CS 70: Homework #11

Abhijay Bhatnagar

November 19, 2018

1   Safeway Monopoly Cards

It's that time of the year again - Safeway is offering its Monopoly Card promotion. Each time you visit Safeway, you are given one of n different Monopoly Cards with equal probability. You need to collect them all to redeem the grand prize.

Let X be the number of visits you have to make before you can redeem the grand prize. Show that $\mathrm{Var}(X) = n^2 \left(\sum_{i=1}^{n} i^{-2} \right) - \mathbb{E}(X)$. *[Hint: Does this remind you of a particular problem? What is the expectation for this problem?]*

Solution:

Proof. This is variable similar to geometric r.v. baseball card problem. For this type of problem:

$$\mathbb{E}[x] = n \sum_{i=1}^{n} i^{-1}$$

Using this, we can solve for variance in the given form:

$$\mathrm{Var}(X) = n^2 \left(\sum_{i=1}^{n} i^{-2} \right) - n \sum_{i=1}^{n} i^{-1} = n \left(\sum_{i=1}^{n} \frac{(n-i)}{i^2} \right)$$

We will use this value to solve the problem. Now, let $X = X_1 + ... + X_n$ where X_i is the number of trips needed to get the ith new card.

$$\mathrm{Var}(X_i) = \frac{1 - p_i}{p_i^2} = \frac{1 - (\frac{n-i+1}{n})}{\frac{n-i+1}{n}^2} = n \frac{i-1}{(n-i+1)^2}$$

Since each card is independent, the variances are linear and we can easily find $\mathrm{Var}(X)$.

$$\mathrm{Var}(X) = n \sum_{i=1}^{n} \frac{i-1}{(n-i+1)^2} = n \left[\frac{0}{n^2} + \frac{1}{(n-1)^2} + ... + \frac{n-1}{1^2} \right] = n \sum_{i=1}^{n} \frac{n-i}{i^2}$$

Which is equal, as desired. $\qquad \square$

2   Geometric Distribution

Two faulty machines, M_1 and M_2, are repeatedly run synchronously in parallel (i.e., both machines execute one run, then both execute a second run, and so on). On each run, M_1 fails with probability p_1 and M_2 fails with probability p_2, all failure events being independent. Let the random variables X_1, X_2 denote the number of runs until the first failure of M_1, M_2 respectively; thus X_1, X_2 have geometric distributions with parameters p_1, p_2 respectively. Let X denote the number of runs until the first failure of *either* machine.

a.) Show that X also has a geometric distribution, with parameter $p_1 + p_2 - p_1 p_2$.

Solution: You can interpret X as a geometric r.v. with the probability of (1 - probability that neither machine fails) $= (1 - (1 - p_1)(1 - p_2)) = p_1 + p_2 - p_1 p_2$.

b.) Now, two technicians are hired to check on the machines every run. They decide to take turns checking on the machines every hour. What is the probability that the first technician is the first one to find a faulty machine?

Solution: Let $p = p_1 + p_2 - p_1 p_2$. The probability that the first technician finds it is $\sum_{k=1}^{\infty} p(1-p)^{2k} = p(\frac{1}{1-(1-p)^2}) = \frac{p}{2p - p^2} = \frac{1}{2-p}$

3   Geometric and Poisson

Let X be geometric with parameter p, Y be Poisson with parameter λ, and $Z = \max(X,Y)$. Assume X and Y are independent. For each of the following parts, your final answers should not have summations.

a.) Compute $P(X > Y)$.

Solution:

$$P(X > Y) = \sum_{y=0}^{\infty} [P(X > Y | Y = y) \times P(Y = y)]$$

$$= \sum_{y=0}^{\infty} (1-p)^y \times \frac{e^{-\lambda}\lambda^y}{y!}$$

In order to eliminate the (1-p) term, we can use some algebra to convert it into a Poisson probability:

$$P(X > Y) = e^{-\lambda p} \sum_{y=0}^{\infty} e^{\lambda p}(1-p)^y \times \frac{e^{-\lambda}\lambda^y}{y!}$$

$$= e^{-\lambda p} \sum_{y=0}^{\infty} e^{-\lambda(1-p)} \frac{[(1-p)\lambda]^y}{y!}$$

$$= e^{-\lambda p}$$

b.) Compute $P(Z \geq X)$.

Solution: 1. Z is the max of the two variables, and the max function is always greater than or equal to X.

c.) Compute $P(Z \leq Y)$.

Solution: $P(Z \leq Y) = P(X \leq Y) = 1 - P(X > Y) = 1 - e^{-\lambda p}$

4 Darts

Alvin is playing darts. His aim follows an exponential distribution; that is, the probability density that the dart is x distance from the center is $f_X(x) = e^{-x}$. The board's radius is 4 units.

a.) What is the probability the dart will stay within the board?

Solution: $P[X \leq 4] = \int_0^4 e^{-x} dx = 1 - e^{-4}$

b.) Say you know Alvin made it on the board. What is the probability he is within 1 unit from the center?

