# INTRODUCTION TO PROBABILITY

by

**Dimitri P. Bertsekas and John N. Tsitsiklis**

## CHAPTER 5: ADDITIONAL PROBLEMS†

Last updated: August 6, 2005

† Problems marked with "[D]" are from "Fundamentals of Applied Probability", by Alvin Drake, and are included here with the author's permission.

## SECTION 5.1. The Bernoulli Process

**Problem 1.** We are given a coin for which the probability of heads is $p$ $(0 < p < 1)$ and the probability of tails is $1 - p$. Consider a sequence of independent flips of the coin.

(a) Let $Y$ be the number of flips up to and including the flip on which the first head occurs. Write down the PMF of $Y$.

(b) Let $X$ be the number of heads that occur on any particular flip. Write down $\mathbf{E}[X]$ and $\mathrm{var}(X)$.

(c) Let $K$ be the number of heads that occur on the first $n$ flips of the coin. Determine the PMF, mean, and variance of $K$.

(d) Given that a total of exactly six heads resulted from the first nine flips, what is the conditional probability that both the first and seventh flips were tails?

(e) Let $H$ be the number of heads that occur on the first twenty flips. Let C be the event that a total of exactly ten heads resulted from the first eighteen flips. Find $\mathbf{E}[H \mid C]$ and the conditional variance $\mathrm{var}(H \mid C)$.

**Solution.** (a) $p_Y(k) = (1-p)^{k-1}p \qquad k = 1, 2, \ldots.$

(b) $\mathbf{E}[X] = p$, $\mathrm{var}(X) = p(1-p)$

(c) $p_K(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n, \quad \mathbf{E}[K] = np, \quad \mathrm{var}(K) = np(1-p).$

(d) There are $\binom{9}{6}$ 9-flip sequences with 6 heads. Out of these, there are $\binom{7}{6}$ in which the first and seventh flips are tails. The desired probability is $\binom{7}{6}/\binom{9}{6} = 1/12$.

(e) We have $\mathbf{E}[H \mid C] = 10 + 2p$ and $\mathrm{var}(H \mid C) = 2p(1-p)$.

**Problem 2.** At each trial of a game, Don and Greg flip biased coins, simultaneously but independently. For each trial, the probability of heads is $p_D$ and $p_G$ for Don and Greg, respectively.

(a) Given that the flips on a particular trial resulted in 2 heads, find the PMF of the number of additional trials up to and including the next trial on which 2 heads result.

(b) Given that the flips on a particular trial resulted in at least one head, find the probability that Don flipped a head on that trial.

(c) Starting from a trial on which no heads result, find the probability that Don's next flip of a head will occur before Greg's next flip of a head.

(d) Given that Don receives \$$d$ for each head he flips, and Greg receives \$$g$ for each head he flips, find the transform associated with the total amount of money earned by the two players during the first $n$ trials.

**Solution.** (a) Each trial has probability $p_D p_G$ of resulting in two heads, independently of other trials. Therefore, the number of trials until the first time that two heads are obtained is geometric with parameter $p_D p_G$, and

$$p_M(m) = p_D p_G (1 - p_D p_G)^{m-1}, \qquad m = 1, 2, 3, \ldots.$$

(b) The probability of at least one head is $p_D + p_G - p_D p_G$. The probability that Don flipped a head is $p_D$. Using the definition of conditional probability, the answer is

$$\frac{p_D}{p_D + p_G - p_D p_G}.$$

(c) Let $N$ be the trial at which heads was obtained for the first time. We are looking for the probability that at trial $N$, Don's coin resulted in heads, while Greg's coin resulted in tails. Let $A$ be this event.

Let us first find $\mathbf{P}(A \mid N = k)$. The conditioning information tells us that there is at least one head at trial $k$, and no heads before that. The information about what happened in earlier trials is inconsequential, because different trials are independent. Thus, we can just calculate the probability that Greg's coin resulted in tails, given that there was at least one head, and obtain

$$\mathbf{P}(A \mid N = k) = \frac{p_D(1 - p_G)}{p_D + p_G - p_D p_G}.$$

By the total probability theorem,

$$\mathbf{P}(A) = \sum_{k=1}^{\infty} \mathbf{P}(N = k)\, \mathbf{P}(A \mid N = k) = \frac{p_D(1 - p_G)}{p_D + p_G - p_D p_G}.$$

(d) Let $M$ be the total amount of money earned by Don and Greg in the first $n$ trials. Let $D_i$ and $G_i$ be the amount of money earned by Don and Greg, respectively, in the $i$th trial, so that $M = D_1 + D_2 + \cdots + D_n + G_1 + G_2 + \cdots + G_n$. The transforms associated with $D_i$ and $G_i$ are

$$M_{D_i}(s) = 1 - p_D + p_D e^{sd}, \qquad M_{G_i}(s) = 1 - p_G + p_G e^{sg}.$$

Thus,

$$M_M(s) = \left[(1 - p_D + p_D e^{sd})(1 - p_G + p_G e^{sg})\right]^n.$$

**Problem 3.** **[D]** To cross a single lane of moving traffic, we require at least a duration $d$. Successive car interarrival times are independently and identically distributed with probability density function $f_T(t)$. If an interval between successive cars is longer than $d$, we say that the interval represents a single opportunity to cross the lane. Assume that car lengths are small relative to intercar spacing and that our experiment begins the instant after the zeroth car goes by. Determine, in as simple form as possible, expressions for the probability that:

(a) We can cross for the first time just before the $n$th car goes by.

(b) We shall have had exactly $m$ opportunities by the instant the $n$th car goes by.

(c) The occurrence of the $m$th opportunity is immediately followed by the arrival of the $n$th car.

**Solution.** (a) Let $p$ be the probability that a given interval is at least as large as $d$, so that it represents a crossing opportunity ("success"). We have

$$p = \int_t^\infty f_T(t)\,dt.$$

In particular, the probability that the $n$th interval (before the $n$th car arrives) is at least $d$ is equal to $p$.

(b) Each interval corresponds to an independent trial with success probability $p$. Thus, the probability of having exactly $m$ successes in $n$ trials is

$$\binom{n}{m} p^m (1-p)^{n-m}, \qquad m = 0, 1, \ldots, n.$$

(c) This is the probability that the $m$th success occurs at the $n$th trial and is given by the Pascal PMF:

$$\binom{n-1}{m-1} p^m (1-p)^{n-m}, \qquad m \geq 1, \quad n \geq m.$$

**Problem 4.** Let $Y_{17}$ be a Pascal random variable of order 17. Find the numerical values of $a$ and $b$ in the equation

$$\sum_{l=42}^{\infty} p_{Y_{17}}(l) = \sum_{k=0}^{a} \binom{b}{k} p^k (1-p)^{(b-k)},$$

and explain.

**Solution.** The left-hand side is the probability that we need at least 42 trials to get the 17th success. It is the same as the probability of having at most 16 successes in the first 41 trials, which is

$$\sum_{k=0}^{16} \binom{41}{k} p^k (1-p)^{41-k},$$

and therefore, $a = 16$, $b = 41$.

**Problem 5.** **[D]** Fred is giving out samples of dog food. He makes calls door to door, but he leaves a sample (one can) only on those calls for which the door is answered **a**nd a dog is in residence. On any call the probability of the door being answered is 3/4, and the probability that any household has a dog is 2/3. Assume that the events "Door answered" and "A dog lives here" are independent and also that the outcomes of all calls are independent.

(a) Determine the probability that Fred gives away his first sample on his third call.

(b) Given that he has given away exactly four samples on his first eight calls, determine the conditional probability that Fred will give away his fifth sample on his eleventh call.

(c) Determine the probability that he gives away his second sample on his fifth call.

(d) Given that he did not give away his second sample on his second call, determine the conditional probability that he will leave his second sample on his fifth call.

(e) We will say that Fred "needs a new supply" immediately **a**fter the call on which he gives away his last can. If he starts out with two cans, determine the probability that he completes at least five calls before he needs a new supply.

(f) If he starts out with exactly $m$ cans, determine the expected value and variance of $D_m$, the number of homes with dogs which he passes up (because of no answer) before he needs a new supply.

**Solution.** For all but the last part of this problem, we may consider each call to be a Bernoulli trial where the probability of success (door answered and dog in residence) is given by $p = (3/4) \cdot (2/3) = 1/2$.

(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 $(1-p)(1-p)p = 1/8$.

(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 $(1 - p)(1-p)p = 1/8$.

