CS 70 Discrete Mathematics and Probability Theory
Summer 2025 Tate

DIS 04D

Random Variables Intro

Note 15
Note 19

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.

Bernoulli Distribution: $X \sim \text{Bernoulli}(p)$; X represents the outcome of a biased coin flip. X is oftentimes also called an *indicator random variable* of an event with probability p. The distribution is described by the following:

$$\mathbb{P}[X = k] = \begin{cases} p & \text{if } k = 1 \\ 1 - p & \text{if } k = 0 \end{cases}$$

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

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:

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

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 Cookie Jars

Note 15

You have two jars of cookies, each of which starts with n cookies initially. Every day, when you come home, you pick one of the two jars randomly (each jar is chosen with probability $1/2$) and eat one cookie from that jar. One day, you come home and reach inside one of the jars of cookies, but you find that is empty! Let X be the random variable representing the number of remaining cookies in non-empty jar at that time. What is the distribution of X?

Solution: Assume that you found jar 1 empty; the probability that $X = k$ and you found jar 1 empty is computed as follows.

In order for there to be k cookies remaining, you must have eaten a cookie for $2n-k$ days, and then you must have chosen jar 1 (to discover that it is empty). Within those $2n-k$ days, exactly n of those days you chose jar 1. The probability of this is $\binom{2n-k}{n}2^{-(2n-k)}$.

Furthermore, the probability that you then discover jar 1 is empty the day after is $1/2$. So, the probability that $X = k$ and you discover jar 1 empty is $\binom{2n-k}{n}2^{-(2n-k+1)}$. However, we assumed that we discovered jar 1 to be empty; the probability that $X = k$ and jar 2 is empty is the same by symmetry, so the overall probability that $X = k$ is:

$$\mathbb{P}[X = k] = \binom{2n-k}{n}\frac{1}{2^{2n-k}}, \qquad k \in \{0,\ldots,n\}.$$

2 Class Enrollment

Note 15
Note 19

Lydia has just started her CalCentral enrollment appointment. She needs to register for a geography class and a history class. There are no waitlists, and she can attempt to enroll once per day in either class or both. The CalCentral enrollment system is strange and picky, so the probability of enrolling successfully in the geography class on each attempt is p_g and the probability of enrolling successfully in the history class on each attempt is p_h. Also, these events are independent.

(a) Suppose Lydia begins by attempting to enroll in the geography class everyday and gets enrolled in it on day G. What is the distribution of G?

(b) Suppose she is not enrolled in the geography class after attempting each day for the first 7 days. What is $\mathbb{P}[G = i \mid G > 7]$, the conditional distribution of G given $G > 7$?

(c) Once she is enrolled in the geography class, she starts attempting to enroll in the history class from day $G+1$ and gets enrolled in it on day H. Find the expected number of days it takes Lydia to enroll in both the classes, i.e. $\mathbb{E}[H]$.

Suppose instead of attempting one by one, Lydia decides to attempt enrolling in both the classes from day 1. Let G be the number of days it takes to enroll in the geography class, and H be the number of days it takes to enroll in the history class.

(d) What is the distribution of G and H now? Are they independent?

(e) Let A denote the day she gets enrolled in her first class and let B denote the day she gets enrolled in both the classes. What is the distribution of A?

(f) What is the expected number of days it takes Lydia to enroll in both classes now, i.e. $\mathbb{E}[B]$?

(g) What is the expected number of classes she will be enrolled in by the end of 30 days?

Solution:

(a) $G \sim \text{Geometric}(p_g)$.

(b) Given that $G > 7$, the random variable G takes values in $\{8, 9, \dots\}$. For $i = 8, 9, \dots$,

$$\begin{aligned}\mathbb{P}[G = i \mid G > 7] &= \frac{\mathbb{P}[G = i \wedge G > 7]}{\mathbb{P}[G > 7]} = \frac{\mathbb{P}[G = i]}{\mathbb{P}[G > 7]} \\ &= \frac{p_g(1-p_g)^{i-1}}{(1-p_g)^7} = p_g(1-p_g)^{i-8}\end{aligned}$$

If K denotes the additional number of days it takes to get enrolled in the geography class after day 7, i.e. $K = G - 7$, then conditioned on $G > 7$, the random variable K has the geometric distribution with parameter p_g. Note that this is the same as the distribution of G. This is known as the memoryless property of geometric distribution.

(c) We have $H - G \sim \text{Geometric}(p_h)$. This means that $\mathbb{E}[G] = \frac{1}{p_g}$ and $\mathbb{E}[H-G] = \frac{1}{p_h}$, and as such

$$\mathbb{E}[H] = \mathbb{E}[G] + \mathbb{E}[H - G] = \frac{1}{p_g} + \frac{1}{p_h}.$$

(d) $G \sim \text{Geometric}(p_g)$, $H \sim \text{Geometric}(p_h)$. Yes they are independent.

(e) We have $A = \min\{G, H\}$ and $B = \max\{G, H\}$. We also use the following definition of the minimum:

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

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

$$\begin{aligned}\mathbb{P}[A = k] &= \mathbb{P}[\min(G,H) = k] \\ &= \mathbb{P}[(G = k) \cap (H \ge k)] + \mathbb{P}[(H = k) \cap (G > k)] \\ &= \mathbb{P}[G = k] \cdot \mathbb{P}[H \ge k] + \mathbb{P}[H = k] \cdot \mathbb{P}[G > k] && (G,\ H \text{ are independent}) \\ &= [(1-p_g)^{k-1} p_g](1-p_h)^{k-1} + [(1-p_h)^{k-1} p_h](1-p_g)^k && (G,\ H \text{ are geometric}) \\ &= ((1-p_g)(1-p_h))^{k-1}(p_g + p_h(1-p_g)) \\ &= (1 - p_g - p_h + p_h p_g)^{k-1}(p_g + p_h - p_g p_h).\end{aligned}$$

But this final expression is precisely the probability that a geometric RV with parameter $p_g + p_h - p_g p_h$ takes the value k. Hence $A \sim \text{Geom}(p_g + p_h - p_g p_h)$.

