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 $\text{var}(X) = n^2 \left(\sum_{i=1}^{n} i^{-2}\right) - \mathbb{E}(X)$. *[Hint: Does this remind you of a particular problem? What is the expectation for this problem?]*

Solution:

Let X_i be the number of visits we need to make before we have collected the ith unique Monopoly card actually obtained, given that we have already collected $i-1$ unique Monopoly cards. Then $X = \sum_{i=1}^{n} X_i$ and each X_i is geometrically distributed with $p = (n-i+1)/n$. Then

$$\begin{aligned}
\text{var}(X) &= \sum_{i=1}^{n} \text{var}(X_i) && \text{(as the } X_i \text{ are independent)} \\
&= \sum_{i=1}^{n} \frac{1-(n-i+1)/n}{[(n-i+1)/n]^2} && \text{(variance of a geometric r.v. is } (1-p)/p^2) \\
&= \sum_{j=1}^{n} \frac{1-j/n}{(j/n)^2} && \text{(by noticing that } n-i+1 \text{ takes on all values from 1 to } n) \\
&= \sum_{j=1}^{n} \frac{n(n-j)}{j^2} \\
&= \sum_{j=1}^{n} \frac{n^2}{j^2} - \sum_{j=1}^{n} \frac{n}{j} \\
&= n^2 \left(\sum_{j=1}^{n} \frac{1}{j^2} \right) - \mathbb{E}(X) && \text{(using the coupon collector problem expected value).}
\end{aligned}$$

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 $p_1 + p_2 - p_1 p_2$.

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

Solution:

(a) We have that $X_1 \sim \text{Geom}(p_1)$ and $X_2 \sim \text{Geom}(p_2)$. Also, X_1, X_2 are independent r.v.'s. We also use the following definition of the minimum:

$$\min(x,y) = \begin{cases} x & \text{if } x \le y; \\ y & \text{if } x > y. \end{cases}$$

Now, for all $k \in \{1,2,\dots\}$, $\min(X_1,X_2) = k$ is equivalent to $(X_1 = k) \cap (X_2 \ge k)$ or $(X_2 = k) \cap (X_1 > k)$. Hence,

$$\begin{aligned} \mathbb{P}[X = k] &= \mathbb{P}[\min(X_1,X_2) = k] \\ &= \mathbb{P}[(X_1 = k) \cap (X_2 \ge k)] + \mathbb{P}[(X_2 = k) \cap (X_1 > k)] \\ &= \mathbb{P}[X_1 = k] \cdot \mathbb{P}[X_2 \ge k] + \mathbb{P}[X_2 = k] \cdot \mathbb{P}[X_1 > k] \end{aligned}$$

(since X_1 and X_2 are independent)

$$= [(1-p_1)^{k-1} p_1](1-p_2)^{k-1} + [(1-p_2)^{k-1} p_2](1-p_1)^k$$

(since X_1 and X_2 are geometric)

$$\begin{aligned} &= ((1-p_1)(1-p_2))^{k-1} (p_1 + p_2(1-p_1)) \\ &= (1 - p_1 - p_2 + p_1 p_2)^{k-1} (p_1 + p_2 - p_1 p_2). \end{aligned}$$

But this final expression is precisely the probability that a geometric r.v. with parameter $p_1 + p_2 - p_1 p_2$ takes the value k. Hence $X \sim \text{Geom}(p_1 + p_2 - p_1 p_2)$, and $\mathbb{E}[X] = (p_1 + p_2 - p_1 p_2)^{-1}$.

An alternative, slightly cleaner approach is to work with the *tail probabilities* of the geometric distribution, rather than with the usual point probabilities as above. In other words, we can work with $\mathbb{P}[X \ge k]$ rather than with $\mathbb{P}[X = k]$; clearly the values $\mathbb{P}[X \ge k]$ specify the values $\mathbb{P}[X = k]$ since $\mathbb{P}[X = k] = \mathbb{P}[X \ge k] - \mathbb{P}[X \ge (k+1)]$, so it suffices to calculate them instead.