(c) We desire the probability that $Y_2$, the time to the second arrival, is equal to five trials. We know that $Y_2$ is a Pascal random variable, and we have

$$p_{Y_2}(5) = \binom{5-1}{2-1} p^2 (1-p)^{5-2} = \frac{4}{32} = \frac{1}{8}.$$

(d) Here we require the conditional probability that $Y_2$ is equal to 5, given that it is greater than 2. We have

$$\begin{aligned} p_{Y_2 \mid Y_2 > 2}(5 \mid Y_2 > 2) &= \frac{p_{Y_2}(5)}{P(Y_2 > 2)} = \frac{p_{Y_2}(5)}{1 - p_{Y_2}(2)} \\ &= \frac{\binom{5-1}{2-1} p^2 (1-p)^{5-2}}{1 - \binom{2-1}{2-1} p^2 (1-p)^0} = \frac{1/8}{3/4} \end{aligned}$$

(e) The probability that Fred will complete at least five calls before he needs a new supply is equal to the probability that $Y_2$ is greater than or equal to 5. We have

$$\mathbf{P}(Y_2 \geq 5) = 1 - \mathbf{P}(Y_2 \leq 4) = 1 - \sum_{l=2}^{4} \binom{l-1}{2-1} p^2 (1-p)^{l-2} = \frac{5}{16}.$$

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

$$\begin{aligned} p' &= \mathbf{P}(\text{dog lives there} \mid \text{Fred did not leave a sample}) \\ &= \frac{\frac{1}{4} \cdot \frac{2}{3}}{\frac{3}{4} \cdot \frac{1}{3} + \frac{1}{4} \cdot \frac{2}{3} + \frac{1}{4} \cdot \frac{1}{3}} = \frac{1}{3}. \end{aligned}$$

We define $X$ to be a Bernoulli random variable with parameter $p'$. 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 $p' = 1/3$. 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. We can calculate its expectation and variance using the formulas

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

We note that

$$\begin{aligned} \mathbf{E}[F] &= \mathbf{E}[Y_m - m] = \frac{m}{p} - m = m, \\ \mathbf{E}[X] &= p' = \frac{1}{3}, \\ \text{var}(F) &= \text{var}(L_m - m) = \text{var}(L_m) = m\frac{(1-p)}{p^2} = 2m, \\ \text{var}(X) &= p'(1-p') = \frac{2}{9}. \end{aligned}$$

We therefore obtain

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

**Problem 6.** Alice and Bob alternate playing at the casino table. (Alice starts and plays at odd times $i = 1, 3, \ldots$; Bob plays at even times $i = 2, 4, \ldots$.) At each time $i$, the net gain of whoever is playing is a random variable $G_i$ with the following PMF:

$$p_G(g) = \begin{cases} 1/3, & g = -2, \\ 1/2, & g = 1, \\ 1/6, & g = 3. \end{cases}$$

Assume that the net gains at different times are independent. We refer to an outcome of $-2$ as a "loss," and an outcome of 1 or 3 as a "win."

(a) They keep gambling until the first time where a loss by Bob immediately follows a loss by Alice. Write down the PMF of the total number of rounds played. (A round consists of two plays, one by Alice and then one by Bob.)

(b) Write down an expression for the transform of the net gain of Alice up to the time of the first loss by Bob.

(c) Write down the PMF for $Z$, defined as the time at which Bob has his third loss.

(d) Let $N$ be the number of rounds until each one of them has won at least once. Find $\mathbf{E}[N]$.

**Solution.** (a) For each round, the probability that both Alice and Bob have a loss is $(1/3) \cdot (1/3) = (1/9)$. Let $X$ be the total number of rounds played until the first time that both have a loss. Then, $X$ is a geometric random variable, and its PMF is

$$p_X(k) = (1-p)^{k-1} p = \left(\frac{8}{9}\right)^{k-1} \left(\frac{1}{9}\right), \qquad k = 1, 2, \ldots$$

(b) Denote by $Y_i$ the gain from Alice's $i$th game. Let $Z$ be a random variable representing the total gain of Alice up to the time of the first loss by Bob. Thus,

$$Z = Y_1 + Y_2 + Y_3 + \cdots + Y_K,$$

where the random variable $K$ indicates the number of games Bob played up to and including his first loss (Alice will play exactly $K$ games because she plays before Bob in each round). The transform of $Z$ is obtained by

$$M_Z(s) = M_K(s)\Big|_{e^s = M_Y(s)}.$$

Note that $K$ is a geometric random variable with parameter $p = 1/3$. Therefore, the transform of $K$ is

$$M_K(s) = \frac{pe^s}{1 - (1-p)e^s} = \frac{\frac{1}{3}e^s}{1 - \frac{2}{3}e^s}.$$

The transform of $Y$ is

$$M_Y(s) = M_G(s) = \frac{1}{3}e^{-2s} + \frac{1}{2}e^s + \frac{1}{6}e^{3s}.$$

Hence,

$$M_Z(s) = \frac{\frac{1}{3}e^s}{1 - \frac{2}{3}e^s}\Bigg|_{e^s = \frac{1}{3}e^{-2s} + \frac{1}{2}e^s + \frac{1}{6}e^{3s}} = \frac{\frac{1}{3}(\frac{1}{3}e^{-2s} + \frac{1}{2}e^s + \frac{1}{6}e^{3s})}{1 - \frac{2}{3}(\frac{1}{3}e^{-2s} + \frac{1}{2}e^s + \frac{1}{6}e^{3s})}$$

(c) Consider the number of games, $K_3$, that Bob played until his third loss. The random variable $K_3$ is a Pascal random variable, with PMF

$$p_{K_3}(k) = \binom{k-1}{3-1}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{k-3} \qquad k = 3, 4, 5, \ldots$$

In this question, we are interested in another random variable $Z$ defined as the time at which Bob has his third loss. Note that $Z = 2K_3$. Thus,

$$p_Z(z) = \binom{(z/2)-1}{3-1}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{(z/2)-3} \qquad z = 6, 8, 10, \ldots$$

(d) Suppose we observe this gambling process, and let $U$ be a random variable indicating the number of rounds we see until at least one of them wins. Note that $U$ is a geometric random variable with parameter $1 - (1/3)\cdot(1/3) = 8/9$.

Consider another random variable $V$ representing the number of additional rounds we have to observe until the other one wins. If both Alice and Bob win at $U$th round, then $V = 0$. Otherwise, $V$ is geometrically distributed with $p = (1/2) + (1/6) = (2/3)$. Clearly, the number $N$ of rounds until each one of them has won at least once is

$$N = U + V.$$

Thus,

$$\begin{aligned}
\mathbf{E}[N] &= \mathbf{E}[U+V] = \mathbf{E}[U] + \mathbf{E}[V] \\
&= \frac{1}{8/9} + \left(0 \cdot \mathbf{P}(\text{both win} \,|\, \text{at least one wins}) + \frac{1}{2/3} \cdot \mathbf{P}(\text{one wins} \,|\, \text{at least one wins})\right) \\
&= \frac{9}{8} + 0 \cdot \frac{(2/3)\cdot(2/3)}{8/9} + \frac{3}{2} \cdot \frac{2\cdot(1/3)\cdot(2/3)}{(8/9)} \\
&= \frac{9}{8} + \frac{3}{2}\cdot\frac{1}{2} = \frac{15}{8}
\end{aligned}$$

There is another approach to this problem. We partition the sample space into three disjoint events:

$A_1$: both win first round;

$A_2$: only one wins first round;

$A_3$: both lose first round.

By the total expectation theorem,

$$\begin{aligned}
\mathbf{E}[N] &= \mathbf{E}[N \mid A_1]\mathbf{P}(A_1) + \mathbf{E}[N \mid A_2]\mathbf{P}(A_2) + \mathbf{E}[N \mid A_3]\mathbf{P}(A_3) \\
&= 1 \cdot \left(\frac{2}{3}\cdot\frac{2}{3}\right) + \left(1 + \frac{1}{\frac{2}{3}}\right)\cdot\left(2\cdot\frac{1}{3}\cdot\frac{2}{3}\right) + (1 + \mathbf{E}[N])\cdot\left(\frac{1}{3}\cdot\frac{1}{3}\right).
\end{aligned}$$

Solving for $\mathbf{E}[N]$, we obtain $\mathbf{E}[N] = 15/8$.

**Problem 7.** Each night, the probability of a robbery attempt at the local warehouse is 1/5. A robbery attempt is successful with probability 3/4, independent of other nights. After any particular successful robbery, the robber celebrates by taking off either the next 2 or 4 nights (with equal probability), during which time there will be no robbery attempts. After that, the robber returns to his original routine.

(a) Let $K$ be the number of robbery attempts up to (and including) the first successful robbery. Find the PMF of $K$.

(b) Let $D$ be the number of days until (and including) the second successful robbery, including the days of celebration after the first robbery. Find the PMF of $D$, or its transform (whichever you find more convenient).

During a successful robbery, the robber steals a random number of candy bars, which is 1, 2, or 3, with equal probabilities. This number is independent for each successful robbery and independent of everything else. No candy bars are stolen in unsuccessful robberies.

(c) Let $S$ be the number of candy bars collected in two successful robberies. Find the PMF of $S$.

(d) Let $Q$ be the number of candy bars collected in ten robbery attempts (whether successful or not). Find the PMF of $Q$, or its transform, whichever is easier. Find the expectation and the variance of $Q$.

**Solution.** We define the following random variables:

$X$: the number of days up to and including the first successful robbery;

$B$: the number of candy bars stolen during a successful robbery;

$C$: the number of days of rest after a successful robbery.

Note that $X$ is a geometric random variable with parameter 3/20 (the probability of a successful robbery on a given night). Also, it is given that $B$ is uniform over the set $\{1, 2, 3\}$, and that $C$ is uniform over the set $\{2, 4\}$.

(a) Since the probability that any given robbery attempt succeeds is 3/4, the random variable $K$ is geometric with parameter $p = 3/4$. Thus,

$$p_K(k) = (1-p)^{k-1}p = \left(\frac{1}{4}\right)^{k-1}\frac{3}{4} = \frac{3}{4^k}, \qquad \text{for } k = 1, 2, \ldots.$$

(b) We will derive both the PMF and the transform of $D$, the number of days up to and including the second successful robbery.