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 $\mathbb{P}[A \ge k]$ rather than with $\mathbb{P}[A = k]$; clearly the values $\mathbb{P}[A \ge k]$ specify the values $\mathbb{P}[A = k]$ since $\mathbb{P}[A = k] = \mathbb{P}[A \ge k] - \mathbb{P}[A \ge (k+1)]$, so it suffices to calculate them instead.

We then get the following argument:

$$\begin{aligned}
\mathbb{P}[A \geq k] &= \mathbb{P}[\min(G,H) \geq k] \\
&= \mathbb{P}[(G \geq k) \cap (H \geq k)] \\
&= \mathbb{P}[G \geq k] \cdot \mathbb{P}[H \geq k] && (G,\ H \text{ are independent}) \\
&= (1-p_g)^{k-1}(1-p_h)^{k-1} && (G,\ H \text{ are geometric}) \\
&= ((1-p_g)(1-p_h))^{k-1} \\
&= (1-p_g-p_h+p_g p_h)^{k-1}.
\end{aligned}$$

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

(f) From part (e) we get $\mathbb{E}[A] = \frac{1}{p_g+p_h-p_g p_h}$. From part (d) we have $\mathbb{E}[G] = \frac{1}{p_g}$ and $\mathbb{E}[H] = \frac{1}{p_h}$.

We now observe that $\min\{g,h\} + \max\{g,h\} = g + h$; using linearity of expectation, this means that $\mathbb{E}[A] + \mathbb{E}[B] = \mathbb{E}[G] + \mathbb{E}[H]$. As such, we have

$$\mathbb{E}[B] = \frac{1}{p_g} + \frac{1}{p_h} - \frac{1}{p_g + p_h - p_g p_h}.$$

(g) Let I_G and I_H be the indicator random variables of the events "$G \leq 30$" and "$H \leq 30$" respectively. Then $I_G + I_H$ is the number of classes she will be enrolled in within 30 days. Hence the answer is

$$\mathbb{E}[I_G] + \mathbb{E}[I_H] = \mathbb{P}[G \leq 30] + \mathbb{P}[H \leq 30] = 1 - (1-p_g)^{30} + 1 - (1-p_h)^{30}.$$

3 Fishy Computations

Note 19

Assume for each part that the random variable can be modelled by a Poisson distribution.

(a) Suppose that on average, a fisherman catches 20 salmon per week. What is the probability that he will catch exactly 7 salmon this week?

(b) Suppose that on average, you go to Fisherman's Wharf twice a year. What is the probability that you will go at most once in 2024?

(c) Suppose that in March, on average, there are 5.7 boats that sail in Laguna Beach per day. What is the probability there will be *at least* 3 boats sailing throughout the *next two days* in Laguna?

(d) Denote $X \sim \text{Pois}(\lambda)$. Prove that

$$\mathbb{E}[X f(X)] = \lambda\, \mathbb{E}[f(X+1)]$$

for any function f.

Solution:

(a) Let X be the number of salmon the fisherman catches per week. $X \sim \text{Poisson}(20 \text{ salmon/week})$, so

$$\mathbb{P}[X = 7 \text{ salmon/week}] = \frac{20^7}{7!} e^{-20} \approx 5.23 \cdot 10^{-4}.$$

(b) Similarly $X \sim \text{Poisson}(2)$, so

$$\mathbb{P}[X \le 1] = \frac{2^0}{0!} e^{-2} + \frac{2^1}{1!} e^{-2} \approx 0.41.$$

(c) Let X_1 be the number of sailing boats on the next day, and X_2 be the number of sailing boats on the day after next. Now, we can model sailing boats on day i as a Poisson distribution $X_i \sim \text{Poisson}(\lambda = 5.7)$. Let Y be the number of boats that sail in the next two days. We are interested in $Y = X_1 + X_2$. We know that the sum of two independent Poisson random variables is Poisson. Thus, we have $Y \sim \text{Poisson}(\lambda = 5.7 + 5.7 = 11.4)$.

$$\begin{aligned}
\mathbb{P}[Y \ge 3] &= 1 - \mathbb{P}[Y < 3] \\
&= 1 - \mathbb{P}[Y = 0 \cup Y = 1 \cup Y = 2] \\
&= 1 - (\mathbb{P}[Y = 0] + \mathbb{P}[Y = 1] + \mathbb{P}[Y = 2]) \\
&= 1 - \left(\frac{11.4^0}{0!} e^{-11.4} + \frac{11.4^1}{1!} e^{-11.4} + \frac{11.4^2}{2!} e^{-11.4} \right) \\
&\approx 0.999.
\end{aligned}$$

(d) We apply the Law of the Unconscious Statistician,

$$\begin{aligned}
\mathbb{E}[Xf(X)] &= \sum_{x=0}^{\infty} x f(x)\, \mathbb{P}[X = x] \\
&= \sum_{x=0}^{\infty} x f(x) \frac{e^{-\lambda} \lambda^x}{x!} \\
&= \sum_{x=1}^{\infty} x f(x) \frac{e^{-\lambda} \lambda^x}{x!} \\
&= \lambda \sum_{x=1}^{\infty} f(x) \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \\
&= \lambda \sum_{x=0}^{\infty} f(x+1) \frac{e^{-\lambda} \lambda^x}{x!} \\
&= \lambda\, \mathbb{E}[f(X+1)]
\end{aligned}$$

as desired.