We then get the following argument:

$$\begin{aligned}
\mathbb{P}[X \geq k] &= \mathbb{P}[\min(X_1, X_2) \geq k] \\
&= \mathbb{P}[(X_1 \geq k) \cap (X_2 \geq k)] \\
&= \mathbb{P}[X_1 \geq k] \cdot \mathbb{P}[X_2 \geq k] && \text{since } X_1, X_2 \text{ are independent} \\
&= (1-p_1)^{k-1}(1-p_2)^{k-1} && \text{since } X_1, X_2 \text{ are geometric} \\
&= ((1-p_1)(1-p_2))^{k-1} \\
&= (1-p_1-p_2+p_1p_2)^{k-1}.
\end{aligned}$$

This is the tail probability of a geometric distribution with parameter $p_1 + p_2 - p_1p_2$, so we are done.

(b) Let the required probability be denoted by p (the probability that the first technician is the first to find a faulty machine) and X denote the first point of failure of either of the machines. Now, we have:

$$p = \mathbb{P}[X = 1] + \mathbb{P}[X = 3] + \mathbb{P}[X = 5] + \ldots$$

as the first technician will be the first to find a faulty machine if and only if the first failure occurs in an odd run. Now, let us decompose the sum as follows:

$$p = \mathbb{P}[X = 1] + \sum_{i=1}^{\infty} \mathbb{P}[\{X = 2i+1\} \cap \{X \neq 1\}] = \mathbb{P}[X = 1] + \sum_{i=1}^{\infty} \mathbb{P}[X \neq 1]\mathbb{P}[X = 2i+1 | X \neq 1]$$

Using the memoryless property of the geometric distribution to further simplify the sum, we get:

$$\begin{aligned}
p &= \mathbb{P}[X = 1] + \sum_{i=1}^{\infty} \mathbb{P}[X \neq 1]\mathbb{P}[X = 2i] = \mathbb{P}[X = 1] + \mathbb{P}[X \neq 1] \sum_{i=1}^{\infty} \mathbb{P}[X = 2i] \\
&= \mathbb{P}[X = 1] + \mathbb{P}[X \neq 1] \left(1 - \sum_{i=1}^{\infty} \mathbb{P}[X = 2i-1]\right) = \mathbb{P}[X = 1] + \mathbb{P}[X \neq 1](1-p)
\end{aligned}$$

where the second equality follows because the probability that X is even is the complement of the event that X is odd. The final equation is intuitive as in the event that the first technician doesn't find either of the machines broken in the first run, the memoryless property of the geometric distribution ensures that the probability that the second technician finds the machines broken first is the same as the probability that the first technician does when we have no knowledge of the first run. That is, we have:

$$p = \mathbb{P}[\text{Second technician finds the machines broken first} \mid \text{No machine fails in run 1}]$$

.

By solving the above equation, we get that:

$$p = \frac{1}{2 - p_1 - p_2 + p_1 p_2}$$

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 $P(X > Y)$.

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

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

Solution:

(a) Condition on Y so you can use the nice property of geometric random variables that $P(X > k) = (1-p)^k$:

$$\begin{aligned}
P(Y < X) &= \sum_{y=0}^{\infty} P(X > Y | Y = y) P(Y = y) \\
&= \sum_{y=0}^{\infty} (1-p)^y \frac{e^{-\lambda}\lambda^y}{y!} \\
&= e^{-\lambda p} e^{\lambda p} \sum_{y=0}^{\infty} \frac{e^{-\lambda}(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p} \sum_{y=0}^{\infty} \frac{e^{-\lambda(1-p)}(\lambda(1-p))^y}{y!} \\
&= e^{-\lambda p}
\end{aligned}$$

To simplify the last summation we observed that the sum could be interpreted as the sum of the probabilities for a Poisson($\lambda(1-p)$) random variable, which is equal to 1.

