1. **Coupon Collector**

Each brand of candy bar has one coupon in it. There are \( n \) different coupons in total; getting at least one coupon of each type entitles you to a prize. Each candy bar you eat can have any one of the coupons in it, with all being equally likely. Let \( X \) be the (random) number of candy bars you eat before you have all coupons. What are the mean and variance of \( X \)?

**Solution**

First, let \( X_1 \) be the number of candy bars which must be eaten to obtain a unique coupon. We will have that a unique coupon is obtained when the first candy bar is eaten, so that \( X_1 = 1 \), always. Now let \( X_2 \) be the number of candy bars after the first which must be eaten before getting another unique coupon. We have \( X_2 \) is a geometric random variable with parameter \( \frac{n-1}{n} \), since there are \( n - 1 \) unique coupons left to be obtain out of the \( n \) possible coupons. We have, then, that \( E[X_2] = \frac{n}{n-1} \). Now similarly, let \( X_3 \) be the number of candy bars eaten after obtaining the second unique coupon before finding a third unique coupon. \( X_3 \) is geometric with parameter \( \frac{n-2}{n} \Rightarrow E[X_3] = \frac{n}{n-2} \).

In general, \( X_i \), or the number of candy bars eaten after obtaining the \((i - 1)\)-th unique coupon before obtaining the \(i\)-th unique coupon, is geometric with parameter \( \frac{n-i+1}{n} \), so that \( E[X_i] = \frac{n}{n-i+1} \).
We have that
\[ X = \sum_{i=1}^{n} X_i \Rightarrow E[X] = \sum_{i=1}^{n} E[X_i] \]
\[ = \sum_{i=1}^{n} \frac{n}{n - i + 1} \]
\[ = n \sum_{i=1}^{n} \frac{1}{i} \]

Since \( X_i \) and \( X_j \) are independent for \( i \neq j \), we have that
\[ Var[X] = Var \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} Var[X_i]. \]

Since
\[ Var[X_i] = \frac{1 - \frac{n-i+1}{n}}{\left( \frac{n-i+1}{n} \right)^2} = \frac{n(i-1)}{(n - i + 1)^2}, \]
then
\[ Var[X] = \sum_{i=1}^{n} \frac{n(i-1)}{(n - i + 1)^2}. \]

2. Minimum of Exponentials

(a) \( X_1 \) is an exponential random variable with parameter \( \lambda_1 \), and \( X_2 \) with \( \lambda_2 \). Let \( Y = \min(X_1, X_2) \). What is the PDF of \( Y \)? Is \( Y \) one of the common random variables?

(b) Use induction to show that the minimum of \( n \) exponential random variables with parameter 1 is an exponential random variable with parameter \( n \).

Solution

(a) We have that the CDF of \( Y \) follows
\[ F_Y(y) = P(Y \leq y) = P(\min(X_1, X_2) \leq y) = 1 - P(\min(X_1, X_2) > y) \]
\[ = 1 - P(X_1 > y \cap X_2 > y) \]
\[ = 1 - P(X_1 > y)P(X_2 > y) \]
We have that
\[ P(X_1 > y) = \int_y^\infty f_{X_1}(x)dx = \int_y^\infty \lambda_1 e^{-\lambda_1 x}dx = e^{-\lambda_1 y}. \]

Similarly, \( P(X_2 > y) = e^{-\lambda_2 y} \), so that
\[ F_Y(y) = 1 - P(X_1 > y)P(X_2 > y) = 1 - e^{-(\lambda_1 + \lambda_2)y} \]
\[ \Rightarrow f_Y(y) = \frac{d}{dy}F_Y(y) = (\lambda_1 + \lambda_2)e^{-(\lambda_1 + \lambda_2)y}. \]

Thus, \( Y \) is an exponential random variable with parameter \( \lambda = \lambda_1 + \lambda_2 \).

(b) We have from part (a) that if \( X_1 \) and \( X_2 \) are iid (independent and identically-distributed) exponential random variables with parameter \( \lambda = 1 \), that \( Z_2 \triangleq \min(X_1, X_2) \) is distributed as an exponential random variable with parameter \( \lambda_2 = 1 + 1 = 2 \). This proves the case for \( n = 2 \).