The PMF of $D$ can be easily found by conditioning on $C$, the number of days the robber rests after the first successful robbery (which only takes on values 2 or 4):

$$p_D(d) = p_{D|C}(d \,|\, C = 2)\mathbf{P}(C = 2) + p_{D|C}(d \,|\, C = 4)\mathbf{P}(C = 4),$$

where

$$p_{D|C}(d \,|\, C = 2) = \begin{cases} \binom{d-3}{1}\left(\frac{3}{20}\right)\left(\frac{17}{20}\right)^{d-4}\left(\frac{3}{20}\right), & \text{if } d = 4, 5, 6, ..., \\ 0, & \text{otherwise.} \end{cases}$$

There must be at least one day before the rest period, since it follows a successful robbery; similarly, there must be at least one day after the rest period. Thus we can view the coefficient in the preceding formula as the number of ways to choose the beginning of a four-day period in a block of $d$ days. Then we multiply the probability of the first success and the probability of $d-4$ failures and finally the probability of the second success at trial $d$. Similarly, we have:

$$p_{D|C}(d \,|\, C=4) = \begin{cases} \binom{d-5}{1}\left(\frac{3}{20}\right)\left(\frac{17}{20}\right)^{d-6}\left(\frac{3}{20}\right), & \text{if } d = 6,7,8..., \\ 0, & \text{otherwise.} \end{cases}$$

Also, $\mathbf{P}(C=2) = \mathbf{P}(C=4) = 1/2$. So plugging into our expression for $p_D(d)$, we get

$$p_D(d) = \begin{cases} \frac{1}{2}\binom{d-3}{1}(\frac{3}{20})^2(\frac{17}{20})^{d-4}, & \text{if } d = 4,5, \\ \frac{1}{2}\binom{d-3}{1}(\frac{3}{20})^2(\frac{17}{20})^{d-4} + \frac{1}{2}\binom{d-5}{1}(\frac{3}{20})^2(\frac{17}{20})^{d-6}, & \text{if } d \geq 6, \\ 0 & \text{otherwise.} \end{cases}$$

Alternatively, we can solve this problem using transforms. Note that $D$ is the sum of three independent random variables: $X_1$, the number of days until the first successful robbery; $C$, the number of days of rest after the first success; and $X_2$, the number of days from the end of the rest period until the second successful robbery. Since $D = X_1 + C + X_2$, and $X_1$, $X_2$, and $C$ are independent, we find that the transform of $D$ is $M_D(s) = [M_X(s)]^2 M_C(s)$. Now, $X_1$ and $X_2$ are (independent) geometric random variables with parameter 3/20 (the probability of a successful robbery). Again, $C$ is equally likely to be 2 or 4. Thus, we conclude that

$$M_D(s) = \big(M_X(s)\big)^2 M_C(s) = \left(\frac{\frac{3}{20}e^s}{1-\frac{17}{20}e^s}\right)^2 \left(\frac{1}{2}e^{2s} + \frac{1}{2}e^{4s}\right).$$

(c) Given a successful robbery, the PMF of $Y$ is $p_Y(y) = 1/3$, for $y = 1,2,3$, and $p_Y(y) = 0$, otherwise. The total number of candybars collected in 2 successful robberies is $S = Y_1 + Y_2$, where $Y_1$ and $Y_2$ are independent and identically distributed with PMF $p_Y(y)$. Therefore, the PMF of $S$ is

$$p_S(s) = \begin{cases} 1/9, & \text{if } s = 2,6, \\ 2/9, & \text{if } s = 3,5, \\ 3/9, & \text{if } s = 4, \\ 0, & \text{otherwise.} \end{cases}$$

(d) Since the probability of a robbery attempt being successful is $\frac{3}{4}$, and since the number of candy bars taken in a successful attempt is equally likely to be 1, 2, or 3, we can view each attempt as resulting in $B$ candy bars, with the following PMF for $B$:

$$p_B(b) = \begin{cases} 1/4, & \text{if } b = 0,1,2,3, \\ 0, & \text{otherwise.} \end{cases}$$

In this case, finding the transform is much easier than finding the PMF. Since $Q$ is the sum of ten independent values of $B$, the transform is simply

$$M_Q(s) = \big(M_B(s)\big)^{10} = \left(\frac{1 + e^s + e^{2s} + e^{3s}}{4}\right)^{10}.$$

Since $Q = B_1 + \cdots + B_{10}$, we have $\mathbf{E}[Q] = \mathbf{E}[B_1] + \cdots + \mathbf{E}[B_{10}] = 10 \cdot \mathbf{E}[B] = 15$. Similarly, using the independence of the $B_i$, we see that $\mathrm{var}(Q) = \mathrm{var}(B_1) + \cdots + \mathrm{var}(B_{10}) = 10 \cdot \mathrm{var}(B) = 10 \cdot (5/4) = 50/4$.

**Problem 8.** A particular medical operation proves fatal in 1% of the cases. Find an approximation to the probability that there will be at least 2 fatalities in 200 operations.

**Solution.** We could find an exact value by using the binomial probability mass function. A reasonable, and much more efficient method is to use the Poisson approximation to the binomial, which tells us that for a binomial random variable with parameters $n$ and $p$, we have:

$$\mathbf{P}(k \text{ successes}) \approx \frac{\lambda^k}{k!}e^{-\lambda}$$

where $\lambda = np$. The desired probability is

$$\begin{aligned}
\mathbf{P}(2 \text{ or more fatalities}) &= 1 - \mathbf{P}(0 \text{ or } 1 \text{ fatality}) \\
&= 1 - \frac{2^0}{0!}e^{-2} - \frac{2^1}{1!}e^{-2} \\
&= .594
\end{aligned}$$

**Problem 9.** You drive to work 5 days a week for a full year (50 weeks), and on any given day, you get a traffic ticket with probability $p = 0.02$, independently of other days. Let $X$ be the total number of tickets you get in the year.

(a) What is the probability that the number of tickets you get is exactly equal to the expected value of $X$?

(b) Calculate approximately the probability in (a), using a Poisson approximation.

(c) The fine for a ticket is \$10, or \$20, or \$50, with respective probabilities 0.5, 0.3, and 0.2, and independently of other tickets. Find the mean and variance of the total amount you pay for traffic tickets during the year?

(d) Suppose you do not know the probability $p$ of getting a ticket, but you got 5 tickets during the year, and you estimate $p$ by the sample mean

$$\hat{P} = \frac{5}{250}.$$

What is the range of possible values of $p$ assuming that the difference between $p$ and the sample mean $\hat{P}$ is within 5 times the standard deviation of the sample mean? (See Example 2.21 in the text for more detail on the use and properties of the sample mean.)

**Solution.** (a) The random variable $X$ has a binomial PMF, with parameters $p = 0.02$ and $n = 250$. The mean is $\mathbf{E}[X] = np = 250 \cdot 0.02 = 5$. The desired probability is

$$\mathbf{P}(X = 5) = \binom{250}{5}(0.02)^5(0.98)^{245} = 0.177.$$

(b) The Poisson approximation has parameter $\lambda = np = 5$, so the probability in (a) is approximated by

$$\tilde{p} = e^{-\lambda}\frac{\lambda^5}{5!} = 0.175.$$

(c) We have

$$\mathbf{E}[Y] = \sum_{i=1}^{250} \mathbf{E}[Y_i],$$

where $Y_i$ is the amount of money you pay on the $i$th day. The PMF of $Y_i$ is

$$P(Y_i = y) = \begin{cases} 0.98, & \text{if } y = 0, \\ 0.01, & \text{if } y = 10, \\ 0.006, & \text{if } y = 20, \\ 0.004, & \text{if } y = 50. \end{cases}$$

The mean of $Y$ is then

$$\mathbf{E}[Y] = 250 \cdot \mathbf{E}[Y_i] = 250 \cdot (0.01 \cdot 10 + 0.006 \cdot 20 + 0.004 \cdot 50) = 250 \cdot 0.42 = 105.$$

Using the independence of the random variables $Y_i$, the variance of $Y$ is

$$\begin{aligned} \text{var}(Y) &= 250 \cdot \text{var}(Y_i) = 250 \cdot \Big(\mathbf{E}[Y_i^2] - \big(\mathbf{E}[Y_i]\big)^2\Big) \\ &= 250 \cdot (0.01 \cdot 10^2 + 0.006 \cdot 20^2 + 0.004 \cdot 50^2 - 0.42^2) = 3305.9. \end{aligned}$$

(d) The variance of the sample mean is

$$\frac{p(1-p)}{250}$$

(cf. Example 2.21 in the text), so assuming that $|p - \hat{P}|$ is within 5 times the standard deviation, the possible values of $p$ are those that satisfy $p \in [0, 1]$ and

$$(p - 0.02)^2 \leq \frac{25p(1-p)}{250}.$$

Solving the quadratic inequality, we see that if $\hat{P}$ is assumed to be within 5 standard deviations from the true mean, then $p$ must be in the range $(0.0029, 0.124)$.

## SECTION 5.2. The Poisson Process

**Problem 10.** A train bridge is constructed across a wide river. Trains arrive at the bridge according to a Poisson process of rate $\lambda = 3$ per day.

(a) If a train arrives on day 0, find the probability that there will be no trains on days 1, 2, and 3.

(b) Find the probability that no trains arrive in the first 2 days, but 4 trains arrive on the 4th day.

(c) Find the probability that it takes more than 2 days for the 5th train to arrive at the bridge.

**Solution.** (a) The probability that there are no trains in 3 days is the probability that the interarrival time is greater than 3, so the probability is: $e^{-9} = .000123$.

(b) The events that no trains arrive in the first 2 days, and 4 arrive on the 4th day are independent, by the properties of the Poisson process. Therefore the probability of both events occurring is the product of their probabilities, which is:

$$e^{-6} \cdot e^{-3} \frac{3^4}{4!} = 1.543 \cdot 10^{-5}$$

(c) The probability that the 5th train has not arrived by the second day is the probability that the sum of the first 5 interarrival times is greater than 2, which equals the probability that there are at most 4 arrivals in the first 2 days, thus giving:

$$\mathbf{P}(\text{at most 4 arrivals in 2 days}) = e^{-6}\left(1 + 6 + \frac{6^2}{2!} + \frac{6^3}{3!} + \frac{6^4}{4!}\right) = .2851.$$

**Problem 11.** An amateur criminal is contemplating shoplifting from a store. Police officers walk by the store according to a Poisson process of rate $\lambda$ per minute. If an officer walks by while the crime is in progress, the criminal will be caught.

(a) If it takes the criminal $t$ seconds to commit the crime, find the probability that the criminal will be caught.

(b) Repeat part (a) under the new assumption that the criminal will only be caught if two police officers happen to walk by while the crime is in progress.

