MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering & Computer Science
6.041/6.431: Probabilistic Systems Analysis
(Fall 2010)

Problem Set 7: Solutions

1. (a) The event of the ith success occuring before the jth failure is equivalent to the ith success occurring within the first $(i + j - 1)$ trials (since the ith success must occur no later than the trial right before the jth failure). This is equivalent to event that i or more successes occur in the first $(i + j - 1)$ trials (where we can have, at most, $(i + j - 1)$ successes). Let S_i be the time of the ith success, F_j be the time of the jth failure, and N_k be the number of successes in the first k trials (so N_k is a binomial random variable over k trials). So we have:

$$\mathbf{P}(S_i < F_j) = \mathbf{P}(N_{i+j-1} \geq i) = \sum_{k=i}^{i+j-1} \binom{i+j-1}{k} p^k (1-p)^{i+j-1-k}$$

(b) Let K be the number of successes which occur before the jth failure, and L be the number of trials to get to the jth failure. L is simply a jth order Pascal, with probability of $1 - p$ (since we are now interested in the failures, not the successes.) Plugging into the formula for jth order Pascal random variable,

$$\mathbf{E}[L] = \frac{j}{1-p}, \sigma_K^2 = \frac{p}{(1-p)^2}j$$

Since $K = L - j$,

$$\mathbf{E}[K] = \frac{p}{1-p}j, \sigma_K^2 = \frac{p}{(1-p)^2}j$$

(c) This expression is the same as saying we need at least 42 trials to get the 17th success. Therefore, it can be rephrased as having a maximum of 16 successes in the first 41 trials. Hence $b = 41$, $a = 16$.

2. A successful call occurs with probability $p = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$.

(a) Fred will give away his first sample on the third call if the first two calls are failures and the third is a success. Since the trials are independent, the probability of this sequence of events is simply

$$(1-p)(1-p)p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

(b) The event of interest requires failures on the ninth and tenth trials and a success on the eleventh trial. For a Bernoulli process, the outcomes of these three trials are independent of the results of any other trials and again our answer is

$$(1-p)(1-p)p = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$$

(c) We desire the probability that L_2, the time to the second arrival is equal to five trials. We know that $p_{L_2}(\ell)$ is a Pascal PMF of order 2, and we have

$$p_{L_2}(5) = \binom{5-1}{2-1} p^2 (1-p)^{5-2} = 4 \cdot \left(\frac{1}{2}\right)^5 = \frac{1}{8}$$

(d) Here we require the conditional probability that the experimental value of L_2 is equal to 5, given that it is greater than 2.

$$\mathbf{P}(L_2 = 5 | L_2 > 2) = \frac{p_{L_2}(5)}{P(L_2 > 2)} = \frac{p_{L_2}(5)}{1 - p_{L_2}(2)}$$

$$= \frac{\binom{5-1}{2-1}p^2(1-p)^{5-2}}{1 - \binom{2-1}{2-1}p^2(1-p)^0} = \frac{4 \cdot \left(\frac{1}{2}\right)^5}{1 - \left(\frac{1}{2}\right)^2} = \frac{1}{6}$$

(e) The probability that Fred will complete at least five calls before he needs a new supply is equal to the probability that the experimental value of L_2 is greater than or equal to 5.

$$\mathbf{P}(L_2 \geq 5) = 1 - P(L_2 \leq 4) = 1 - \sum_{\ell=2}^{4} \binom{\ell-1}{2-1}p^2(1-p)^{\ell-2}$$

$$= 1 - (\frac{1}{2})^2 - \binom{2}{1}(\frac{1}{2})^3 - \binom{3}{1}(\frac{1}{2})^4 = \frac{5}{16}$$

(f) Let discrete random variable F represent the number of failures before Fred runs out of samples on his mth successful call. Since L_m is the number of trials up to and including the mth success, we have $F = L_m - m$. Given that Fred makes L_m calls before he needs a new supply, we can regard each of the F unsuccessful calls as trials in another Bernoulli process with parameter r, where r is the probability of a success (a disappointed dog) obtained by

$$r = \mathbf{P}(\text{dog lives there | Fred did not leave a sample})$$

$$= \frac{\mathbf{P}(\text{dog lives there AND door not answered})}{1 - \mathbf{P}(\text{giving away a sample})} = \frac{\frac{1}{4} \cdot \frac{2}{3}}{1 - \frac{1}{2}} = \frac{1}{3}$$

