Solutions to Review Problems for Exam #2

1. **Poisson Processes and Random Mutations**. It was shown in class and in the lecture notes that, if $M(t)$ denotes the number of mutations that occur in a bacterial colony in the time interval $[0, t]$, then $M(t)$ can be modeled by a Poisson process; in other words, for each $t > 0$, $M(t)$ is a modeled by a Poisson random variable with parameter λt, where the parameter λ denotes the (constant) average number of mutations per unit time. Hence,

$$\Pr[M(t) = m] = \begin{cases} \dfrac{(\lambda t)^m}{m!}\, e^{-\lambda t}, & \text{for } m = 0, 1, 2, 3, \ldots \text{ and } t \geqslant 0; \\ 0 & \text{elsewhere.} \end{cases} \tag{1}$$

(a) Let T_1 denote the time of occurrence of the first mutation. Give the probability density function for T_1 and compute its expected value.

Solution: First, note that

$$\Pr(T_1 > t) = \Pr[M(t) = 0]; \tag{2}$$

since $t < T_1$ if and only if there are no mutations in $[0, t]$.
It follows from (2) and (1) with $m = 0$ that

$$\Pr(T_1 > t) = e^{-\lambda t},$$

so that

$$\Pr(T_1 \leqslant t) = 1 - e^{-\lambda t}, \quad \text{for } t > 0. \tag{3}$$

We conclude from (3) that T_1 has cdf given by

$$F_{T_1}(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{for } t > 0; \\ 0 & \text{for } t \leqslant 0. \end{cases} \tag{4}$$

Differentiating $F_{T_1}(t)$ with respect to t, for $t \neq 0$ in (4), yields the pdf

$$f_{T_1}(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{for } t > 0; \\ 0 & \text{for } t \leqslant 0, \end{cases}$$

so that T_1 has an exponential distribution with parameter $1/\lambda$. Hence,

$$E(T_1) = \frac{1}{\lambda}.$$

□

(b) Compute the limits $\lim_{t\to 0}\dfrac{\Pr[M(t)=1]}{t}$ and $\lim_{t\to 0}\dfrac{\Pr[M(t)\geqslant 2]}{t}$ and give interpretations to your results.

***Solution**:* Use 1 with $m=1$ to compute

$$\frac{\Pr[M(t)=1]}{t}=\frac{\lambda t\ e^{-\lambda t}}{t}=\lambda e^{-\lambda t},\qquad \text{for } t\neq 0,$$

so that

$$\lim_{t\to 0}\frac{\Pr[M(t)=1]}{t}=\lambda. \tag{5}$$

An interpretation of (5) is that, when $t>0$ is very small, the probability that there will be exactly one mutation in $[0,1]$ is approximately proportional to t, with constant of proportionality λ.

Next, observe that

$$\Pr[M(t)\geqslant 2]=1-\Pr[M(t)\leqslant 1]=1-\Pr[M(t)=0]-\Pr[M(t)=1],$$

so that

$$\Pr[M(t)\geqslant 2]=1-e^{-\lambda t}-\lambda t\ e^{-\lambda t}. \tag{6}$$

Dividing on both sides of (6) by $t\neq 0$ we then obtain

$$\frac{\Pr[M(t)\geqslant 2]}{t}=\frac{1-e^{-\lambda t}}{t}-\lambda\ e^{-\lambda t}. \tag{7}$$

Note that, by L'Hospital's Rule,

$$\lim_{t\to 0}\frac{1-e^{-\lambda t}}{t}=\lim_{t\to 0}\lambda e^{-\lambda t}=\lambda; \tag{8}$$

so that, combining (7) and (8),

$$\lim_{t\to 0}\frac{\Pr[M(t)\geqslant 2]}{t}=0. \tag{9}$$

Thus, the probability that there will be two or more mutations in $[0,t]$, when $|t|$ is very small, is close to 0. □

(c) For each real pair of real numbers, t_1 and t_2, with $t_1<t_2$, define $Y=M(t_2)-M(t_1)$. Compute the expected value, $E(Y)$, of Y, and give and interpretation for your result.

