[V \mid B \cap A^c]$ is the time until each sends an email after time 1, given Bob sends an email before Alice. So $\mathbf{E}[V \mid B \cap A^c] = \frac{1}{\lambda_A+\lambda_B} + \frac{1}{\lambda_A}$.
Also, $\mathbf{E}[V \mid A]$ is the expected time it takes for Bob to send his first email after time 1 (since, given A, Alice already sent an email in the interval $[0,1]$). So $\mathbf{E}[V \mid A] = \mathbf{E}[S] = 1/\lambda_B$. Combining all of this with the above, we have:

$$\begin{aligned}
\mathbf{E}[U] &= (1-e^{-\lambda_A})(1+1/\lambda_B) \\
&\quad + e^{-\lambda_A}\left(1 + \frac{\lambda_B}{\lambda_A+\lambda_B}\left(\frac{1}{\lambda_A+\lambda_B} + \frac{1}{\lambda_A}\right) + \frac{\lambda_A}{\lambda_A+\lambda_B}\left(\frac{1}{\lambda_A+\lambda_B} + \frac{1}{\lambda_B}\right)\right).
\end{aligned}$$

(d) **(5 points)** Given that a total of 10 emails were sent during the interval $[0,2]$, what is the probability that exactly 4 of them were sent by Alice?

Solution:

$$\begin{aligned}
\mathbf{P}(\text{Alice sent 4 in } [0,2] \mid \text{total 10 sent in } [0,2]) &= \frac{\mathbf{P}(\text{Alice sent 4 in } [0,2] \cap \text{total 10 sent in } [0,2])}{\mathbf{P}(\text{total 10 sent in } [0,2])} \\
&= \frac{\mathbf{P}(\text{Alice sent 4 in } [0,2] \cap \text{Bob sent 6 } [0,2])}{\mathbf{P}(\text{total 10 sent in } [0,2])} \\
&= \frac{\left(\frac{(2\lambda_A)^4 e^{-2\lambda_A}}{4!}\right)\left(\frac{(\lambda_B)^6 e^{-\lambda_B}}{6!}\right)}{\frac{(2\lambda_A+\lambda_B)^{10} e^{-2\lambda_A+\lambda_B}}{10!}} \\
&= \binom{10}{4}\left(\frac{2\lambda_A}{2\lambda_A+\lambda_B}\right)^4\left(\frac{\lambda_B}{2\lambda_A+\lambda_B}\right)^6.
\end{aligned}$$

As the form of the solution suggests, the problem can be solved alternatively by computing the probability of a single email being sent by Alice, given it was sent in the interval $[0,2]$. This can be found by viewing the number of emails sent by Alice in $[0,2]$ as the number of arrivals arising from a Poisson process with twice the rate ($2\lambda_A$) in an interval of half the duration (particularly, the interval $[1,2]$), then merging this process with Bob's process. Then the probability that an email sent in the interval $[0,2]$ was sent by Alice is the probability that an arrival in this new merged process came from the newly constructed $2\lambda_A$ rate process:

$$p = \frac{2\lambda_A}{2\lambda_A + \lambda_B}.$$

Then, out of 10 emails, the probability that 4 came from Alice is simply a binomial probability with 4 successes in 10 trials, which agrees with the solution above.

5. **(5 points)** Suppose that $\lambda_A = 4$. Use Chebyshev's inequality to find an upper bound on the probability that Alice sent at least 5 emails during the time interval $[0, 1]$. Does the Markov inequality provide a better bound?

   Solution:

   Let N be the number of emails Alice sent in the interval $[0, 1]$. Since N is a Poisson random variable with parameter λ_A,

   $$\mathbf{E}[N] = \text{var}(N) = \lambda_A = 4.$$

   To apply the Chebyshev inequality, we recognize:

   $$\mathbf{P}(N \geq 5) = \mathbf{P}(N - 4 \geq 1) \leq \mathbf{P}(|N - 4| \geq 1) \leq \frac{\text{var}(N)}{1^2} = 4.$$

   In this case, the upper-bound of 4 found by application of the Chebyshev inequality is uninformative, as we already knew $\mathbf{P}(N \geq 5) \leq 1$.