We define X to be a Bernoulli random variable with parameter r. Then, the number of dogs passed up before Fred runs out, D_m, is equal to the sum of F Bernoulli random variables each with parameter $r = \frac{1}{3}$, where F is a random variable. In other words,

$$D_m = X_1 + X_2 + X_3 + \cdots + X_F.$$

Note that D_m is a sum of a random number of independent random variables. Further, F is independent of the X_i's since the X_i's are defined in the conditional universe where the door is not answered, in which case, whether there is a dog or not does not affect the probability of that trial being a failed trial or not. From our results in class, we can calculate its expectation and variance by

$$\mathbf{E}[D_m] = \mathbf{E}[F]\mathbf{E}[X]$$
$$\text{var}(D_m) = \mathbf{E}[F]\text{var}(X) + (\mathbf{E}[X])^2\text{var}(F),$$

where we make the following substitutions.

$$\mathbf{E}[F] = \mathbf{E}[L_m - m] = \frac{m}{p} - m = m.$$

$$\text{var}(F) = \text{var}(L_m - m) = \text{var}(L_m) = \frac{m(1-p)}{p^2} = 2m.$$

$$\mathbf{E}[X] = r = \frac{1}{3}.$$

$$\text{var}(X) = r(1-r) = \frac{2}{9}.$$

Finally, substituting these values, we have

$$
\begin{aligned}
\mathbf{E}[D_m] &= m \cdot \frac{1}{3} = \frac{m}{3} \\
\mathrm{var}(D_m) &= m \cdot \frac{2}{9} + \left(\frac{1}{3}\right)^2 \cdot 2m = \frac{4m}{9}
\end{aligned}
$$

3. We view the random variables T_1 and T_2 as interarrival times in two independent Poisson processes both with rate λ S as the interarrival time in a third Poisson process (independent from the first two) with rate μ. We are interested in the expected value of the time Z until either the first process has had two arrivals or the second process has had an arrival.

Given that the first arrival was from the second process, the expected wait time for that arrival would be $\frac{1}{\mu+\lambda}$. The probability of an arrival from the second process is $\frac{\mu}{\mu+\lambda}$. Given that the first arrival time was from the first process, the expected wait time would be that for first arrival, $\frac{1}{\mu+\lambda}$, plus the expected wait time for another arrival from the merged process. Similarily, the probability of an arrival from the first process is $\frac{\lambda}{\mu+\lambda}$. Thus,

$$
\mathbf{E}[Z] = \mathbf{P}(\text{Arrival from second process})\mathbf{E}[\text{wait time}|\text{Arrival from second process}] +
$$
$$
\mathbf{P}(\text{Arrival from first process})\mathbf{E}[\text{wait time}|\text{Arrival from first process}]
$$
$$
= \frac{\mu}{\mu+\lambda} \cdot \frac{1}{\mu+\lambda} + \frac{\lambda}{\mu+\lambda} \cdot \left(\frac{1}{\mu+\lambda} + \frac{1}{\mu+\lambda}\right).
$$

After some simplifications, we see that

$$
\mathbf{E}[Z] = \frac{1}{\mu+\lambda} + \frac{\lambda}{\mu+\lambda} \cdot \frac{1}{\mu+\lambda}
$$

4. The dot location of the yarn, as related to the size of the pieces of the yarn cut for any particular customer, can be viewed in light of the random incident paradox.

   (a) Here, the length of each piece of yarn is exponentially distributed. As explained on page 298 of the text, due to the memorylessness of the exponential, the distribution of the length of the piece of yarn containing the red dot is a second order Erlang. Thus, the $\mathbf{E}[R] = 2\mathbf{E}[L] = \frac{2}{\lambda}$.