**Solution.** (a) The criminal will be caught if the first officer comes by in fewer than $t$ seconds. Since the time until the first arrival is exponentially distributed, the desired probability is $1 - e^{-t\lambda}$.

(b) We are interested in the probability that the second arrival occurs before time $t$. By integrating the Erlang PDF of order 2, this probability is $1 - e^{-\lambda t}(\lambda t + 1)$.

**Problem 12.** The MIT soccer team needs at least 8 players to avoid forfeiting a game. Assume that each player has some chance of getting injured for the season, but her playing lifetime for a given season is exponentially distributed with parameter $\lambda$. For simplicity, assume that the coach insists on only playing 8 players at a time, and then replaces a player as soon as she gets hurt. Find:

(a) The expected time until the first substitution.

(b) The distribution of total time the team can play in a season, given that there are $n$ women on the team.

**Solution.** (a) We may view the time until a particular player is injured as the time until the first arrival in a Poisson process of rate $\lambda$. Since each player is independent, and since we have 8 players, we have 8 independent Poisson processes of rate $\lambda$. Thus, we may view the time until any player is injured as the time until the first arrival in a merged Poisson process, which has rate $8\lambda$. The expected time until the first arrival is therefore

$$\frac{1}{8 \cdot \lambda}.$$

(b) The time until a next injury is exponential with rate $8\lambda$. The time $T$ until a game is forfeited is the time until there are 7 players left, i.e., the time until there are $n - 7$ injuries. This time has an Erlang distribution of order $n - 7$ and parameter $8\lambda$. Thus,

$$f_T(t) = \frac{(8\lambda)^{n-7} t^{(n-7)-1} e^{-8\lambda t}}{((n-7)-1)!}, \qquad t \geq 0.$$

**Problem 13.** A certain police officer stops cars for speeding. The number of red sports cars she stops in one hour is a Poisson process with rate 4, while the number of other cars she stops is a Poisson process with rate 1. Assume that these two processes are independent of each other. Find the probability that this police officer stops at least 2 ordinary cars before she stops 3 red sports cars.

**Solution.** The process of stopping cars is a Poisson process, obtained by merging two independent Poisson processes Each time the police officer stops a car, there is a $4/(2+4) = 2/3$ probability that it will be a red sports car. Therefore the probability she stops at least 2 ordinary cars before stopping 2 sports cars is

$$\begin{aligned} &1-\mathbf{P}(\text{she stops 3 sports cars before stopping 2 other cars}) \\ &\quad = 1 - \binom{4}{1} = \left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^3 = \frac{49}{81}. \end{aligned}$$

**Problem 14.** Consider two independent Poisson processes, with arrival rates $\alpha$ and $\beta$, respectively. Determine:

(a) The probability $q$ that the next three arrivals come from the same process.

(b) The PMF of $N$, the number of arrivals from the first process that occur before the fourth arrival from the second process.

**Solution.** (a) Whenever an arrival occurs, it comes from the first Poisson process with probability $\alpha/(\alpha+\beta)$, and this is independent from one arrival to the next. Therefore, the probability that all three arrivals come from the same process is equal to

$$\left(\frac{\alpha}{\alpha+\beta}\right)^3 + \left(\frac{\beta}{\alpha+\beta}\right)^3.$$

(b) The probability that $N = n$ is the probability that out of the first $n+3$ arrivals, there were exactly $n$ coming from the first process and that the $(n+4)$th arrival was from process 2. Thus,

$$p_N(n) = \binom{n+3}{3}\left(\frac{\beta}{\alpha+\beta}\right)^4\left(\frac{\alpha}{\alpha+\beta}\right)^n, \qquad n = 0, 1, \ldots$$

**Problem 15.** Suppose the waiting time until the next bus at a particular bus stop is exponentially distributed with parameter $\lambda = 1/15$. Suppose that a bus pulls out just as you arrive at the stop. Find the probability that:

(a) You wait more than 15 minutes for a bus.

(b) You wait between 15 and 30 minutes for a bus.

(c) What are the probabilities in (a) and (b) assuming the bus left 5 minutes before you arrive?

**Solution.** (a) The probability that you wait more than 15 minutes is

$$\int_{15}^{\infty} \frac{1}{15} e^{-x/15}\, dx = -e^{-x/15}\Big|_{15}^{\infty} = e^{-1}.$$

(b) The probability that you wait between 15 and 30 minutes is

$$\int_{15}^{30} \frac{1}{15} e^{-x/15}\, dx = -e^{-x/15}\Big|_{15}^{30} = e^{-1} - e^{-2}.$$

(c) The same as in (a) and (b), because of the memoryless property.

**Problem 16.** A phone at a telephone exchange rings according to a Poisson process of rate $\lambda$. If 3 calls arrive in the first ninety minutes, find:

(a) The probability that all 3 calls arrived in the first 30 minutes.

(b) The probability that at least one arrived in the first 30 minutes.

**Solution.** (a) The arrival time of each of the three calls is uniformly distributed in the interval of 90 minutes (see Problem 16 in Chapter 5 of the text). Furthermore, the three arrival times are independent of each other. This follows intuitively from the definition of the Poisson process: given that there was an arrival at some particular time, this gives us no information on what may have happened at other times. Therefore the probability that all three occur within the first 30 minutes is: $(1/3)^3 = 1/27$.

(b) The probability that at least one ocurs in the first 30 minutes is, by the same reasoning as above, $1 - (8/27) = 19/27$.

**Problem 17.** The time to finish a problem set is exponentially distributed with parameter $\lambda = 1/2$.

(a) Find the probability that a particular problem set takes more than 2 hours to finish.

(b) Given that you have been working on a problem set already for 7 hours, find the probability that this problem set will take more than 9 hours total (i.e., two hours more).

**Solution.** (a) This probability is

$$\int_{2}^{\infty} \frac{1}{2} e^{-x/2}\, dx = 1 - e^{-1}.$$

(b) Using the memorylessness property of the exponential distribution, we see that the conditional probability that the problem set will take more than 9 hours given that it has already taken 7, is just the answer to part (a).

**Problem 18.** Based on your understanding of the Poisson process, determine the numerical values of $a$ and $b$ in the following expression and explain your reasoning.

$$\int_t^\infty \frac{\lambda^6 \tau^5 e^{-\lambda\tau}}{5!}\,d\tau = \sum_{k=a}^{b} \frac{(\lambda t)^k e^{-\lambda t}}{k!}.$$

**Solution.** The left-hand side is the probability that an Erlang random variable of order 6 and rate $\lambda$ is larger than $t$, i.e., the probability of at most 5 arrivals over an interval of length $t$. The right-hand side is the probability that the number of arrivals in a Poisson process with rate $\lambda$, over an interval of length $t$, is between $a$ and $b$ (inclusive). Thus, $a = 0$ and $b = 5$.

**Problem 19.** **[D]** A woman is seated beside a conveyor belt, and her job is to remove certain items from the belt. She has a narrow line of vision and can get these items only when they are right in front of her. She has noted that the probability that exactly $k$ of her items will arrive in a minute is given by

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

and she assumes that the arrivals of her items constitute a Poisson process.

(a) If she wishes to sneak out to have a beer but will not allow the expected value of the number of items she misses to be greater than 5, how much time may she take?

(b) If she leaves for two minutes, what is the probability that she will miss exactly two items the first minute and exactly one item the second minute?

(c) If she leaves for two minutes, what is the probability that she will miss a total of exactly three items?

(d) The union has installed a bell which rings once a minute with precisely one-minute intervals between gongs. If, between two successive gongs, more than three items come along the belt, she will handle only three of them properly and will destroy the rest. Under this system, what is the expected fraction of items that will be destroyed?

**Solution.** (a) Item arrivals are a Poisson process with parameter $\lambda = 2$ per minute. The expected number of arrivals in $t$ minutes is equal to $2t$. For that number to be no larger than 5, we need $t \leq 2.5$.

(b) Using independence of the arrivals in different one-minute intervals, the desired probability is equal to

$$p_K(2)p_K(1) = \frac{2^3 e^{-4}}{2}.$$

(c) The number of arrivals in a two-minute interval has a Poisson PMF with parameter $2\lambda = 4$. Therefore, the probability of 3 arrivals in 2 minutes is

$$\frac{4^3 e^{-4}}{3!}.$$

(d) The expected number of processed items in a one-minute interval is equal to

$$\begin{aligned}
p_K(1) + 2p_K(2) + 3\big(p_K(3) + p_K(4) + \cdots\big) \\
&= p_K(1) + 2p_K(2) + 3 - 3p_K(0) - 3p_K(1) - 3p_K(2) \\
&= 3 - 3p_K(0) - 2p_K(1) - p_K(2) \\
&= 3 - 9e^{-2}.
\end{aligned}$$

Therefore, the expected number of unprocessed items is equal to

$$\lambda - 3 + 9e^{-2} = 9e^{-2} - 1.$$

Since $\lambda = 2$, this represents a fraction $(9e^{-2} - 1)/2$ of the total.

**Problem 20.** **[D]** Arrivals of certain events at points in time are known to constitute a Poisson process, but it is not known which of two possible values of $\lambda$, the average arrival rate, describes the process. Our a priori estimate is that $\lambda = 2$ or $\lambda = 4$ with equal probability. We observe the process for $t$ units of time and observe exactly $k$ arrivals. Given this information, determine the conditional probability that $\lambda = 2$. Check to see whether or not your answer is reasonable for some simple limiting values for $k$ and $t$.