Solution: Compute

$$\begin{aligned} E(Y) &= E[M(t_2) - M(t_1)] \\ &= E[M(t_2)] - E[M(t_1)] \\ &= \lambda t_2 - \lambda t_1 \\ &= \lambda(t_2 - t_1); \end{aligned}$$

so that the expected number of mutations in the interval $(t_1, t_2]$ is proportional to the length of the interval, $t_2 - t_1$, with constant of proportionality λ. □

2. **Random Walk on the Integers**. A particle starts at $x = 0$ and, after one unit of time, it moves one unit to the right with probability p, for $0 < p < 1$, or to the left with probability $1 - p$. Assume that at each time step, whether a particle will move to the right or to the left is independent of where it has been.

   (a) Let X_1 denote the position of the particle after one unit of time and X_2 denote that after 2 units of time. Give the probability distributions for X_1 and X_2 and compute their expectations and variances.

   Solution: Let S denote the random variable with values -1 and 1, and probability distribution function given by

$$p_S(x) = \begin{cases} 1 - p & \text{if } x = -1; \\ p & \text{if } x = \ 1. \end{cases} \tag{10}$$

   Then, the expected value of S is

$$E(S) = (-1)(1 - p) + (1)p = 2p - 1, \tag{11}$$

   and

$$\text{Var}(S) = E(S^2) - [E(S)]^2, \tag{12}$$

   where

$$E(S^2) = (-1)^2(1 - p) + (1)^2 p = 1. \tag{13}$$

   Next, use (13) and (11) to obtain from (12) that

$$\text{Var}(S) = 1 - [2p - 1]^2 = 4p(1 - p). \tag{14}$$

Set $X_0 = 0$, so that $X_1 = X_0 + S$. Thus, X_1 has the same probability distribution as that of S; thus, in view of (10), (11) and (13),

$$p_{X_1}(x) = \begin{cases} 1-p & \text{if } x = -1; \\ p & \text{if } x = \ 1, \end{cases} \tag{15}$$

$$E(X_1) = 2p - 1, \tag{16}$$

and

$$\text{Var}(X_1) = 4p(1-p). \tag{17}$$

Next, observe that

$$X_2 = X_1 + S, \tag{18}$$

and that possible values for X_2 are -2, 0, and 2. For those values of X_2 we compute

$$\begin{aligned} p_{X_2}(k) &= \Pr(X_1 + S = k) \\ &= \sum_{\ell} \Pr(S = \ell, X_1 = k - \ell) \\ &= \sum_{\ell} \Pr(S = \ell) \cdot \Pr(X_1 = k - \ell), \end{aligned}$$

since X_1 and S are independent. We then have that

$$p_{X_2}(k) \ = \ p_S(-1) \cdot p_{X_1}(k+1) + p_S(1) \cdot p_{X_1}(k-1). \tag{19}$$

Using (10) and (15), we obtain from (19) that

$$p_{X_2}(x) = \begin{cases} (1-p)^2 & \text{if } x = -2; \\ 2(1-p)p & \text{if } x = \ 0; \\ p^2 & \text{if } x = \ 2. \end{cases} \tag{20}$$

Finally, use (18) to get

$$E(X_2) = E(X_1) + E(S) = 2(2p - 1), \tag{21}$$

where we have used (11) and (16); and

$$\text{Var}(X_2) = \text{Var}(X_1) + \text{Var}(S) = 8p(1-p), \tag{22}$$

since X_1 and S are independent, where we have used (14) and (17). □

(b) Let X_3 denote the position of the particle in the previous part after 3 units of time. Give probability distribution, expectation and variance of X_3. Generalize this result to X_n, the position of the particle after n units of time. The set of random variables $\{X_n \mid n = 0, 1, 2, 3, \ldots\}$ is an example of a discrete–time random process

Solution: Let S and X_2 be as defined in part (a), and note that

$$X_3 = X_2 + S. \tag{23}$$

Then, the calculations leading to (19) in part (a) imply that

$$p_{X_3}(k) \quad = \quad p_S(-1) \cdot p_{X_2}(k+1) + p_S(1) \cdot p_{X_2}(k-1), \tag{24}$$

where $k = -3, -1, 1, 3$. Thus, using (10) and (20), we obtain from (24) that

$$p_{X_3}(x) = \begin{cases} (1-p)^3 & \text{if } x = -3; \\ 3(1-p)^2 p & \text{if } x = -1; \\ 3(1-p)p^2 & \text{if } x = \ \ 1; \\ p^3 & \text{if } x = \ \ 3. \end{cases} \tag{25}$$