   (b) Think of exponentially-spaced marks being made on the yarn, so the length requested by the customers each involve *three* such sections of exponentially distributed lengths (since the PDF of L is third-order Erlang). The piece of yarn with the dot will have the dot in any one of these three sections, and the expected length of that section, by (a), will be $2/\lambda$, while the expected lengths of the other two sections will be $1/\lambda$. Thus, the total expected length containing the dot is $4/\lambda$.

   In general, for processes, in which the interarrival intervals with distribution $F_X(x)$ are IID, the expected length of an arbitrarily chosen interval is $\frac{\mathbf{E}[X^2]}{\mathbf{E}[X]}$. We see that for the above parts, this formula is certainly valid.

   (c) Using the formula stated above, $\mathbf{E}[L] = \int_0^1 \ell^2 e^\ell \, d\ell = e^\ell(\ell^2 - 2\ell + 2)]_0^1 = e - 2$
   $E[L^2] = \int_0^1 \ell^3 e^\ell \, d\ell = e^\ell(\ell^3 - 3\ell^2 + 6\ell - 6)]_0^1 = 6 - 2e$
   Hence,

   $$
   \mathbf{E}[R] = \boxed{\frac{6 - 2e}{e - 2}}.
   $$

5. (a) We know there are n arrivals in t amount of time, so we are looking for how many extra arrivals there are in s amount of time.

$$p_{M|N}(m|n) = \frac{(\lambda s)^{m-n} e^{-\lambda s}}{(m-n)!} \quad \text{for } m \geq n \geq 0$$

(b) By definition:

$$
\begin{aligned}
p_{N,M}(n,m) &= p_{M|N}(m|n) p_N(n) \\
&= \frac{\lambda^m s^{m-n} t^n e^{-\lambda(s+t)}}{(m-n)! n!} \quad \text{for } m \geq n \geq 0
\end{aligned}
$$

(c) By definition:

$$
\begin{aligned}
p_{N|M}(n|m) &= \frac{p_{M,N}(m,n)}{p_M(m)} \\
&= \binom{m}{n} \frac{s^{m-n} t^n}{(s+t)^m} \quad \text{for } m \geq n \geq 0
\end{aligned}
$$

(d) We want to find: $\mathbf{P}(N = n | M = m)$. Given M=m, we know that the m arrivals are uniformly distributed between 0 and t+s. Consider each arrival a success if it occurs before time t, and a failure otherwise. Therefore given M=m, N is a binomial random variable with m trials and probability of success $\frac{t}{t+s}$. We have the desired probability:

$$\mathbf{P}(N = n | M = m) = \binom{m}{n} \left(\frac{t}{t+s}\right)^n \left(\frac{s}{t+s}\right)^{m-n} \quad \text{for } m \geq n \geq 0$$

(e) We can rewrite the expectatation as:

$$
\begin{aligned}
\mathbf{E}[NM] &= \mathbf{E}[N(M-N) + N^2] \\
&= \mathbf{E}[N]\mathbf{E}[M-N] + \mathbf{E}[N^2] \\
&= (\lambda t)(\lambda s) + \left(\text{var}(N) + (\mathbf{E}[N])^2\right) \\
&= (\lambda t)(\lambda s) + \lambda t + (\lambda t)^2
\end{aligned}
$$

where the second equality is obtained via the independent increment property of the poisson process.

6. The described process for cars passing the checkpoint is a Poisson process with an arrival rate of $\lambda = 2$ cars per minute.

(a) The first and third moments are, respectively,

$$\mathbf{E}[T] = \frac{1}{\lambda} = \frac{1}{2} \qquad \mathbf{E}[T^3] = \int_0^\infty t^3 2 e^{-2t} dt = \frac{3!}{2^3} \underbrace{\int_0^\infty \frac{2^4 t^3 e^{-2t}}{3!} dt}_{=1} = \frac{3}{4}$$

where we recognized the integrand to be a 4th-order Erlang PDF and therefore integrating it over the entire range of the random variable must sum to unity.