**Solution.** We have a Poisson process with an average arrival rate $\lambda$ which is equally likely to be either 2 or 4. Thus,

$$\mathbf{P}(\lambda = 2) = \mathbf{P}(\lambda = 4) = \frac{1}{2}.$$

We observe the process for $t$ time units and observe $k$ arrivals. The conditional probability that $\lambda = 2$ is, by definition

$$\mathbf{P}(\lambda = 2 \,|\, k \text{ arrivals in time } t) = \frac{\mathbf{P}(\lambda = 2 \text{ and } k \text{ arrivals in time } t)}{\mathbf{P}(k \text{ arrivals in time } t)}.$$

Now, we know that

$$\begin{aligned}\mathbf{P}(\lambda = 2 \text{ and } k \text{ arrivals in time } t) &= \mathbf{P}(k \text{ arrivals in time } t \mid \lambda = 2) \cdot \mathbf{P}(\lambda = 2)\\ &= \frac{(2t)^k e^{-2t}}{k!} \cdot \frac{1}{2}.\end{aligned}$$

Similarly,

$$\mathbf{P}(\lambda = 2 \text{ and } k \text{ arrivals in time } t) = \frac{(4t)^k e^{-4t}}{k!} \cdot \frac{1}{4}.$$

Thus,

$$\begin{aligned}\mathbf{P}(\lambda = 2 \mid k \text{ arrivals in time } t) &= \frac{\frac{(2t)^k e^{-2t}}{k!}\left(\frac{1}{2}\right)}{\frac{(2t)^k e^{-2t}}{k!}\left(\frac{1}{2}\right) + \frac{(4t)^k e^{-4t}}{k!}\left(\frac{1}{2}\right)}\\ &= \frac{(2t)^k e^{-2t}}{(2t)^k e^{-2t} + (4t)^k e^{-4t}}\\ &= \frac{1}{1 + 2^K e^{-2T}}.\end{aligned}$$

To check whether this answer is reasonable, suppose $t$ is large and $k = 2t$ (observed arrival rate equals 2). Then, $\mathbf{P}(\lambda = 2 \,|\, k \text{ arrivals in time } t)$ approaches 1 as $t$ goes to $\infty$. Similarly, if $t$ is large and $k = 4t$ (observed arrival rate equals 4), then, $\mathbf{P}(\lambda = 2 \,|\, k \text{ arrivals in time } t)$ approaches 0 as $t$ goes to $\infty$.

**Problem 21.** **[D]** Let $K_1, , K_2, \ldots$ be independent identically distributed geometric random variables. Random variable $R_i$ is defined by

$$R_i = \sum_{j=1}^{i} K_i, \qquad i = 1, 2, \ldots.$$

If we eliminate arrivals number $R_1, R_2, \ldots$ in a Poisson process, do the remaining arrivals constitute a Poisson process?

**Solution.** We can think of the $R_i$ as the trial numbers at which a Bernoulli process registers a success. Eliminating arrivals numbered $R_i$ is the same as using independent Bernoulli trials to decide which arrivals are to be eliminated. As discussed in the text, the resulting process is Poisson.

**Problem 22.** Determine, in an efficient manner (without using integration by parts), the fourth moment of a continuous random variable described by the probability density function

$$f_X(x) = \frac{4^3 x^2 e^{-4x}}{2}, \qquad x \geq 0.$$

**Solution.** We have

$$\mathbf{E}(X^4) = \int_0^\infty x^4 f_X(x) dx = \int_0^\infty \frac{4^3 x^6 e^{-4x}}{2}\, dx = \frac{6!}{4^4 \cdot 2} \int_0^\infty \frac{4^7 x^6 e^{-4x}}{6!}\, dx = \frac{6!}{4^4 \cdot 2}.$$

To see that the integral in the final step evaluates to 1, notice that the integrand corresponds to an Erlang PDF of order seven.

**Problem 23.** A room has two lamps that use bulbs of type A and B, respectively. The lifetime, $X$, of any particular bulb of a particular type is a random variable, independent of everything else, with the following PDF:

$$\text{for type-A Bulbs: } f_X(x) = \begin{cases} e^{-x}, & \text{if } x \geq 0, \\ 0, & \text{otherwise;} \end{cases}$$

$$\text{for type-B Bulbs: } f_X(x) = \begin{cases} 3e^{-3x}, & \text{if } x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Both lamps are lit at time zero. Whenever a bulb is burned out it is immediately replaced by a new bulb.

(a) What is the expected value of the number of type-B bulb failures until time $t$?

(b) What is the PDF of the time until the first failure of *either* bulb type?

(c) Find the expected value and variance of the time until the third failure of a type-B bulb.

(d) Suppose that a type-A bulb has just failed. How long do we expect to wait until a subsequent type-B bulb failure?

**Solution.** (a) This is the expected number of "arrivals" in a Poisson process with parameter 3, and is equal to $3t$.

(b) The process of bulb failures of either bulb type is obtained by merging two Poisson processes. It is therefore Poisson with parameter $1 + 3 = 4$. The time until the first failure of either bulb type is exponentially distributed with parameter 4 and its PDF is

$$4e^{-4t}, \qquad t \geq 0.$$

(c) The time until the third failure of a type-B bulb is an Erlang random variable of order 3, with parameter 3. Therefore, its mean and variance is equal to 1 and 1/3, respectively.

(d) Using the fresh-start property, the time until the first subsequent type-B bulb failure is exponentially distributed with parameter 3, and its expected value is 1/3.

**Problem 24.** [**D**] Dave is taking a multiple-choice exam. You may assume that the number of questions is infinite. Simultaneously, but independently, his conscious and subconscious faculties are generating answers for him, each in a Poisson manner. (His conscious and subconscious are always working on different questions.) Conscious responses are generated at a rate of $\lambda_c$ responses per minute. Subconscious responses are generated at a rate of $\lambda_s$ responses per minute. Each conscious response is an independent Bernoulli trial with probability $p_c$ of being correct. Similarly, each subconscious response is an independent Bernoulli trial with probability $p_s$ of being correct. Dave responds only once to each question, and you can assume that his time for recording these conscious and subconscious responses is negligible.

(a) Determine $p_K(k)$, the probability mass function for the number of conscious responses Dave makes in an interval of $t$ minutes.

(b) If we pick any question to which Dave has responded, what is the probability that his answer to that question:

  (i) Represents a conscious response.

  (ii) Represents a conscious correct response.

(c) If we pick an interval of $t$ minutes, what is the probability that in that interval Dave will make exactly $r$ conscious responses *and* exactly $s$ subconscious responses?

(d) Determine the transform for the probability density function for random variable $X$, where $X$ is the time from the start of the exam until Dave makes his first conscious response which is preceded by at least one subconscious response.

(e) Determine the probability mass function for the total number of responses up to and including his third conscious response.

(f) The papers are to be collected as soon as Dave has completed exactly $n$ responses. Determine:

  (i) The expected number of questions he will answer correctly

  (ii) The probability mass function for $L$, the number of questions he answers correctly.

(g) Repeat part (f) for the case in which the exam papers are to be collected at the end of a fixed interval of $t$ minutes.

**Solution.** (a) The random variable $K$ has a Poisson PMF with parameter $\lambda_c t$. Therefore, for $t \geq 0$,

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

(b) We have

$$\mathbf{P}(\text{conscious response}) = \frac{\lambda_c}{\lambda_c + \lambda_s},$$

$$\begin{aligned}\mathbf{P}(\text{conscious correct response}) &= \mathbf{P}(\text{conscious resp}) \cdot \mathbf{P}(\text{correct resp} \mid \text{conscious resp}) \\ &= \frac{\lambda_c}{\lambda_c + \lambda_s} p_c.\end{aligned}$$

(c) Since the two conscious and subconscious responses are generated independently, we have

$$\begin{aligned}
&\mathbf{P}(r \text{ conscious responses and } s \text{ subconscious responses}) \\
&\qquad = \mathbf{P}(r \text{ conscious responses }) \cdot \mathbf{P}(s \text{ unconscious responses}) \\
&\qquad = \frac{(\lambda_c t)^r e^{-\lambda_c t}}{r!} \cdot \frac{(\lambda_s t)^s e^{-\lambda_s t}}{s!}.
\end{aligned}$$

(d) Let $X_s$ be the time from the start of the exam to the time of the first subconscious response, and $X_c$ be the time from the first subconscious response to the time of the next conscious response. Then $X = X_s + X_c$ and

$$M_X(s) = M_{X_s}(s) \cdot M_{X_c}(s) = \left(\frac{\lambda_s}{\lambda_s - s}\right)\left(\frac{\lambda_c}{\lambda_c - s}\right).$$

(e) Let "success" indicate a conscious response and "failure" indicate a subconscious response. Then

$$\mathbf{P}(\text{ success}) = \frac{\lambda_c}{\lambda_c + \lambda_s}, \quad \mathbf{P}(\text{ failure}) = \frac{\lambda_s}{\lambda_c + \lambda_s}.$$

The total number of responses $N$ up to and including the third success is the time until the third arrival in a Bernoulli process, and is described by a Pascal PMF, so that

$$p_N(n) = \binom{n-1}{2}\left(\frac{\lambda_c}{\lambda_c + \lambda_s}\right)^3 \left(\frac{\lambda_s}{\lambda_c + \lambda_s}\right)^{n-3}, \quad n = 3, 4, \ldots.$$

(f) The probability $p$ that a particular question is answered correctly is

$$p = \frac{\lambda_c}{\lambda_c + \lambda_s} p_c + \frac{\lambda_s}{\lambda_c + \lambda_s} p_s.$$

The expected number of questions answered correctly is $np$. Each question answered can be viewed as an independent Bernoulli trial, with probability $p$ of being successful. Thus, the probability of $l$ correct answers is

$$p_L(l) = \binom{n}{l}\left(\frac{\lambda_c}{\lambda_c + \lambda_s} p_c + \frac{\lambda_s}{\lambda_c + \lambda_s} p_s\right)^l \left(1 - \left(\frac{\lambda_c}{\lambda_c + \lambda_s} p_c + \frac{\lambda_s}{\lambda_c + \lambda_s} p_s\right)\right)^{n-l},$$

for $l = 0, 1, \ldots, n$.

