ECS 132 Spring 2024: Assignment 5
Solutions

This problem set covers:

- Problem 1: Discrete random variables and conditional probability
- Problem 2: A general pdf and a discrete counting problem
- Problem 3: Minimum of two exponential processes
- Problem 4: Normal distribution and Z-scores
- Problem 5: Sensitivity and specificity

1 Problem

Recall the Medium Access Control (MAC) Protocol discussed in lecture (at the end of Sec V.6.) We saw that each time slot could be in one of three states: I (idle), C (collision) or S (Success). Consider here that there are 6 nodes who want to send packets and that all of the nodes act independently of one another (the example in lecture had 5 nodes). Each node has a probability $p = 0.4$ of being active in a time slot. Let Y be the random variable that denotes the number of nodes that are active in a time slot. For an idle slot $Y = 0$, for a successful slot $Y = 1$, and if $Y \geq 2$ the slot has a collision.

1. For a single time slot, calculate these two things: (a) $P(I)$ and (b) $P(S)$.
2. Now consider the event E, that there is an eavesdropper who sometimes monitors the time slot to read the the content being sent by the nodes, and that $P(E) = 0.5$. By looking through data logs from your security software you learn about the behavior of the eavesdropper. Specifically, that with probability 0.2 the the slot will be an idle slot given that the eavesdropper is listening, and that with probability 0.3 the slot will be a successful slot given that the eavesdropper is listening. What is the probability that the eavesdropper is listening given that the time slot is a collision? (Note, you calculated $P(I)$ and $P(S)$ above and are given information about conditional probabilities here.)

Answer

Since the sensors are independent the number of sensors that are active in any one time slot $Y \sim \text{binom}(6, p)$.

1. a) $P(I) = P(Y = 0) = \binom{6}{0} p^0(1-p)^6 = (1-p)^6 = 0.046656$
2. b) $P(S) = P(Y = 1) = \binom{6}{1} p(1-p)^5 = 6p(1-p)^5 = 0.186624$

3. We need to solve for $P(E|C)$.

   Givens:

   - $P(I)$ calculated above
   - $P(S)$ calculated above
   - $P(I|E) = 0.2$
   - $P(S|E) = 0.3$

   From these givens we can deduce:

   - $P(C) = 1 - P(I) - P(S) = 0.76672$
   - $P(C|E) = 1 - P(I|E) - P(S|E) = 0.5$

   Now to solve the problem:

$$\begin{aligned} P(E|C) &= \frac{P(E)P(C|E)}{P(C)} \\ &= \frac{(0.5)(0.5)}{0.76672} \\ &= 0.326 \end{aligned}$$

2 Problem

The probability density function of X, the lifetime of a certain type of electronic device (measured in hours) is given by

$$f(x) = \begin{cases} \frac{10}{x^2} & x > 10 \\ \\ 0 & x \leq 10 \end{cases}$$

1. Calculate the mathematical expression for the cumulative distribution function of X.
2. What is the probability that the device will fail within the first 15 hours?
3. Now consider that you have a large collection of these devices lined up in a row, numbered device1, device2, etc. They are all independent of one another. What is the probability that the 4th device will be the first one that does not fail within the first 15 hours?

Answers

1. Calculate the cumulative distribution function of X.

$$\begin{aligned} P\{X \leq a\} &= \int_{-\infty}^{a} f(x)dx = \int_{10}^{a} f(x)dx \\ &= \int_{10}^{a} \frac{10}{x^2}\,dx \\ &= -10x^{-1}\Big|_{10}^{a} \\ &= 1 - \frac{10}{a} \quad \text{ for } \; a \geq 10 \; \text{ and } \; 0 \quad \text{otherwise} \end{aligned}$$

2. Solve for $F_X(15)$.

$$\begin{aligned}
P\{X \leq 15\} &= 1 - \int_{10}^{15} \frac{10}{x^2} dx \\
&= -10x^{-1}\Big|_{10}^{15} \\
&= -(2/3) + 1 \\
&= 1/3
\end{aligned}$$

3. What is the probability that the 4th device will be the first one that does not fail within the first 15 hours?

   Each device is independent and follows a Bernoulli process where success is that the device does not fail within the first 15 hours, $p = P(X > 15)$. The random variable Y corresponding to the first success in a series of $X \sim Bern(p)$ processes is the geometric distribution, $Y \sim Geom(p)$.

