CS 70 Discrete Mathematics and Probability Theory
Spring 2026 Sinclair, Song

DIS 12A

Geometric and Poisson Distributions

Note 18

First, let's recap some key terms!

Random Variable: A random variable X is a function from $\Omega \to \mathbb{R}$, mapping the possible outcomes to real numbers. Note that this function itself is not random; the *outcomes* are random. We define

$$\mathbb{P}[X = k] = \mathbb{P}[\{\omega \in \Omega : X(\omega) = k\}].$$

Distribution of a random variable: the set of all $(k, \mathbb{P}[X = k])$, describing the probability of attaining each value of the random variable.

And now, the focus of today's discussion:

Geometric Distribution: $X \sim \text{Geometric}(p)$; X represents the number of independent trials until the first success (including the success), where p is the probability of success in each trial.

Poisson Distribution: $X \sim \text{Poisson}(\lambda)$; X represents the number of occurrences of an event in one unit of time, if on average there are λ occurrences in one unit of time. The distribution is described by the following:

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad \mathbb{E}[X] = \lambda, \qquad \text{Var}(X) = \lambda.$$

Further, if $X \sim \text{Poisson}(\lambda_x)$ and $Y \sim \text{Poisson}(\lambda_y)$ are independent, then $X + Y \sim \text{Poisson}(\lambda_x + \lambda_y)$.

1 Head Count II

Note 18

Consider a coin with $\mathbb{P}[\text{Heads}] = 3/4$. Suppose you flip the coin until you see heads for the first time, and define X to be the number of times you flipped the coin.

(a) What is $\mathbb{P}[X = k]$, for some $k \geq 1$? Express your answer in terms of k. (Do not just copy down a formula—re-derive it yourself!)

(b) What is the name of the distribution of X, and what are its parameters?

(c) What is $\mathbb{P}[X > k]$, for some $k \geq 0$? (You should not have any summations.)

(d) What is $\mathbb{P}[X < k]$, for some $k \geq 1$? (You should not have any summations.)

(e) What is $\mathbb{P}[X > k \mid X > m]$, for some $k \geq m \geq 0$? Show that your answer is equal to $\mathbb{P}[X > k - m]$. Why do we call this the memoryless property?

(f) Suppose $Y \sim \text{Geometric}(p)$ and $Z \sim \text{Geometric}(q)$ are independent. Find the distribution of $\min(Y,Z)$ and justify your answer.

Hint: consider flipping two coins (with $\mathbb{P}[\text{Heads}] = p$ and $\mathbb{P}[\text{Heads}] = q$ respectively) simultaneously.

Solution:

(a) If we flipped k times, then we had $k-1$ tails and 1 head, in that order, giving us

$$\mathbb{P}[X = k] = \frac{3}{4}\left(1 - \frac{3}{4}\right)^{k-1} = \frac{3}{4}\left(\frac{1}{4}\right)^{k-1}.$$

(b) $X \sim \text{Geometric}(\frac{3}{4})$

(c) If we had to flip *more than* k times before seeing our first heads, then our first k flips must have been tails, giving us

$$\mathbb{P}[X > k] = \left(1 - \frac{3}{4}\right)^{k} = \left(\frac{1}{4}\right)^{k}.$$

You can alternatively write as the sum $\sum_{i=k+1}^{\infty} \mathbb{P}[X = i] = \sum_{i=k+1}^{\infty} \frac{3}{4} * (\frac{1}{4})^{i-1} = \frac{3}{4} * (\frac{1}{4})^{k} * \frac{1}{1-1/4} = (\frac{1}{4})^{k}$ using the formula for an infinite geometric sum

(d) Notice $\mathbb{P}[X < k] = 1 - \mathbb{P}[X \geq k] = 1 - \mathbb{P}[X > k-1]$ since X can only take on integer values. Along similar lines to the previous part, we then have

$$\mathbb{P}[X < k] = 1 - \mathbb{P}[X > k-1] = 1 - \left(1 - \frac{3}{4}\right)^{k-1} = 1 - \left(\frac{1}{4}\right)^{k-1}.$$