(g) Correct answers arrive as a Poisson process with rate $\lambda_c p_c + \lambda_s p_s$. Therefore, the expected number of correct answers is $(\lambda_c p_c + \lambda_s p_s)t$, and the probability of $l$ correct answers is

$$p_L(l) = \frac{(\lambda t)^l e^{-\lambda T}}{l!}, \quad l = 0, 1, \ldots,$$

where $\lambda = \lambda_c p_c + \lambda_s p_s$.

**Problem 25.** There are two types of calls to the MIT Campus Patrol. Type A calls (distress calls) arrive as a Poisson process with rate $\lambda_A$. Type B calls (professors who have lost their keys) arrive as an independent Poisson process with rate $\lambda_B$. Let us fix $t$ to be 12 o'clock.

(a) What is the expected length of the interval that $t$ belongs to? (That is, the interval from the last event before $t$ until the first event after $t$.)

(b) What is the probability that $t$ belongs to an AA interval? (That is, the first event before, as well as the first event after time $t$ are both of type A.)

(c) Let $c$ be a constant. What is the probability that between $t$ and $t+c$, we have exactly two events, one of type $A$, followed by one of type $B$?

**Solution.** (a) Events arrive as a Poisson process with rate $\lambda = \lambda_A + \lambda_B$. The expected time until the next event is $1/\lambda$, and by looking at the process backwards (and noting that a Poisson process run backwards is also Poisson), the expected time since the last event is also $1/\lambda$. Thus, the answer is $2/\lambda$.

(b) Each event is of type $A$ independent of any other event, and with probability $\lambda_A/\lambda$. Thus, the answer is $\lambda_A^2/\lambda^2$.

(c) This is the probability that there are exactly two arrivals in a Poisson process with rate $\lambda$, which is $e^{-\lambda c}(\lambda c)^2/2$, times the probability $\lambda_A\lambda_B/\lambda^2$ that the arrivals are of the specified types.

**Problem 26.** **[D]** The interarrival times for cars passing a checkpoint are independent random variables with PDF

$$f_T(t) = \begin{cases} 2e^{-2t}, & t > 0, \\ 0, & \text{otherwise}, \end{cases}$$

where the interarrival times are measured in minutes. The successive values of the durations of these interarrival times are recorded on small computer cards. The recording operation occupies a negligible time period following each arrival. Each card has space for three entries. As soon as a card is filled, it is replaced by the next card.

(a) Determine the mean and the third moment of the interarrival times.

(b) Given that no car has arrived in the last four minutes, determine the PMF for random variable $K$, the number of cars to arrive in the next six minutes.

(c) Determine the PDF, the expected value, and the transform for the total time required to use up the first dozen computer cards.

(d) Consider the following two experiments:

  (i) Pick a card at random from a group of completed cards and note the total time, $U$, the card was in service. Find $\mathbf{E}[U]$ and $\sigma_U^2$.

  (ii) Come to the corner at a certain time. When the card in use at the time of your arrival is completed, note the total time it was in service (the time from the start of its service to its completion). Call this time $V$. Determine $\mathbf{E}[V]$ and $\sigma_V^2$.

(e) Given that the computer card presently in use contains exactly two entries and also that it has been in service for exactly 0.5 minute, determine the PDF for the remaining time until the card is completed.

**Solution.** (a) An interarrival time $T$ is an exponential random variable with parameter $\lambda = 2$. Therefore, $\mathbf{E}[T] = 1/\lambda = 1/2$, and

$$\mathbf{E}[T^3] = \int_0^\infty t^3 2e^{-2t} dt = \frac{(2\cdot 3)}{2^3} \int_0^\infty \frac{2^4 t_0{}^3 e^{-2t}}{3!} dt = \frac{3}{4}.$$

The last equality was obtained because the integrand is an Erlang PDF of order 4, and therefore this integral is equal to 1.

(b) The Poisson process is memoryless, and thus the history of events in the previous 4 minutes does not affect $p_K(k)$. So, the number of arrivals in the next 6 minutes corresponds to the number of arrivals in a Poisson process with rate $\lambda = 2$,and

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

(c) Let $D$ denote the total time to use up the first dozen computer cards, so that

$$D = X_1 + X_2 + X_3 + \cdots + X_{36},$$

where the $X_i$ are independent exponentially distributed random variables with parameter $\lambda = 2$. Therefore $D$ has an Erlang PDF of order 36, and

$$f_D(d) = \frac{(2)^{36}(d)^{35}e^{-2d}}{35!}, \quad d \geq 0, \qquad \mathbf{E}[D] = 36\mathbf{E}[T_1] = 18 \qquad M_D(s) = \left(\frac{2}{s+2}\right)^{36}.$$

(d)(i) Choosing a card at random is the same as choosing some integer $N$ and recording the sum of the interarrival times of cars $N$, $N + 1$, $N + 2$. The selection of $N$ is independent of the interarrival times of the various cars. Therefore, the conditional distribution of the sum of these interarrival times is the same for every value $n$ of $N$, and is an Erlang PDF of order 3. It follows that the expected value is $3/\lambda = 3/2$ and the variance is $3/\lambda^2 = 3/4$.

(d)(ii) For the particular card that was picked, let $S_0$ be the time that this card started service, and let $S_1$, $S_2$, $S_3$ be the arrival times within the service operation, with service ending at time $S_3$. Let $t$ be the time that you arrived to observe the process.

Suppose that $S_0 < t < S_1$. Then, the service time of a card can be split into intervals of length $t - S_0$, $S_1 - t$, $S_2 - S_1$, $S_3 - S_1$. The length of each one of these intervals is an independent exponentially distributed random variable. (See the discussion of random incidence in Section 5.3, and Problem 24 in the text.) Therefore, the service time $S_3 - S_0$ has an Erlang PDF of order 4. For the other possible cases, i.e., $S_0 < S_1 < t < S_2 < S_3$ and $S_0 < S_1 < S_2 < t < S_3$, the same conclusion is reached. it follows that the mean of $S_3 - S_0$ is $4/\lambda = 4/2 = 2$ and its variance is $4/\lambda^2 = 1$.

(e) The Poisson process is memoryless so any event during some interval of time is independent of all events within a previous *non-overlapping* interval of time. Therefore, the remaining time until the next arrival (which is also the time that the card is completed) is still described by the same exponential PDF, $2e^{-2t}$, for $t \geq 0$.

**Problem 27.**

(a) Shuttles depart from New York to Boston every hour on the hour. Passengers arrive according to a Poisson process of rate $\lambda$ per hour. Find the expected number of passengers on a shuttle. (Ignore issues of limited seating.)

(b) Now, and for the rest of this problem, suppose that the shuttles are not operating on a deterministic schedule, but rather their interdeparture times are exponentially distributed with rate $\mu$ per hour, and independent of the process of passenger arrivals. Find the PMF of the number shuttle departures in one hour.

(c) Let us define an "event" in the airport to be either the arrival of a passenger, or the departure of a plane. Find the expected number of "events" that occur in one hour.

(d) If a passenger arrives at the gate, and sees $2\lambda$ people waiting, find his/her expected time to wait until the next shuttle.

(e) Find the PMF of the number of people on a shuttle.

**Solution.** (a) The number of people that arrive within an hour is Poisson-distributed with parameter $\lambda$, and its expected value is $\lambda$.

(b) If the interarrival times for the shuttles are exponentially distributed, then shuttle departures form a Poisson process of rate $\mu$. Thus, the number of departures in one hour has a Poisson PMF with parameter $\mu$.

(c) Here, we are merging two independent Poisson processes, which results in a Poisson process of rate $\mu + \lambda$. Therefore, the expected number of "events" occurring in one hour will be $\mu + \lambda$.

(d) The number of people waiting conveys some information on the time since the last departure. On the other hand, because of memorylessness of the exponential distribution, this number is independent from the time until the next departure. Thus, the expected waiting time is just $1/\mu$, irrespective of how many people are waiting.

(e) This is essentially the same problem as Problem 21 in the text. Every event at the airport has probability $\lambda/(\lambda+\mu)$ of being a passenger arrival ("failure") and probability $\mu/(\lambda + \mu)$ of being a shuttle departure ("success"). Furthermore, different events are independent. The number of passengers on a shuttle is the number of failures until the first success and is distributed as $K - 1$, where $K$ is a geometric random variable with parameter $\mu/(\lambda + \mu)$. Thus, the PMF of the number of people on the shuttle is

$$\left(\frac{\lambda}{\lambda+\mu}\right)^k \left(\frac{\mu}{\lambda+\mu}\right), \qquad k = 0, 1, \ldots$$

**Problem 28.** Type A, B, and C items are placed in a common buffer, each type arriving as part of an independent Poisson process with average arrival rates, respectively, of $a$, $b$, and $c$ items per minute.

For the first four parts of this problem, assume the buffer is discharged immediately whenever it contains a total of ten items.

(a) What is the probability that, of the first ten items to arrive at the buffer, only the first and one other are type A?

(b) What is the probability that any particular discharge of the buffer contains five times as many type A items as type B items?

(c) Determine the PDF, expectation, and variance of the total time between consecutive discharges of the buffer.

(d) Determine the probability that during a particular five minute interval there exactly two arrivals of each type.

For the rest of this problem, a different rule is used for discharging the buffer: namely, the buffer is discharged immediately whenever it contains a total of three type A items.

(e) Determine the PDF, expectation, and variance of the total time between consecutive discharges of the buffer.