   Using the part above we solve for $p = 1 - P(X \leq 15) = 2/3$. The probability that $Y = 4$ is

$$\begin{aligned}
P(Y = 4) &= (1-p)^3 p \\
&= (1/3)^3(2/3) = 0.0247
\end{aligned}$$

3 Problem

You arrive at a bus stop to find there are no busses currently there. Two different lines, line A and line B service the bus stop and both will take you to your destination so you get onto the first bus that arrives. Both busses have arrival times corresponding to exponential distributions and the busses are independent of one another. Bus A on average comes by once every ten minutes. Bus B on average comes by once every 5 mins. What is the probability that you will have to wait more than 8 mins for a bus to arrive?

Answer

The arrival time of Bus A is random variable $X_A \sim \text{Expo}(\lambda_A)$ and the arrival time of Bus B is random variable $X_B \sim \text{Expo}(\lambda_B)$. Here the rates $\lambda_A = 1/10$ and $\lambda_B = 1/5$. Just like the example in lecture of the two cores handling arriving jobs, we want to define the random variable $W = \min(X_A, X_B)$ and calculate the probability that $W > 8$.

$$\begin{aligned}
P(W > 8) &= P(X_A > 8 \cap X_B > 8) \\
&= P(X_A > 8) \cdot P(X_B > 8) \\
&= \exp(-\lambda_A \cdot 8) \cdot \exp(-\lambda_B \cdot 8) \\
&= \exp(-\frac{1}{10}8) \cdot \exp(-\frac{1}{5}8) \\
&= (0.449)(0.202) = 0.0907
\end{aligned}$$

Recall that the CDF for the exponential distribution $F_X(y) = 1 - \exp(-\lambda y)$. So $P(X > y) = 1 - F_X(y) = \exp(-\lambda y)$

4 Problem

Suppose the size of a network flow is a normal random variable with parameters $\mu = 71$ GBytes and $\sigma = 2.5$ GBytes. Use a Z-table to determine the following. (This requires mapping your normal distribution onto the standard normal distribution.)

1. What percentage of the flows are greater than 72 GBytes?
2. Let m denote the size of the flow. What is the value of m for which 88.30% of the flows are smaller than m?

Answer

1. Let X be a random variable that denotes the network flow size. $X \sim \mathrm{Norm}(71, 2.5)$. We want to find $P(X > 72)$ which is $1 - P(X \leq 72)$.

$$\begin{aligned} P(X \leq 72) &= P(\frac{X-71}{2.5} \leq \frac{72-71}{2.5}) \\ &= \Phi(0.4) \\ &= 0.6554 \end{aligned}$$

   Thus the required probability is $1 - 0.6554 = 0.3446$.

2. We want to know for what value of m is it that case that $F_X(m) = 0.8830$. From the Z-table we read off that $\Phi(Z) = 0.8830$ for $Z = 1.19$. Now we translate that into the X value. Recall

$$Z = \frac{X - \mu}{\sigma} \quad \text{so} \quad X \; = \; Z\sigma + \mu$$

   And here we set $X = m$ so

$$m = \; Z\sigma + \mu \; = \; (1.19)(2.5) + 71 = 73.975$$

5 Problem

In a newspaper trivia column, L. M. Boyd (Boyd, L. M.: The Grab Bag (syndicated newspaper column), The San Francisco Chronicle (July 17, 1999)) ponders why lie detector results are not admissible in court. His answer is that "lie detector tests pass 10 percent of the liars and fail 20 percent of the truth-tellers." If you use these percentages and take $L = 1$ to mean being a liar and $F = 1$ to mean failing the test, what are the numerical values of the (1) sensitivity and (2) specificity for such a lie detector test?

Answer

We define the following two random variables

$$L = \begin{cases} 1 & \text{person is a liar} \\ 0 & \text{person is a truth teller} \end{cases}$$

$$F = \begin{cases} 1 & \text{fails lie detector test} \\ 0 & \text{passes lie detector test} \end{cases}$$

1. 10% of the liars pass test. This implies $P(F = 0)|(L = 1) = 0.1$, which further implies that $P(F = 1)|(L = 1) = 0.9$. Thus the sensitivity $\eta = 0.9$.

2. Similarly, 20% of truth tellers fail the test. This implies that $P(F = 1)|(L = 0) = 0.2$. This implies that the specificity $\theta = P(F = 0)|(L = 0) = 0.8$.

(Remember $P(A^{\mathsf{C}}|B) + P(A|B) = 1$.)