(e) By part (c), we have

$$\mathbb{P}[X > k \mid X > m] = \frac{\mathbb{P}[X > k \cap X > m]}{\mathbb{P}[X > m]} = \frac{\mathbb{P}[X > k]}{\mathbb{P}[X > m]} = \left(\frac{1}{4}\right)^{k-m}.$$

However, note that this is exactly $\mathbb{P}[X > k-m]$. The reason this makes sense is that if we want to compute the probability that the first heads occurs after k flips, and we know that the first heads occurs after m flips, then the first m flips are tails. Thus, by the independence of the coin flips, the first m flips don't matter, and so we only need to compute the probability that the first heads occurs after $k-m$ flips. This is called the **memorylessness property** of the geometric distribution because having flipped the coin m times without seeing heads doesn't affect the distribution of future flips.

(f) Let Y be the number of coins we flip until we see a heads from flipping a coin with bias p, and let Z similarly be the number of coins we flip until we see a heads from flipping a coin with bias q.

Imagine we flip the bias p coin and the bias q coin at the same time. The minimum of the two random variables represents how many simultaneous flips occur before at least one head is seen.

The probability of not seeing a head at all on any given simultaneous flip is $(1-p)(1-q)$; this corresponds to a failure. This means that the probability that there will be a success on any particular trial is $1-(1-p)(1-q) = p+q-pq$. Therefore, $\min(Y,Z) \sim \text{Geometric}(p+q-pq)$.

Alternative 1: We can also solve this algebraically. The probability that $\min(Y,Z) = k$ for some positive integer k is the probability that the first $k-1$ coin flips for both Y and Z were tails, and we get heads on the kth toss (this can come from either Y or Z). Specifically, this occurs with probability

$$((1-p)(1-q))^{k-1} \cdot (p+q-pq)$$

We recognize this as the formula for a geometric random variable with parameter $p+q-pq$.

Alternative 2: 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. Let $W = \min(Y,Z)$. We can work with $\mathbb{P}[W \geq k]$ rather than with $\mathbb{P}[W = k]$; clearly the values $\mathbb{P}[W \geq k]$ specify the values $\mathbb{P}[W = k]$ since $\mathbb{P}[W = k] = \mathbb{P}[W \geq k] - \mathbb{P}[W \geq (k+1)]$, so it suffices to calculate them instead. We then get the following argument:

$$\begin{aligned}
\mathbb{P}[W \geq k] &= \mathbb{P}[\min(Y,Z) \geq k] \\
&= \mathbb{P}[(Y \geq k) \cap (Z \geq k)] \\
&= \mathbb{P}[Y \geq k] \cdot \mathbb{P}[Z \geq k] && \text{since } Y,Z \text{ are independent} \\
&= (1-p)^{k-1}(1-q)^{k-1} && \text{since } Y,Z \text{ are geometric} \\
&= ((1-p)(1-q))^{k-1} \\
&= (1-p-q+pq)^{k-1}.
\end{aligned}$$

This is the tail probability of a geometric distribution with parameter $p+q-pq$, thus we can conclude that $W \sim Geom(p+q-pq)$, which is the same result as before!

2 Unreliable Servers

Note 18

A Google competitor owns a warehouse consisting of a very large number of servers (a server farm). On any given day, every server in the farm may go down with the same probability, independently of all other servers and of what happens on other days. On average, 4 servers go down in the cluster per day.

(a) What is an appropriate distribution to model the number of servers that crash on any given day for a certain cluster? (Give the name and parameter(s) of the distribution.)

(b) Compute the expected value and variance of the number of crashed servers on a given day for a certain cluster.

(c) Compute the probability that strictly less than 3 servers crashed on a given day for a certain cluster.

(d) Compute the probability that at least 3 servers crashed on a given day for a certain cluster.

(e) Compute the probability that exactly 6 servers crashed over a given two-day period for a certain cluster.

(f) After a recent software update, servers never crash except when rebooting. Every morning each server is rebooted and each time a server is rebooted, it has a 5% chance of crashing. Which distribution can we use to model the day of the first crash?

(g) Compute the expected day of a computer's first crash.