(f) For an observer arriving at a random time, long after the process began, obtain the PDFs of:

(i) $U$, the time until the arrival of the next item at the buffer input

(ii) $V$, the time until the next discharge of the buffer

**Solution.** (a) The three independent Poisson processes can be merged into one Poisson process with an arrival rate of $a+b+c$ items per minute. Any arrival in the merged process has probability $\mathbf{P}(A) = a/(a+b+c)$ of coming from process A, independently of other arrivals. Thus, for each arrival we have an independent Bernoulli trial that determines whether it is type A or not. Hence the probability that the first item is type A, and exactly one of the next items is type A, is

$$\mathbf{P}(A)\binom{9}{1}\mathbf{P}(A)(1-\mathbf{P}(A))^8 = \binom{9}{1}\mathbf{P}(A)^2(1-\mathbf{P}(A))^8.$$

(b) The are two ways for there to be 5 times as many type A items as type B items. One possibility is that there are 5 type A items, 1 type B item, and 4 type C items. The probability of this event is

$$\frac{10!}{5!\,1!\,4!}\mathbf{P}(A)^5\mathbf{P}(B)\mathbf{P}(C)^4,$$

where

$$\mathbf{P}(A) = \frac{a}{a+b+c}, \qquad \mathbf{P}(B) = \frac{b}{a+b+c}, \qquad \mathbf{P}(C) = \frac{c}{a+b+c}.$$

The other possibility is that there are 0 type A items, 0 type B items, and 10 type C items. The probability of this event is $\mathbf{P}(C)^{10}$. The desired probability is obtained by adding the probabilities of the two possibilities.

(c) The total time between consecutive discharges is an Erlang random variable of order 10 and parameter $a+b+c$, with PDF

$$p_T(t) = \frac{(a+b+c)^{10} t^9 e^{-(a+b+c)t}}{9!}, \qquad t \geq 0.$$

The mean and variance are $\mathbf{E}[T] = 10/(a+b+c)$ and $\text{var}(T) = 10/(a+b+c)^2$ respectively.

(d) The probability of exactly two arrivals of type A in 5 minutes is found from the Poisson PMF and is $(5a)^2 e^{-5a}/2$, and similarly for the other types. Since the different types correspond to independent Poisson processes, the probability of interest is the product

$$\frac{(5a)^2 e^{-5a}}{2} \cdot \frac{(5b)^2 e^{-5b}}{2} \cdot \frac{(5c)^2 e^{-5c}}{2} = \frac{(125abc)^2 e^{-5(a+b+c)}}{8}.$$

(e) The total time between consecutive discharges is an Erlang random variable of order 3 and parameter $a$, with PDF

$$\frac{a^3 t^2 e^{-at}}{2!}, \qquad t \geq 0.$$

The mean and variance are $3/a$ and $3/a^2$, respectively.

(f)(i) Since the arrival of items is a Poisson process, the PDF for the time until the next arrival is an exponential PDF with parameter $a+b+c$. Hence,

$$p_U(u) = (a+b+c)e^{-(a+b+c)u}, \qquad u \geq 0.$$

(f)(ii) There is an equal probability that the observer arrives before the first arrival of a type A item, between the first and second arrivals of type A items, or between the second and third arrivals of type A items. We thus use the "divide and conquer" strategy.

If the observation is made before the first type A arrival, then $V$ is a third order Erlang random variable with parameter $a$. If the observation is made between the first and second type A arrivals, then $V$ is a second order Erlang random variable with parameter $a$. If the observation is made after the second type A arrival, then $V$ is an exponential random variable with parameter $a$. Thus,

$$p_V(v) = \frac{1}{3}\frac{a^3 v^2 e^{-av}}{2!} + \frac{1}{3}a^2 v e^{-av} + \frac{1}{3}a e^{-av}, \qquad v \geq 0.$$

**Problem 29.** Let $T_1, T_2$ (respectively, $S$) be exponential random variables with parameter $\lambda$ (respectively, $\mu$). We assume that all three of these random variables are independent. Derive an expression for the expected value of $\min\{T_1 + T_2, S\}$.

**Solution.** We view the random variables $T_1$, $T_2$, and $S$ as interarrival times in two independent Poisson processes with rates $\lambda$ and $\mu$, respectively. 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.

The expected time until the first arrival is $1/(\lambda+\mu)$. With probability $\mu/(\lambda+\mu)$ this arrival comes from the second process and we are done. If it comes from the first process, we have to wait until an arrival from either process. The expected additional waiting time is $1/(\lambda + \mu)$. Using the total expectation theorem, we obtain

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

**Problem 30.** Let $Y$ be exponentially distributed with parameter $\lambda_1$. Let $Z_k$ be Erlang of order $k$. with parameter $\lambda_2$. Assume that $Y$ and $Z_k$ are independent. Let $M_k = \max\{Y, Z_k\}$. Find a recursive formula for $\mathbf{E}[M_k]$, in terms of $E[M_{k-1}]$.

**Solution.** We proceed as in Problem 17 in the text. Consider two independent Poisson processes with rates $\lambda_1$ and $\lambda_2$, respectively. We interpret $Y$ as the first arrival time in the first process and $Z_k$ as the $k$th arrival time in the second process. Let $T$ be the first time at which one of the processes registers an arrival. Since the merged process is Poisson with rate $\lambda_1 + \lambda_2$, we have $\mathbf{E}[T] = 1/(\lambda_1 + \lambda_2)$. There are two cases to consider.

(i) The arrival at time $T$ comes from the first process; this happens with probability $\lambda_1/(\lambda_1 + \lambda_2)$. In this case, we have to wait an additional time $S$ until the second process registers $k$ arrivals. This additional time $S$ is Erlang of order $k$, with parameter $\lambda_2$, and its expected value is $k/\lambda_2$.

(ii) The arrival at time $T$ comes from the second process; this happens with probability $\lambda_2/(\lambda_1 + \lambda_2) = 1/3$. In this case, the additional time $S$ we have to wait is the time until the first process registers an arrival and the second process registers $k - 1$ arrivals. Thus, $\mathbf{E}[S] = \mathbf{E}[M_{k-1}]$. Putting everything together, we have

$$\mathbf{E}[\max\{Y, Z\}] = \frac{1}{\lambda_1 + \lambda_2} + \frac{\lambda_1}{\lambda_1 + \lambda_2} \cdot \frac{k}{\lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2} \cdot \mathbf{E}[M_{k-1}],$$

which is the desired recursion.

**Problem 31.** Consider the random variable $Z = X - Y$, where $X$ and $Y$ are independent and exponentially distributed with parameter $\lambda$.

(a) Find the PDF of $Z$ by conditioning on the events $\{X \leq Y\}$ and $\{X \geq Y\}$, and using an interpretation in terms of Poisson arrivals.

(b) Repeat part (a) for the case where $X$ and $Y$ are independent and exponentially distributed, but with different parameters $\lambda_X$ and $\lambda_Y$.

**Solution.** (a) We interpret $X$ and $Y$ as the first arrival times in two independent Poisson processes with the same rate $\lambda$. Let $Z = Y - X$. With probability 1/2, the first process has the first arrival ($X < Y$). Then, $Y - X$ is the remaining time until the second process registers an arrival. By the fresh-start property, $Z = Y - X$ has an exponential distribution:

$$f_{Z|\{X \leq Y\}}(z) = e^{-\lambda z}, \qquad z \geq 0.$$

By a symmetrical argument, with probability 1/2, the first arrival comes from the second process and $-Z = X - Y$ has an exponential distribution:

$$f_{Z|\{X \geq Y\}}(z) = e^{\lambda z}, \qquad z \leq 0.$$

Thus,

$$f_Z(z) = \mathbf{P}(X \leq Y) f_{Z|\{X \leq Y\}}(z) + \mathbf{P}(X \geq Y) f_{Z|\{X \leq Y\}}(z) = \frac{1}{2} e^{-\lambda |z|}.$$

(b) For the general case of different parameters, the argument is similar, except that

$$\mathbf{P}(X \leq Y) = \lambda_X / (\lambda_X + \lambda_Y).$$

Furthermore, conditioned on $X \leq Y$, $Z$ is exponential with parameter $\lambda_Y$; conditioned on $Y \leq X$, $Z$ is the negative of an exponential with parameter $\lambda_X$. It follows that

$$f_Z(z) = \begin{cases} \dfrac{\lambda_Y}{\lambda_X + \lambda_Y} \lambda_X e^{-\lambda_X z}, & z \geq 0, \\[2ex] \dfrac{\lambda_X}{\lambda_X + \lambda_Y} \lambda_Y e^{\lambda_Y z}, & z < 0. \end{cases}$$

**Problem 32.** **[D]** Al makes cigars, placing each cigar on a constant-velocity conveyor belt as soon as it is finished. Bo packs the cigars into boxes of four cigars each, placing each box back on the belt as soon as it is filled. The time Al takes to construct any particular cigar is, believe it or not, an independent exponential random variable with an expected value of five minutes.

(a) Let $K$ be the number of cigars that Al makes in $t$ minutes. Determine $P_A(k,t)$, the probability that Al makes exactly $k$ cigars in $t$ minutes. Determine the mean and variance of $K$ as a function of $t$.

(b) Determine the probability density function $f_T(t)$, where $T$ is the interarrival time (measured in minutes) between placing successive cigars on the conveyor belt.

(c) Determine $P_B(r,t)$, the probability that Bo places exactly $r$ boxes of cigars back on the belt during an interval of $t$ minutes.

(d) Determine the probability density function $f_S(s)$ where $S$ is the interarrival time (measured in minutes) between placing successive boxes of cigars on the conveyor belt.

(e) If we arrive at a certain point in time, long after the process began, determine the PDF $f_R(r)$, where $R$ is the duration of our wait until we see a box of cigars.

**Solution.** Let $V$ be the time to construct any particular cigar. Since the times to construct a cigar are independent and exponentially distributed, we have a Poisson process with rate $\lambda = 1/\mathbf{E}[V] = 0.2$.