Solution: This is equivalent to:

$$
\begin{aligned}
P(X \leq 1 | X \leq 4) &= \frac{P[X \leq 1 \cap X \leq 4]}{P[X \leq 4]} \\
&= \frac{P[X \leq 1]}{P[X \leq 4]} \\
&= \frac{1 - e^{-1}}{1 - e^{-4}}
\end{aligned}
$$

c.) If Alvin is within 1 unit from the center, he scores 4 points, if he is within 2 units, he scores 3, etc. In other words, Alvin scores $\lfloor 5 - x \rfloor$, where x is the distance from the center. What is Alvin's expected score after one throw?

Solution:

$$
\begin{aligned}
E[x] &= \sum_{i=1}^{4} \left[(5 - i) \times \int_{i-1}^{i} e^{-x} dx \right] \\
&= 4 - \sum_{k=1}^{4} e^{-k}
\end{aligned}
$$

5 Exponential Practice

a.) Let $X_1, X_2 \sim \text{Exponential}(\lambda)$ be independent, $\lambda > 0$. Calculate the density of $Y := X_1 + X_2$.
[*Hint*: One way to approach this problem would be to compute the CDF of Y and then differentiate the CDF.]

Solution: The CDF of $Y = P[Y \leq y]$. For exponential distributions, we can consider $y > 0$. If $X_1 + X_2 \leq y, X_1 \leq y$ and $X_2 \leq y - X_1$. From this it follows:

$$
\begin{aligned}
P[Y \leq y] &= P[X_1 \leq y, X_2 \leq y - X_1] \\
&= \int_0^y \int_0^{y-x_1} \lambda e^{-\lambda x_1} \lambda e^{-\lambda x_2} dx_2 dx_1 \\
&= \lambda^2 \int_0^y e^{-\lambda x_1} \times \frac{1}{\lambda} [-e^{-\lambda(y-x_1)} + 1] dx_1 \\
&= \lambda \int_0^y e^{-\lambda x_1} - e^{-\lambda y} dx_1 \\
&= \lambda \left(\frac{1}{\lambda}[-e^{-\lambda(y-x_1)} + 1] - y e^{-\lambda y} \right)
\end{aligned}
$$

Now to differentiate that mess...

$$
\begin{aligned}
f_Y(y) &= \frac{d}{dy} P[Y \leq y] = \lambda e^{-\lambda y} - \lambda e^{-\lambda y} + \lambda^2 y e^{-\lambda y} \\
&= \lambda^2 e^{-\lambda y}
\end{aligned}
$$

b.) Let $t > 0$. What is the density of X_1, conditioned on $X_1 + X_2 = t$? [*Hint*: Once again, it may be helpful to consider the CDF $\Pr(X_1 \leq x \mid X_1 + X_2 = t)$. To tackle the conditioning part, try conditioning instead on the event $\{X_1 + X_2 \in [t, t + \varepsilon]\}$, where $\varepsilon > 0$ is small.]

Solution: Conditioning on the hint:

$$
\begin{aligned}
P[X_1 \leq x | X_1 + X_2 \in [t, t + \varepsilon]] &= \frac{P(X_1 \leq x \cap X_1 + X_2 \in [t, t + \varepsilon]}{P[X_1 + X_2 \in [t, t + \varepsilon]]} \\
&= \frac{\int_0^x \int_{t-x_1}^{t-x_1+\varepsilon} \lambda^2 e^{-\lambda x_1} e^{-\lambda x_2} dx_2 dx_1}{f_Y(t)\varepsilon} \\
&= \frac{\int_0^x \int_{t-x_1}^{t-x_1+\varepsilon} \lambda^2 e^{-\lambda x_1} e^{-\lambda x_2} dx_2 dx_1}{\lambda^2 t e^{-\lambda t} \varepsilon} \\
&= \frac{x}{t}
\end{aligned}
$$

The density is the derivative $\implies = \frac{1}{t}$.

6   Uniform Means

Let $X_1, X_2, ..., X_n$ be n independent and identically distributed uniform random variables on the interval $[0, 1]$ (where n is a positive integer).

a.) Let $Y = \min\{X_1, X_2, ..., X_n\}$. Find $\mathbb{E}[Y]$. [*Hint*: Use the tail sum formula, which says the expected value of a nonnegative random variable is $\mathbb{E}[X] = \int_0^\infty \Pr(X > x)\, dx$. Note that we can use the tail sum formula since $Y \geq 0$.]

Solution: $\frac{1}{n+1}$

$$E[\min\{X_1, X_2, \ldots, X_n\}] = n \int_0^1 x(1-x)^{n-1} dx = \frac{1}{n+1}$$

b.) Let $Z = \max\{X_1, X_2, \ldots, X_n\}$. Find $\mathbb{E}Z$. [*Hint*: Find the CDF.]

Solution: $1 - \frac{1}{n+1}$

Parts (a) and (b) are symmetric as min and max functions of the same variables, so the solution to (b) is the inverse of the solution to (a).