Now for the inductive case, assume that \( Z_{n-1} \triangleq \min(X_1, X_2, \ldots, X_{n-1}) \) is an exponential random variable with parameter \( \lambda_{n-1} = n - 1 \), and \( X_1, X_2, \ldots, X_{n-1} \) are exponential random variables with parameter \( \lambda = 1 \). Let \( X_n \) be an additional exponential random variable with parameter \( \lambda = 1 \). We wish to find the distribution of \( Z_n \triangleq \min(X_1, X_2, \ldots, X_{n-1}, X_n) \). We find, however, that
\[ Z_n = \min(X_1, X_2, \ldots, X_{n-1}, X_n) = \min(\min(X_1, X_2, \ldots, X_{n-1}), X_n) = \min(Z_{n-1}, X_n). \]

Since \( Z_{n-1} \) is exponential with parameter \( \lambda_{n-1} = n - 1 \), and \( X_n \) is exponential with parameter \( \lambda = 1 \), we have from part (a) that \( Z_n \) is exponential with parameter \( \lambda_{n-1} + \lambda = n \). This proves the inductive case, and completes the proof.

3. Random Chord

A circle has radius \( r \). Any chord of the circle is at distance at most \( r \) from the center. A random chord is drawn by first choosing its distance \( D \) from the center uniformly from the interval \([0, r]\), and then choosing any chord at that distance from the center. Find the PDF of \( L \), the length of the chord. Draw a figure to illustrate.
**Solution** Without loss of generality, $D$ is a random point in the interval $[0, r]$ on the horizontal axis, and the chord is parallel to the vertical axis as shown in the diagram.

Note that $L$ takes on values in $[0, 2r]$ and hence $F_L(l) = 0$ for $l < 0$, and $F_L(l) = 1$ for $l > 2r$. Furthermore, $D^2 + (L/2)^2 = r^2$, and hence for any $l \in [0, 2r],
\[
F_L(l) = P(L \leq l) = P(2\sqrt{r^2 - D^2} \leq l) = P(D^2 \geq r^2 - l^2/4) = P(D \geq \sqrt{r^2 - l^2/4}) = \frac{1}{r}(r - \sqrt{r^2 - l^2/4})
\]
since $D$ is uniformly distributed on $[0, r]$. It follows that
\[
f_L(l) = \frac{d}{dl}F_L(l) = \begin{cases} \frac{l}{2r\sqrt{2r^2-l^2}}, & 0 \leq l \leq 2r \\ 0, & \text{else} \end{cases}
\]

4. Fire Station

(a) A fire station is to be located at a point $a$ along a road of length $A$, $0 < A < \infty$. If fires will occur at points uniformly chosen on $(0, A)$, where should the station be located so as to minimize the expected distance from the fire? That is, choose $a$ so as to minimize the quantity $E[|X-a|]$ when $X$ is uniformly distributed over $(0, A)$. 
(b) Now suppose that the road is of infinite length—stretching from point 0 outward to \( \infty \). If the distance of a fire from point 0 is exponentially distributed with rate \( \lambda \), where should the fire station now be located? That is, we want to minimize \( E[|X - a|] \) with respect to \( a \) when \( X \) is now an exponential random variable with parameter \( \lambda \).

Solution

(a) We have that \( f_X(x) = 1/A, \) for \( x \in [0, A] \), and 0 otherwise. Thus

\[
E[|X - a|] = \int_0^a (a-x)\frac{1}{A} dx + \int_a^A (x-a)\frac{1}{A} dx \\
= \frac{1}{A} \left( \frac{A^2}{2} - aA + a^2 \right) \\
= \frac{1}{A} \left( \left( a - \frac{A}{2} \right)^2 + \frac{A^2}{4} \right).
\]

Minimizing in terms of \( a \), we will have that \( a - \frac{A}{2} = 0 \Rightarrow a = \frac{A}{2} \).

(b) We now have that \( f_X(x) = \lambda e^{-\lambda x}, \) for \( x \geq 0 \), so that

\[
E[|X - a|] = \int_0^a (a-x)\lambda e^{-\lambda x} dx + \int_a^A (x-a)\lambda e^{-\lambda x} dx \\
= a + \frac{1}{\lambda} 2e^{-a\lambda} + \frac{1}{\lambda}.
\]

To minimize in terms of \( a \), we compute

\[
\frac{d}{da} E[|X - a|] = 1 - 2e^{-a\lambda} = 0 \\
\Rightarrow a = \frac{\ln 2}{\lambda}.
\]

To check that this is a minimizer, we find that \( \frac{d^2}{da^2} E[|X - a|] = 2\lambda e^{-a\lambda} > 0 \) at \( a = \frac{\ln 2}{\lambda} \). Thus, we have found the minimizer.