(h) Five of these upgraded computers are connected to create a server farm; these computers are rebooted at the same time every morning. What is the expected day of the cluster's first crash?

Solution:

(a) Because each server goes down independently of the other servers, and with the same probability, the number of servers that crash on a given day follows a binomial distribution Binomial(n, p), where n is the number of servers and p is the probability of each individual server crashing on any given day.

Since on average, 4 servers crash per day, we have $p = \frac{4}{n}$. We are given that the number of servers in the cluster is large, so $n \gg p$ and we can model the number of servers that crash on a given day as a Poisson distribution with $\lambda = 4$.

(b) Recall that the expectation and variance of a Poisson distribution with parameter λ are both equal to λ and in this case $\lambda = 4$.

(c) To compute the probability that fewer than 3 servers went down, we must add the probabilities that 0 servers go down, 1 server goes down, and the probability that 2 servers go down. The PMF of the Poisson distribution is

$$\mathbb{P}[X = i] = \frac{\lambda^i}{i!} e^{-\lambda},$$

thus

$$\begin{aligned}\mathbb{P}[X = 0 \text{ or } X = 1 \text{ or } X = 2] &= e^{-4} + 4e^{-4} + \frac{4^2}{2} e^{-4} \\ &= e^{-4} + 4e^{-4} + 8e^{-4} \\ &= 13e^{-4}\end{aligned}$$

(d) The probability that at least 3 servers crashed is equal to complement of the probability that fewer than 3 servers crashed. This gives us

$$\mathbb{P}[\text{at least 3 servers crashed}] = 1 - \mathbb{P}[\text{fewer than 3 servers crashed}] = 1 - 13e^{-4}.$$

(e) Let X_1 represent the servers that crash on day 1, and let X_2 represent the servers that crash on day 2. The total number of crashes on a particular day can be modeled as a Poisson random variable, so we have $X_1 \sim \text{Poisson}(\lambda_1 = 4)$ and $X_2 \sim \text{Poisson}(\lambda_2 = 4)$ with X_1 and X_2 independent of each other, since crashes on different days are independent.

Thus, the total number of crashes over both days is $Y = X_1 + X_2$. Since the sum of two independent Poisson random variables is also Poisson, we know that Y is Poisson with parameter $\lambda = \lambda_1 + \lambda_2 = 4 + 4 = 8$.

This means that the probability of exactly 6 crashes among these two days is $\mathbb{P}[Y = 6] = \frac{8^6}{6!}e^{-8}$.

(f) Since there is one reboot per day and each reboot independently crashes with probability 0.05, the day of the first crash is geometric: $X \sim \text{Geometric}(0.05)$.

(g) The expected value of a geometric random variable is $\frac{1}{p}$, so $\mathbb{E}[X] = \frac{1}{0.05} = 20$ days.

(h) On each day, the probability of at least one computer crashing is $1 - \mathbb{P}[\text{all 5 computers start without crashing}]$ $1 - 0.95^5$. The day of the cluster's first crash is therefore $X \sim \text{Geometric}(1 - 0.95^5)$, giving $\mathbb{E}[X] = \frac{1}{1-0.95^5} \approx 4.42$ days.

3 Shuttles and Taxis at Airport

Note 18

In front of terminal 3 at San Francisco Airport is a pickup area where shuttles and taxis arrive according to a Poisson distribution. The shuttles arrive at a rate $\lambda_1 = 1/20$ (i.e. 1 shuttle per 20 minutes) and the taxis arrive at a rate $\lambda_2 = 1/10$ (i.e. 1 taxi per 10 minutes) starting at 00:00. The shuttles and the taxis arrive independently.

(a) Write in terms of a Poisson distribution, the distribution of taxis between 00:00 and 00:01:

(b) What is the distribution of the following:

(i) The number of taxis that arrive between times 00:00 and 00:20?

(ii) The number of shuttles that arrive between times 00:00 and 00:20?

(iii) The total number of vehicles (shuttles and taxis) that arrive between times 00:00 and 00:20?

(c) What is the probability that exactly 1 shuttle and 3 taxis arrive between times 00:00 and 00:20?

(d) Given that exactly 1 pickup vehicle arrived between times 00:00 and 00:20, what is the conditional probability that this vehicle was a taxi?

