Chris Piech and Jerry Cain
CS109

Section #3
February 2, 2021

Section #3: Discrete and Random Variables

Overview of Section Materials

The warmup questions provided will help students practice concepts introduced in lectures. The section problems are meant to apply these concepts in more complex scenarios similar to what you will see in problem sets and quizzes.

1 Warmups

1.1 Website Visits

You have a website where only one visitor can be on the site at a time, but there is an infinite queue of visitors, so that immediately after a visitor leaves, a new visitor will come onto the website. On average, visitors leave your website after 5 minutes. Assume that the length of stay is exponentially distributed. What is the probability that a user stays more than 10 minutes, if we calculate this probability:

a. using the random variable X, defined as the length of stay of the user?

b. using the random variable Y, defined as the number of users who leave your website over a 10-minute interval?

If this problem doesn't convince you that the Poisson and Exponential RVc are coupled, then I'm not sure will! As defined above, $X \sim \text{Exp}(\lambda = \frac{1}{5})$.

$$P(X > 10) = 1 - F_X(10) = 1 - (1 - e^{-\lambda 10}) = e^{-2} \approx 0.1353$$

Alternatively, we have that Y is the number of users leaving on the website in the next 10 minutes. The average number of users leaving is 2 users per 10 minutes. $Y \sim \text{Poi}(\lambda = 2)$.

$$\begin{aligned} P(Y = 0) &= \frac{2^0 e^{-2}}{0!} \\ &= e^{-2} \approx 0.1353 \end{aligned}$$

1.2 Continuous Random Variables

Let X be a continuous random variable with the following probability density function:

$$f_X(x) = \begin{cases} c(e^{x-1} + e^{-x}) & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

a. Find the value of c that makes f_X a valid probability distribution.

b. What is $P(X > 0.75)$?

a. We need $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

$$
\begin{aligned}
\int_{-\infty}^{\infty} f_X(x)dx &= \int_0^1 c(e^{x-1} + e^{-x})dx \\
1 &= c\left[e^{x-1} - e^{-x}\right]_{x=0}^{1} \\
1 &= c(e^{1-1} - e^{-1} - (e^{0-1} - e^{-0})) \\
c &= \frac{1}{1 - e^{-1} - (e^{-1} - 1)} = \frac{1}{2 - \frac{2}{e}}
\end{aligned}
$$

b.

$$
\begin{aligned}
P(X > 0.75) &= \int_{0.75}^{1} c(e^{x-1} + e^{-x})dx \\
&= c\left[e^{x-1} - e^{-x}\right]_{x=0.75}^{1} \\
&= c\left(e^{1-1} - e^{-1} - (e^{0.75-1} - e^{-0.75})\right) \\
&= c\left(1 - e^{-1} - e^{-0.25} + e^{-0.75}\right) = \frac{1 - e^{-1} - e^{-0.25} + e^{-0.75}}{2 - \frac{2}{e}}
\end{aligned}
$$

2 Problems

2.1 Conditional Probabilities: Corrupt Hot-Dog-Eating Contest Judges

Preamble: We have three big tools for manipulating conditional probabilities:

- Definition of conditional probability: $P(EF) = P(E|F)P(F)$
- Law of Total Probability: $P(E) = P(EF) + P(EF^C) = P(E|F)P(F) + P(E|F^C)P(F^C)$
- Bayes Rule: $P(E|F) = \dfrac{P(F|E)P(E)}{P(F)} = \dfrac{P(F|E)P(E)}{P(F|E)P(E) + P(F|E^C)P(E^C)}$

This is a good time to commit these three to memory and start thinking about when each of them is useful.