(a) We have, for any $t \geq 0$,

$$P_A(k,t) = \frac{(0.2t)^k e^{-0.2t}}{k!}, \qquad k = 0, 1, \ldots.$$

Also,

$$\mathbf{E}[K] = \lambda t = 0.2t, \qquad \sigma_K^2 = \lambda t = 0.2t.$$

(b) The random variable $T$ is the interarrival time in a Poisson process and is thus exponentially distributed:

$$f_T(t) = 0.2e^{-0.2t}, \qquad \tau \geq 0.$$

(c) Let $A_i$ be the event Bo has $i$ cigars at the start of the $t$-minute period. We have, for any $t \geq 0$,

$$
\begin{aligned}
P_B(r,t) &= \mathbf{P}(\text{Bo packs } r \text{ boxes in } t \text{ minutes}) \\
&= \sum_{i=0}^{3} \mathbf{P}(A_i)\mathbf{P}(\text{Bo packs } r \text{ boxes in } t \text{ minutes} \mid A_i) \\
&= \sum_{i=0}^{3} P(A_i)\mathbf{P}(\text{Bo receives between } 4r-i \text{ and } 4r-i+3 \text{ cigars, inclusive}) \\
&= \sum_{i=0}^{3} \mathbf{P}(A_i) \sum_{k=4r-i}^{4r-i+3} \mathbf{P}_A(k,T) \\
&= \frac{1}{4}\Big(P_A(4r-3,T) + 2P_A(4r-2,T) + 3P_A(4r-1,T) + 4P_A(4r,T) \\
&\quad + 3P_A(4r+1,T) + 2P_A(4r+2,T) + P_A(4r+3,T)\Big), \quad r = 0,1,\ldots.
\end{aligned}
$$

(d) The interarrival time for boxes is the same as the time until the fourth arrival of a cigar, and has an Erlang PDF of order 4, with parameter 0.2:

$$
f_S(s) = \frac{(0.2)^4 s^3 e^{-0.2s}}{3!}, \qquad s \geq 0.
$$

(e) Using the events $A_i$ defined in part (c), we have

$$
f_R(r) = \sum_{i=0}^{3} \mathbf{P}(A_i) f_{R|A_i}(r \mid A_i) = \sum_{i=0}^{3} \mathbf{P}(A_i) f_{Y_{4-i}}(r_0),
$$

where $Y_{4-i}$ has an Erlang PDF of order $4-i$ and parameter $\lambda = 0.2$. Thus,

$$
f_R(r) = \frac{1}{4}\left(\frac{(0.2)^4 r^3 e^{-0.2r}}{3!} + \frac{(0.2)^3 r^2 e^{-0.2r}}{2!} + \frac{(0.2)^2 r e^{-0.2r}}{1!} + (0.2)e^{-0.2r}\right), \qquad r \geq 0.
$$

**Problem 33.** **[D]** All ships travel at the same speed through a wide canal. Eastbound ships arrive as a Poisson process with an arrival rate of $\lambda_E$ ships per day. Westbound ships arrive as an independent Poisson process with an arrival rate of $\lambda_W$ ships per day. An indicator at a point in the canal is always pointing in the direction of travel of the most recent ship to pass. Each ship takes $t$ days to traverse the canal.

(a) What is the probability that the next ship passing by the indicator causes it to change its direction?

(b) What is the probability that an eastbound ship will see no westbound ships during its eastward journey through the canal?

(c) If we begin observing at an arbitrary time, determine the probability mass function of the total number of ships we observe up to and including the seventh eastbound ship we see.

(d) If we begin observing at an arbitrary time, determine the PDF of the time until we see the seventh eastbound ship.

(e) Given that the pointer is pointing west:

  (i) What is the probability that the next ship to pass it will be westbound?

  (ii) What is the PDF of the remaining time until the pointer changes direction?

**Solution.** (a) This is the probability that a westbound ship passed last (making the indicator point west) times the probability an eastbound ship will pass next, plus the probability an eastbound ship passed last (making the indicator point east) times the probability a westbound ship is next, and is equal to

$$\left(\frac{\lambda_E}{\lambda_E+\lambda_W}\right)\left(\frac{\lambda_W}{\lambda_E+\lambda_W}\right)+\left(\frac{\lambda_W}{\lambda_E+\lambda_W}\right)\left(\frac{\lambda_E}{\lambda_E+\lambda_W}\right)=2\frac{\lambda_E\lambda_W}{(\lambda_E+\lambda_W)^2}.$$

(b) Suppose that an eastbound ship enters the canal at time $t_0$. This ship will meet any westbound ship that entered the canal between times $t_0-t$ and $t_0+t$. Thus, the desired probability is the probability that there are no westbound ship arrivals during an interval of length $2t$, and using the Poisson PMF, it is equal to

$$e^{-\lambda_W 2t}.$$

(c) We view each ship arrival as an independent trial, and each eastbound ship as a success. Each trial is a success with probability

$$p=\lambda_E/(\lambda_E+\lambda_W).$$

We are interested in the PMF of the number of trials until the seventh success. This is a Pascal PMF of order seven, with parameter $p$, of the form

$$\binom{k-1}{6}p^7(1-p)^{k-7}, \qquad k=7,8,\ldots$$

(d) The time until we see the seventh eastbound ship is an Erlang random variable of order 7, with parameter $\lambda_E$, of the form

$$\frac{\lambda_E^7 t^6 e^{-\lambda_E t}}{6!}, \qquad t\geq 0.$$

(e) (i) The direction of the next ship does not depend on the previous ships. Therefore, this is just the probability $\lambda_W/(\lambda_E + \lambda_W)$ that the next ship is westbound.
(ii) The pointer will change directions on the next arrival of an eastbound ship. The time until an eastbound ship arrives is an exponential random variable with parameter $\lambda_E$, and its PDF is

$$\lambda_E e^{-\lambda_E t}, \qquad t \geq 0.$$

**Problem 34.** We are given the following statistics about the number of children in a typical family in a small village. There are 100 families. 10 families have no children, 40 have 1, 30 have 2, 10 have 3, 10 have 4.

(a) If you pick a family at random, what is the expected number of children in that family?

(b) If you pick a child at random (each child is equally likely), what is the expected number of children in that child's family (including the picked child)?

(c) Generalize your approach from part (b) to the case where a fraction $p_k$ of the families has $k$ children, and provide a formula.

**Solution.** (a) Let $K$ be the number of children in a randomly selected family. We have

$$\mathbf{E}[K] = \frac{10}{100} \cdot 0 + \frac{40}{100} \cdot 1 + \frac{30}{100} \cdot 2 + \frac{10}{100} \cdot 3 + \frac{10}{100} \cdot 4 = \frac{17}{10}.$$

(b) A child picked at random is more likely to be in a large family than a small one. So, the probability law for the number of children in the family that includes the randomly selected child must place more weight on larger families. This objective is accomplished by weighting the probability law for the number of children in a randomly selected family by the number of children in the family. Let $W$ be the number of children in the family of the randomly selected child. The PMF of $W$ must be proportional to $wp_K(w)$. For the PMF to sum to 1, we need to normalize by dividing by $\sum_w wp_K(w)$, which is the same as $\mathbf{E}[W]$. Therefore,

$$p_W(w) = w\frac{p_K(w)}{\mathbf{E}[K]}$$

and

$$p_W(w) = \begin{cases} \dfrac{1 \cdot (4/10)}{17/10} = \dfrac{4}{17}, & w = 1, \\[2ex] \dfrac{2 \cdot (3/10)}{17/10} = \dfrac{6}{17}, & w = 2, \\[2ex] \dfrac{3 \cdot 1/10}{17/10} = \dfrac{3}{17}, & w = 3, \\[2ex] \dfrac{4 \cdot 1/10}{17/10} = \dfrac{4}{17}, & w = 4, \\[2ex] 0, & \text{otherwise.} \end{cases}$$

Using this probability law, we can calculate the expected number of children in the family of the randomly selected child as

$$\mathbf{E}[W] = \frac{4}{17} \cdot (1) + \frac{6}{17} \cdot (2) + \frac{3}{17} \cdot (3) + \frac{4}{17} \cdot (4) = \frac{41}{17}.$$

(c) Generalizing from part (b), we have

$$\begin{aligned}
\mathbf{E}[W] &= \sum_w w p_W(w) \\
&= \sum_w w \cdot \frac{w p_K(w)}{\mathbf{E}[K]} \\
&= \sum_k k \cdot \frac{k p_K(k)}{\mathbf{E}[K]} \\
&= \frac{\mathbf{E}[K^2]}{\mathbf{E}[K]},
\end{aligned}$$

where the third equality is obtained by observing that $w$ is just a “dummy variable” that runs through the same values as $k$.

**Problem 35.** Consider a Poisson process of rate $\lambda$. Let $N$ be the number of arrivals in $(0, t]$, and let $M$ be the number of arrivals in $[0, t+s]$.

(a) Find the joint PMF of $N$ and $M$.

(b) Find $\mathbf{E}[NM]$.

**Solution.** (a) The event $\{M = m,\ N = n\}$ occurs when $\{N = n\}$ and $\{M - N = m - n\}$. That is, from $[0, t]$ there have to be $n$ arrivals, and after $t$ but prior to $t+s$ there have to be $m - n$ arrivals. Since the interval $[0, t]$ is disjoint from the interval $(t, t+s]$, the numbers of arrivals in each are independent and have a Poisson distribution with rate $\lambda$. Symbolically,

$$p_{N,M}(n, m) = p_N(n) p_{M \mid N}(m \mid n) = \left[\frac{(\lambda t)^n e^{-\lambda t}}{n!}\right] \left[\frac{(\lambda s)^{m-n} e^{-\lambda s}}{(m-n)!}\right]$$

(b) We rewrite $\mathbf{E}[NM]$ as

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

where the second equality is obtained using the independence of $N$ and $N - M$.