   To find a better bound on this probability, use the Markov inequality, which gives:

   $$\mathbf{P}(N \geq 5) \leq \frac{\mathbf{E}[N]}{5} = \frac{4}{5}.$$

6. **(5 points)** You do not know λ_A but you watch Alice for an hour and see that she sent exactly 5 emails. Derive the maximum likelihood estimate of λ_A based on this information.

   Solution:

   $$\begin{aligned}
   \hat{\lambda}_A &= \arg\max_{\lambda} \log\left(p_N(5;\lambda)\right) \\
   &= \arg\max_{\lambda} \log\left(\frac{\lambda^5 e^{-\lambda}}{5!}\right) \\
   &= \arg\max_{\lambda} -\log(5!) + 5\log(\lambda) - \lambda.
   \end{aligned}$$

   Setting the first derivative to zero

   $$\begin{aligned}
   \frac{5}{\lambda} - 1 &= 0 \\
   \hat{\lambda}_A &= 5.
   \end{aligned}$$

7. **(5 points)** We have reasons to believe that λ_A is a large number. Let N be the number of emails sent during the interval $[0, 1]$. Justify why the CLT can be applied to N, and give a precise statement of the CLT in this case.

   Solution: With λ_A large, we assume $\lambda_A \gg 1$. For simplicity, assume λ_A is an integer. We can divide the interval $[0, 1]$ into λ_A disjoint intervals, each with duration $1/\lambda_A$, so that these intervals span the entire interval from $[0, 1]$. Let N_i be the number of arrivals in the ith such interval, so that the N_i's are independent, identically distributed Poisson random variables with parameter 1. Since N is defined as the number of arrivals in the interval $[0, 1]$, then $N = N_1 + \cdots + N_{\lambda_A}$. Since $\lambda_A \gg 1$, then N is the sum of a large number of independent and identically distributed random variables, where the distribution of N_i does not change as the number of terms in the sum increases. Hence, N is approximately normal with mean λ_A and variance λ_A.

   If λ_A is not an integer, the same argument holds, except that instead of having λ_A intervals, we have an integer number of intervals equal to the integer part of λ_A ($\bar{\lambda}_A$ =floor(λ_A)) of length $1/\lambda_A$ and an extra interval of a shorter length $(\lambda_A - \bar{\lambda}_A)/\lambda_A$.

   Now, N is a sum of λ_A independent, identically distributed Poisson random variables with parameter 1 added to another Poisson random variable (also independent of all the other Poisson random variables) with parameter $(\lambda_A - \bar{\lambda}_A)$. In this case, N would need a small correction to apply the central limit theorem as we are familiar with it; however, it turns out that even without this correction, adding the extra Poisson random variable does not preclude the distribution of N from being approximately normal, for large λ_A, and the central limit theorem still applies.

   To arrive at a precise statement of the CLT, we must "standardize" N by subtracting its mean then dividing by its standard deviation. After having done so, the CDF of the standardized version of N should converge to the standard normal CDF as the number of terms in the sum approaches infinity (as $\lambda_A \to \infty$).

   Therefore, the precise statement of the CLT when applied to N is:

$$\lim_{\lambda_A \to \infty} \mathbf{P}\left(\frac{N - \lambda_A}{\sqrt{\lambda_A}} \leq z\right) = \Phi(z)$$

   where $\Phi(z)$ is the standard normal CDF.

8. **(5 points)** Under the same assumption as in last part, that λ_A is large, you can now pretend that N is a normal random variable. Suppose that you observe the value of N. Give an (approximately) 95% confidence interval for λ_A. State precisely what approximations you are making.

   Possibly useful facts: The cumulative normal distribution satisfies $\Phi(1.645) = 0.95$ and $\Phi(1.96) = 0.975$.