(e) Suppose you reach the pickup area at 00:20. You learn that you missed 3 taxis and 1 shuttle in those 20 minutes. What is the probability that you need to wait for more than 10 mins until either a shuttle or a taxi arrives?

Solution:

(a) Let $T([0,1])$ denote the number of taxis that arrive between times 00:00 and 00:01. This interval has length 1 minute, so the number of taxis $T([0,1])$ arriving in this interval is distributed according to Poisson(λ_2) = Poisson$(\frac{1}{10})$, i.e.

$$\mathbb{P}[T([0,1]) = t] = \frac{\frac{1}{10}^t e^{-\frac{1}{10}}}{t!}, \text{ for } t = 0,1,2,\ldots.$$

(b) (i) Let $T([0,20])$ denote the number of taxis that arrive between times 00:00 and 00:20. This interval has length 20 minutes, so the number of taxis $T([0,20])$ arriving in this interval is distributed according to Poisson$(\lambda_2 \cdot 20)$ = Poisson(2), i.e.

$$\mathbb{P}[T([0,20]) = t] = \frac{2^t e^{-2}}{t!}, \text{ for } t = 0,1,2,\ldots.$$

(ii) Let $S([0,20])$ denote the number of shuttles that arrive between times 00:00 and 00:20. This interval has length 20 minutes, so the number of shuttles $S([0,20])$ arriving in this interval is distributed according to Poisson$(\lambda_1 \cdot 20)$ = Poisson(1), i.e.

$$\mathbb{P}[S([0,20]) = s] = \frac{1^s e^{-1}}{s!}, \text{ for } s = 0,1,2,\ldots.$$

(iii) Let $N([0,20]) = S([0,20]) + T([0,20])$ denote the total number of pickup vehicles (taxis and shuttles) arriving between times 00:00 and 00:20. Since the sum of independent Poisson random variables is Poisson distributed with parameter given by the sum of the individual parameters, we have $N[(0,20)] \sim$ Poisson(3), i.e.

$$\mathbb{P}[N([0,20]) = n] = \frac{3^n e^{-3}}{n!}, \text{ for } n = 0,1,2,\ldots.$$

(c) We have

$$\mathbb{P}[T([0,20]) = 3] = \frac{2^3 e^{-2}}{3!} \text{ and } \mathbb{P}[S([0,20]) = 1] = \frac{1^1 e^{-1}}{1!}.$$

Since the taxis and the shuttles arrive independently, the probability that exactly 3 taxis and 1 shuttle arrive in this interval is given by the product of their individual probabilities, i.e.

$$\frac{2^3 e^{-2}}{3!}\frac{1^1 e^{-1}}{1!} = \frac{4}{3}e^{-3} \approx 0.0664.$$

(d) Let A be the event that exactly 1 taxi arrives between times 00:00 and 00:20. Let B be the event that exactly 1 vehicle arrives between times 00:00 and 00:20. We have

$$\mathbb{P}[B] = \frac{3^1 e^{-3}}{1!}.$$

Event $A \cap B$ is the event that exactly 1 taxi and 0 shuttles arrive between times 00:00 and 00:20. Hence

$$\mathbb{P}[A \cap B] = \frac{2^1 e^{-2}}{1!}\frac{1^0 e^{-1}}{0!}.$$

Thus, we get

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]} = 2/3.$$

(e) The event that you need to wait for more than 10 minutes starting 00:20 is equivalent to the event that no vehicle arrives between times 00:20 and 00:30. Let $N[20,30]$ denote the number of vehicles that arrive between times 00:20 and 00:30. This interval has length 10 minutes, so $N[(20,30)] \sim \text{Poisson}((\lambda_1 + \lambda_2) \cdot 10) = \text{Poisson}(3/2)$. Since Poisson arrivals in disjoint intervals are independent, we have

$$\mathbb{P}[N([20,30]) = 0 | T([0,20]) = 3, S([0,20]) = 1] = \mathbb{P}[N([20,30]) = 0] = \frac{1.5^0 e^{-1.5}}{0!} = e^{-1.5} \approx 0.2231.$$