Next, use (23), (11), (14), (16), (17) and the independence of X_2 and S to get

$$E(X_3) = 3(2p-1), \tag{26}$$

and

$$\text{Var}(X_3) = 12p(1-p). \tag{27}$$

Observe that the probabilities given in (20) and (25) are the ones given by the Binomial(n, p) distribution for $n = 2$ and $n = 3$, respectively. An inductive argument on n, for

$$X_n = X_{n-1} + S, \tag{28}$$

will then yield the following probability distribution for X_n

$$p_{X_n}(x) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \text{for } x = 2k - n,\ k = 0, 1, \ldots, n.$$

Similarly, using (28), the independence of X_n and S, and induction on n, we obtain

$$E(X_n) = n(2p-1),$$

and

$$\text{Var}(X_n) = 4np(1-p).$$

□

3. **Exponential Distributions**. A continuous random variable, T, is said to have and exponential distribution with parameter $\beta > 0$, if its probability density function, f_T, is given by

$$f_T(t)\begin{cases}\dfrac{1}{\beta}\,e^{-t/\beta} & \text{for } t \geqslant 0;\\ 0 & \text{elsewhere.}\end{cases} \tag{29}$$

   (a) Compute the conditional probability

$$\Pr(T > t + s \mid T > t)$$

   for all $t, s > 0$.
   Give and interpretation to your result.
   ***Solution**:* Compute

$$\Pr(T > t + s \mid T > t) \;=\; 1 - \Pr(T \leqslant t + s \mid T > t), \tag{30}$$

   where

$$\Pr(T \leqslant t + s \mid T > t) \;=\; \frac{\Pr(T \leqslant t + s, T > t)}{\Pr(T > t)}. \tag{31}$$

   Next, use the probability density function in (29) to compute

$$\begin{aligned}\Pr(T \leqslant t + s, T > t) &= \Pr(t < T \leqslant t + s)\\ &= \int_t^{t+s} \frac{1}{\beta}\,e^{-t/\beta}\,dt\\ &= e^{-t/\beta} - e^{-(t+s)/\beta},\end{aligned}$$

   so that

$$\Pr(T \leqslant t + s, T > t) \;=\; e^{-t/\beta}\left[1 - e^{-s/\beta}\right]. \tag{32}$$

   Similarly,

$$\begin{aligned}\Pr(T > t) &= \int_t^{\infty} \frac{1}{\beta}\,e^{-t/\beta}\,dt\\ &= \lim_{b\to\infty} \int_t^{b} \frac{1}{\beta}\,e^{-t/\beta}\,dt\\ &= \lim_{b\to\infty} \left[e^{-t/\beta} - e^{-b/\beta}\right],\end{aligned}$$

so that

$$\Pr(T > t) \;=\; e^{-t/\beta}. \tag{33}$$

It then follows from (32), (33) and (31) that

$$\Pr(T \leqslant t + s \mid T > t) \;=\; \frac{e^{-t/\beta}[1 - e^{-s/\beta}]}{e^{-t/\beta}} = 1 - e^{-s/\beta}. \tag{34}$$

Thus, combining (30) and (34)

$$\Pr(T > t + s \mid T > t) \;=\; e^{-s/\beta}, \quad \text{for } s > 0. \tag{35}$$

It follows from (35) that the conditional probability $\Pr(T > t + s \mid T > t)$ is independent of t. □

(b) **Survival Time After a Treatment**. In Problem 5 of Assignment #9 you showed that the survival time, T, after a treatment can be modeled by an exponential random variable with parameter β, where β is the expected time of survival.

The survival function, $S(t)$, is the probability that a randomly selected person will survive for at least t years after receiving treatment. Compute $S(t)$.

Suppose that a patient has a 70% probability of surviving at least two years. Estimate the expected survival time of the treatment.

***Solution**:* Note that

$$S(t) = \Pr(T > t) = e^{-t/\beta}, \quad \text{for } t > 0, \tag{36}$$

where we have used (33).

If we are given that $S(2) = 0.7$, it follows from (36) that

$$e^{-2/\beta} = 0.7 \tag{37}$$

Solving (37) for β yields

$$\beta = -\frac{2}{\ln(0.7)} \doteq 5.6.$$

Thus, the expected survival time after treatment is about 5.6 years. □