Alternative solution: Since we know that Y is the limit of $Y_n \sim \text{Bin}(n, q_n = \lambda/n)$, we can write $\mathbb{P}(X > Y) = \lim_{n\to\infty} \mathbb{P}(X > Y_n)$, where

$$
\begin{aligned}
\mathbb{P}(X > Y_n) &= \sum_{k=0}^{n} \mathbb{P}(Y_n = k)\mathbb{P}(X > k \mid Y_n = k) = \sum_{k=0}^{n} \binom{n}{k} q_n^k (1-q_n)^{n-k} \cdot (1-p)^k \\
&= \sum_{k=0}^{n} \binom{n}{k} (q_n(1-p))^k (1-q_n)^{n-k} = (q_n(1-p) + (1-q_n))^n \\
&= \left(1 - \frac{\lambda p}{n}\right)^n
\end{aligned}
$$

which as $n \to \infty$ becomes $e^{-\lambda p}$ as above.

(b) 1, the max of X, Y is always at least X.

(c) $P(Z \le Y) = P(\max(X,Y) \le Y) = P(X \le Y) = 1 - P(X > Y) = 1 - e^{-\lambda p}$

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 $f_X(x) = \exp(-x)$. The board's radius is 4 units.

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

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

(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:

(a) The CDF of an exponential is $\mathbb{P}[X \le x] = 1 - \exp(-x)$. Therefore,

$$\mathbb{P}[X \le 4] = 1 - \exp(-4).$$

(b) We are given that the dart must be within the board, which means that the dart is at least 4 units away from the center. We can use the definition of conditional probability:

$$\mathbb{P}[X \le 1 \mid X \le 4] = \frac{\mathbb{P}[X \le 1 \cap X \le 4]}{\mathbb{P}[X \le 4]} = \frac{\mathbb{P}[X \le 1]}{\mathbb{P}[X \le 4]} = \frac{1 - \exp(-1)}{1 - \exp(-4)}.$$

(c)

$$
\begin{aligned}
\mathbb{E}[\text{score}] &= \int_0^1 4\exp(-x)\,\mathrm{d}x + \int_1^2 3\exp(-x)\,\mathrm{d}x + \int_2^3 2\exp(-x)\,\mathrm{d}x + \int_3^4 \exp(-x)\,\mathrm{d}x \\
&= 4\big(-\exp(-1) + 1\big) + 3\big(-\exp(-2) + \exp(-1)\big) + 2\big(-\exp(-3) + \exp(-2)\big) \\
&\quad + \big(-\exp(-4) + \exp(-3)\big) \\
&= 4 - \exp(-1) - \exp(-2) - \exp(-3) - \exp(-4).
\end{aligned}
$$

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.]

(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 $\mathbb{P}(X_1 \le x \mid X_1 + X_2 = t)$. 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:

(a) Let $y > 0$. Observe that if $X_1 + X_2 \le y$, then since $X_1, X_2 \ge 0$, it follows that $X_1 \le y$ and $X_2 \le y - X_1$.

$$\begin{aligned}
\mathbb{P}(Y \le y) = \mathbb{P}(X_1 \le y, X_2 \le y - X_1) &= \int_0^y \int_0^{y-x_1} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2)\, \mathrm{d}x_2\, \mathrm{d}x_1 \\
&= \lambda^2 \int_0^y \exp(-\lambda x_1) \cdot \frac{1 - \exp(-\lambda(y - x_1))}{\lambda}\, \mathrm{d}x_1 \\
&= \lambda \int_0^y \big(\exp(-\lambda x_1) - \exp(-\lambda y)\big)\, \mathrm{d}x_1 = \lambda \Big(\frac{1 - \exp(-\lambda y)}{\lambda} - y \exp(-\lambda y)\Big).
\end{aligned}$$

Upon differentiating the CDF, we have

$$\begin{aligned}
f_Y(y) = \frac{\mathrm{d}}{\mathrm{d}y} \mathbb{P}(Y \le y) &= \lambda \exp(-\lambda y) - \lambda \exp(-\lambda y) + \lambda^2 y \exp(-\lambda y) \\
&= \lambda^2 y \exp(-\lambda y), \qquad y > 0.
\end{aligned}$$