(b) The Poisson process is memoryless, and thus the history of events in the previous 4 minutes does not affect the future. So, the conditional PMF for K is equivalent to the unconditional PMF that describes the number of Poisson arrivals in an interval of time, which in this case is $\tau = 6$ minutes and thus $(\lambda\tau) = 12$:

$$p_K(k) = \frac{12^k e^{12}}{k!} \ , \quad k = 0, 1, 2, \ldots \qquad .$$

(c) The first dozen computer cards are used up upon the 36th car arrival. Letting D denote this total time, $D = T_1 + T_2 + \ldots + T_{36}$, where each independent T_i is exponentially distributed with parameter $\lambda = 2$, the distribution for D is therefore a 36th-order Erlang distribution with PDF and expected value of, respectively,

$$f_D(d) = \frac{2^{36} d^{35} e^{-2d}}{35!} \ , \quad d \geq 0 \qquad \mathbf{E}[D] = 36\mathbf{E}[T] = 18$$

(d) In both experiments, because a card completes after registering three cars, we are considering the amount of time it takes for three cars to pass the checkpoint. In the second experiment, however, note that the manner with which the particular card is selected is biased towards cards that are in service longer. That is, the time instant at which we come to the corner is more likely to fall within a longer interarrival period – one of the three interarrival times that adds up to the total time the card is in service is selected by *random incidence* (see the end of Section 6.2 in text).

   i. The service time of any particular completed card is given by $Y = T_1 + T_2 + T_3$, and thus Y is described by a 3rd-order Erlang distribution with paramater $\lambda = 2$:

$$\mathbf{E}[Y] = \frac{3}{\lambda} = \frac{3}{2} \qquad \text{var}(Y) = \frac{3}{\lambda^2} = \frac{3}{4} \qquad .$$

   ii. The service time of a particular completed card with one of the three interarrival times selected by random incidence is $W = T_1 + T_2 + L$, where L is the interarrival period that contains the time instant we arrived at the corner. Following the arguments in the text, L is Erlang of order two and thus W is described by a 4th-order Erlang distribution with parameter $\lambda = 2$:

$$\mathbf{E}[W] = \frac{4}{\lambda} = 2 \qquad \text{var}(W) = \frac{4}{\lambda^2} = 1 \qquad .$$

G1†. For simplicity, introduce the notation $N_i = N(G_i)$ for $i = 1, ..., n$ and $N_G = N(G)$. Then

$$
\begin{aligned}
\mathbf{P}(N_1 = k_1, ..., N_n = k_n | N_G = k) &= \frac{\mathbf{P}(N_1 = k_1, ..., N_n = k_n, N_G = k)}{\mathbf{P}(N_G = k)} \\
&= \frac{\mathbf{P}(N_1 = k_1) \cdots \mathbf{P}(N_n = k_n)}{\mathbf{P}(N_G = k)} \\
&= \frac{\left(\frac{(c_1\lambda)^{k_1} e^{-c_1\lambda}}{k_1!}\right) \cdots \left(\frac{(c_n\lambda)^{k_n} e^{-c_n\lambda}}{k_n!}\right)}{\left(\frac{(c\lambda)^k e^{-c\lambda}}{k!}\right)} \\
&= \frac{k!}{k_1! \ldots k_n!} \left(\frac{c_1}{c}\right)^{k_1} \cdots \left(\frac{c_n}{c}\right)^{k_n} \\
&= \begin{pmatrix} k \\ k_1 \ \cdots \ k_n \end{pmatrix} \left(\frac{c_1}{c}\right)^{k_1} \cdots \left(\frac{c_n}{c}\right)^{k_n}
\end{aligned}
$$

The result can be interpreted as a *multinomial distribution*. Imagine we throw an n-sided die k times, where Side i comes up with probability $p_i = c_i/c$. The probability that side i comes up k_i times is given by the expression above. Now relating it back to the Poisson process that we have, each side corresponds to an interval that we sample, and the probability that we sample it depends directly on its relative length. This is consistent with the intuition that, given a number of Poisson arrivals in a specified interval, the arrivals are uniformly distributed.

6.041 / 6.431 Probabilistic Systems Analysis and Applied Probability
Fall 2010