   Solution: We begin by estimating λ_A with its ML estimator $\hat{\Lambda}_A = N$, where $\mathbf{E}[N] = \lambda_A$. With λ_A large, the CLT applies, and we can assume N has an approximately normal distribution. Since $\text{var}(N) = \lambda_A$, we can also approximate the variance of N with ML estimator for λ_A, so $\text{var}(N) \approx N$, and $\sigma_N \approx \sqrt{N}$.

To find the 95% confidence interval, we find β such that:

$$\begin{aligned} 0.95 &= \mathbf{P}\left(|N-\lambda_A| \leq \beta\right) \\ &= \mathbf{P}\left(\frac{|N-\lambda_A|}{\sqrt{N}} \leq \frac{\beta}{\sqrt{N}}\right) \\ &\approx 2\Phi\left(\frac{\beta}{\sqrt{N}}\right). \end{aligned}$$

So, we find:

$$\beta \approx \sqrt{N}\Phi^{-1}(0.975) = 1.96\sqrt{N}.$$

Thus, we can write:

$$\mathbf{P}(N - 1.96\sqrt{N} \leq \lambda_A \leq N + 1.96\sqrt{N}) \approx 0.95.$$

So, the approximate 95% confidence interval is: $[N - 1.96\sqrt{N}, N + 1.96\sqrt{N}]$.

9. You are now told that λ_A is actually the realized value of an exponential random variable Λ, with parameter 2:

$$f_\Lambda(\lambda) = 2e^{-2\lambda}, \qquad \lambda \geq 0.$$

(a) **(5 points)** Find $\mathbf{E}[N^2]$.
Solution:

$$\begin{aligned} \mathbf{E}[N^2] = \mathbf{E}[\mathbf{E}[N^2 \mid \Lambda]] &= \mathbf{E}[\mathrm{var}(N \mid \Lambda) + (\mathbf{E}[N \mid \Lambda])^2] \\ &= \mathbf{E}[\Lambda + \Lambda^2] \\ &= \mathbf{E}[\Lambda] + \mathrm{var}(\Lambda) + (\mathbf{E}[\Lambda])^2 \\ &= \frac{1}{2} + \frac{2}{2^2} \\ &= 1. \end{aligned}$$

(b) **(5 points)** Find the linear least squares estimator of Λ given N.
Solution:

$$\hat{\Lambda}_{\mathrm{LLMS}} = \mathbf{E}[\Lambda] + \frac{\mathrm{cov}(N,\Lambda)}{\mathrm{var}(N)}(N - \mathbf{E}[N]).$$

Solving for the above quantities:

$$\mathbf{E}[\Lambda] = \frac{1}{2}$$

$$\mathbf{E}[N] = \mathbf{E}[\mathbf{E}[N \mid \Lambda]] = \mathbf{E}[\Lambda] = \frac{1}{2}.$$

$$\mathrm{var}(N) = \mathbf{E}[N^2] - (\mathbf{E}[N])^2 = 1 - \frac{1}{2^2} = \frac{3}{4}.$$

$$\mathrm{cov}(N,\Lambda) = \mathbf{E}[N\Lambda] - \mathbf{E}[N]\mathbf{E}[\Lambda] = \mathbf{E}[\mathbf{E}[N\Lambda \mid \Lambda]] - (\mathbf{E}[\Lambda])^2 = \mathbf{E}[\Lambda^2] - (\mathbf{E}[\Lambda])^2 = \mathrm{var}(\Lambda) = \frac{1}{4}.$$

Substituting these into the equation above:

$$\begin{aligned}\hat{\Lambda}_{\text{LLMS}} &= \mathbf{E}[\Lambda] + \frac{\text{cov}(N,\Lambda)}{\text{var}(N)}(N - \mathbf{E}[N]) \\ &= \frac{1}{2} + \frac{1/4}{3/4}\left(N - \frac{1}{2}\right) \\ &= \frac{1}{3}(N+1)\,.\end{aligned}$$

6.041 / 6.431 Probabilistic Systems Analysis and Applied Probability

Fall 2010