Alternative solution: Since X_1 and X_2 are limits of X_1^n/n and X_2^n/n, where X_1^n and X_2^n are independent $\text{Geom}(p_n = \lambda/n)$, we know that $f_Y(y)\mathrm{d}y = \lim_{n\to\infty} \mathbb{P}\left[(X_1^n + X_2^n)/n = y\right]$, i.e. $f(y) = \lim_{n\to\infty} n\mathbb{P}\left[X_1^n + X_2^n = ny\right]$. But from worksheet 11b we know that

$$n\mathbb{P}[X_1^n + X_2^n = ny] = n(ny-1)(1-p_n)^{ny-2} p_n^2 = \lambda^2 \left(y - \frac{1}{n}\right)\left(1 - \frac{\lambda}{n}\right)^{ny-2},$$

which as $n \to \infty$ converges to $\lambda^2 y e^{-\lambda y}$ as desired.

(b) Let $0 \le x \le t$. Following the hint, we have

$$\begin{aligned}
\mathbb{P}(X_1 \le x \mid X_1 + X_2 \in [t, t+\varepsilon]) &= \frac{\mathbb{P}(X_1 \le x, X_1 + X_2 \in [t, t+\varepsilon])}{\mathbb{P}(X_1 + X_2 \in [t, t+\varepsilon])} \\
&= \frac{\mathbb{P}(X_1 \le x, X_2 \in [t - X_1, t - X_1 + \varepsilon])}{f_Y(t) \cdot \varepsilon} \\
&= \frac{\int_0^x \int_{t-x_1}^{t-x_1+\varepsilon} \lambda \exp(-\lambda x_1) \lambda \exp(-\lambda x_2)\, \mathrm{d}x_2\, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} \\
&= \frac{\lambda^2 \int_0^x \exp(-\lambda x_1) \exp(-\lambda(t - x_1)) \varepsilon\, \mathrm{d}x_1}{\lambda^2 t \exp(-\lambda t) \cdot \varepsilon} = \frac{\int_0^x \mathrm{d}x_1}{t} = \frac{x}{t}.
\end{aligned}$$

This means that the density is

$$f_{X_1|X_1+X_2}(x \mid t) = \frac{\mathrm{d}}{\mathrm{d}x}\mathbb{P}(X \le x \mid X_1 + X_2 = t) = \frac{1}{t}, \qquad x \in [0,t],$$

which means that conditioned on $X_1 + X_2 = t$, X_1 is actually uniform on the interval $[0,t]$!

Alternative solution: Using the discrete approximations X_1^n/n and X_2^n/n as in the alternative solution to part (a), we have

$$\begin{aligned} n \cdot \mathbb{P}(X_1^n = xn \mid X_1^n + X_2^n = tn) &= n\frac{\mathbb{P}(X_1^n = xn \cap X_2^n = tn - xn)}{\mathbb{P}(X_1^n + X_2^n = tn)} = n\frac{(1-p_n)^{xn-1}p_n(1-p_n)^{tn-xn-1}p_n}{(tn-1)(1-p_n)^{tn-2}p_n^2} \\ &= \frac{1}{t - 1/n}, \end{aligned}$$

which converges to $1/t$ as $n \to \infty$ just like before.

6 Uniform Means

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

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

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

Solution:

(a) To calculate $\mathbb{P}(Y > y)$, where $y \in [0,1]$, this means that each X_i is greater than y, for $i = 1, \ldots, n$, so $\mathbb{P}(Y > y) = (1-y)^n$. We then use the tail sum formula:

$$\mathbb{E}(Y) = \int_0^1 \mathbb{P}(Y > y)\,\mathrm{d}y = \int_0^1 (1-y)^n\,\mathrm{d}y = -\frac{1}{n+1}(1-y)^{n+1}\Big|_0^1 = \frac{1}{n+1}.$$

Alternative Solution 1:

As explained above, $\mathbb{P}[Y \le y] = 1 - (1-y)^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(y) = n(1-y)^{n-1}$.