Problem: Corrupted by their power, the judges running the popular game show Americas Next Top Hot Dog Eater have been taking bribes from many of the contestants. During each of two episodes, a given contestant is either allowed to stay on the show or is kicked off. If the contestant has been bribing the judges, she will be allowed to stay with probability 1. If the contestant has not been bribing the judges, she will be allowed to stay with probability 1/3, independent of what happens in earlier episodes. Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds.

a. If you pick a random contestant, what is the probability that she is allowed to stay during the first episode?

b. If you pick a random contestant, what is the probability that she is allowed to stay during both episodes?

c. If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she gets kicked off during the second episode?

d. If you pick a random contestant who was allowed to stay during the first episode, what is the probability that she was bribing the judge?

a. Let S_1 be the event she stayed during the first episode, and B the event she bribed the judges.

$$P(S_1) = P(S_1|B)P(B) + P(S_1|B^C)P(B^C) = 1 \cdot (1/4) + (1/3)(3/4) = 1/2$$

b. Let S_2 be the event she stayed during the second episode. Since staying in episodes are conditionally independent given whether she bribed the judges,

$$P(S_1 \cap S_2) = P(S_1 \cap S_2|B)P(B) + P(S_1 \cap S_2|B^C)P(B^C)$$

$$= P(S_1|B)P(S_2|B)P(B) + P(S_1|B^C)P(S_2|B^C)P(B^C) = 1^2(1/4) + (1/3)^2(3/4) = 1/3$$

c.

$$P(S_1 \cap S_2^C) = P(S_1|B)P(S_2^C|B)P(B) + P(S_1|B^C)P(S_2^C|B^C)P(B^C)$$

$$= 1 \cdot 0 \cdot (1/4) + (1/3) \cdot (2/3) \cdot (3/4) = 1/6$$

$$P(S_2^C|S_1) = \frac{P(S_1 \cap S_2^C)}{P(S_1)} = \frac{1/6}{1/2} = 1/3$$

d.

$$P(B|S_1) = \frac{P(S_1|B)P(B)}{P(S_1)} = \frac{(1)(1/4)}{1/2} = 1/2$$

2.2 More Bit Strings

Once again, we're sending sending bit strings across potentially noisy communication channels, just like last week. However, this week we're identifying bit string corruptions in a slightly different way. Now, whenever we want to send n bits of information, we send an extra as the $n + 1^{st}$ bit. Specifically, if the sum of the n data bits is even, the extra $n + 1^{st}$ bit sent is set to 0. If the sum of the n data bits is odd, then $n + 1^{st}$ bit appended is set to 1. If the recipient of the bit string adds all bits and gets an odd number, that recipient knows there's a problem and can request a repeat transmission. We'll assume that each bit is erroneously inverted with probability nonzero $p \leq 0.5$, and that all bit corruptions are independent of one another.

a. Assuming that n = 4 and p = 0.1, what is the probability the transmitted message has errors without being detected?

b. For arbitrary n and p, what is the probability that a bit string is flagged as bogus? You may leave it as a sum of O(n) terms.

c. Simplify your answer from part b by letting $a = \sum_{\text{odd}k} \binom{n+1}{k} p^k (1-p)^{n+1-k}$ and $b = \sum_{\text{even}k} \binom{n+1}{k} p^k (1-p)^{n+1-k}$ and then considering what $a+b$ and $a-b$ equal. Leverage the fact that, in general, $(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$.

a. Note that a bit string with an odd number of errors will be flagged as erroneous, but those bit strings with an even number of errors, regardless of n, will be taken as correct. That means the probability errors will go undetected is:

$$\binom{5}{2}(0.1)^2(0.9)^3 + \binom{5}{4}(0.1)^4(0.9)^1 = 0.07335 \tag{1}$$

b. The above generalizes to arbitrary n, so that the probability of interest is:

$$\sum_{\text{even } k \geq 2} \binom{n+1}{k} p^k (1-p)^{n+1-k} \tag{2}$$

c. If a and b are defined that way, then:

$$a + b = \sum_{\text{all}k} \binom{n+1}{k} p^k (1-p)^{n+1-k} = 1 \tag{3}$$

and

$$a - b = \sum_{\text{all}k} \binom{n+1}{k} (-p)^k (1-p)^{n+1-k} = (1-2p)^{n+1} \tag{4}$$

Solving for a, we arrive at:

$$a = \sum_{\text{even}k} \binom{n+1}{k} (-p)^k (1-p)^{n+1-k} = \frac{1 + (1-2p)^{n+1}}{2} \tag{5}$$

Now, a includes a term for no errors, so we need to subtract that one term off so it doesn't contribute. That leaves us with our final answer for general n and p, which is:

$$\frac{1 + (1-2p)^{n+1}}{2} - (1-p)^{n+1} \tag{6}$$