Then

$$\mathbb{E}(Y) = \int_0^1 y \cdot n(1-y)^{n-1}\,\mathrm{d}y.$$

Perform a u substitution, where $u = 1 - y$ and $\mathrm{d}u = -\mathrm{d}y$. We see:

$$\begin{aligned} \mathbb{E}(Y) &= n \cdot \int_0^1 -(1-u) \cdot u^{n-1}\,\mathrm{d}u = n \cdot \int_0^1 (u^n - u^{n-1})\,\mathrm{d}u = n\left[\frac{u^{n+1}}{n+1} - \frac{u^n}{n}\right]_{u=0}^1 \\ &= n\left[\frac{(1-y)^{n+1}}{n+1} - \frac{(1-y)^n}{n}\right]_{y=0}^1 = n\left[0 - \left(\frac{1}{n+1} - \frac{1}{n}\right)\right] = n\left[\frac{1}{n} - \frac{1}{n+1}\right] = \frac{1}{n+1}. \end{aligned}$$

Alternative Solution 2:
Consider adding another independent uniform variable X_{n+1}. $\mathbb{P}(X_{n+1} < Y)$ is just the probability that X_{n+1} is the minimum, which is $1/(n+1)$ by symmetry since all the X_i's are identical. It so happens that because X_{n+1} is a uniform variable on [0,1], this probability is equal to $\mathbb{E}(Y)$. Let f_Y denote the PDF of Y.

$$\begin{aligned}
\mathbb{P}(X_{n+1} < Y) &= \int_0^1 \mathbb{P}(X_{n+1} < y \mid Y = y) f_Y(y)\,\mathrm{d}y \\
&= \int_0^1 \mathbb{P}(X_{n+1} < y) f_Y(y)\,\mathrm{d}y && \text{(by independence)} \\
&= \int_0^1 y f_Y(y)\,\mathrm{d}y && \text{(CDF of the uniform distribution)} \\
&= \mathbb{E}(Y).
\end{aligned}$$

Alternative Solution 3:
Since $X_1, \ldots, X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. Therefore, $\mathbb{E}(Y) = 1/(n+1)$, the position of the smallest indicator.

(b) We could use the tail sum formula, but it turns out that the CDF is in a form that makes it easy to take an integral. If $Z \le z$, where $z \in [0,1]$, each X_i must be less than z, which happens with probability z, so $\mathbb{P}[Z \le z] = z^n$. This gives us the CDF, and if we take its derivative we'll get the probability density function $f(z) = nz^{n-1}$. Then

$$\mathbb{E}(Z) = \int_0^1 z \cdot nz^{n-1}\,\mathrm{d}z = \int_0^1 nz^n\,\mathrm{d}z = \left[n \cdot \frac{z^{n+1}}{n+1}\right]_{z=0}^1 = \frac{n}{n+1}.$$

Alternative Solution:
As in the previous part, add another independent uniform random variable X_{n+1}. The probability $\mathbb{P}(X_{n+1} > Z)$ is just the probability that X_{n+1} is the maximum, which is $1/(n+1)$ by symmetry.

$$\begin{aligned}
\mathbb{P}(X_{n+1} > Z) &= \int_0^1 \mathbb{P}(X_{n+1} > z \mid Z = z) f_Z(z)\,\mathrm{d}z = \int_0^1 \mathbb{P}(X_{n+1} > z) f_Z(z)\,\mathrm{d}z \\
&= \int_0^1 (1-z) f_Z(z)\,\mathrm{d}z = \int_0^1 f_Z(z)\,\mathrm{d}z - \int_0^1 z f_Z(z)\,\mathrm{d}z \\
\frac{1}{n+1} &= 1 - \mathbb{E}(Z) \\
\mathbb{E}(Z) &= \frac{n}{n+1}
\end{aligned}$$

Alternative Solution 2:
Since $X_1, \ldots, X_n$ are i.i.d., their values split the interval $[0,1]$ into $n+1$ sections, and we expect these sections to be of equal length because they are uniformly distributed. The expectation of the smallest X_i is $1/(n+1)$, the expectation of the second smallest is $2/(n+1)$, etc. Therefore, $\mathbb{E}(Z) = n/(n+1)$, the position of the largest indicator.