MIT OPENCOURSEWARE

18.440: Probability and Random Variables
Problem Set 5

Travis McVoy

March 3, 2024

CREATIVE COMMONS LICENSE STATEMENT

MIT OCW material is free under a creative commons license. Under said license, I believe I am free to "remix/adapt" work from MIT OCW provided that I follow the same license (people may use my work), I credit the original source, I do not use MIT's name, and I make note of any changes. To be clear, then, the labeling of this work (course number, course name, MIT OCW) is strictly for attribution purposes. Sadly, ***I am not affialited with MIT*** and this work should be treated as such. As for changes, I anticipate no changes other than occassionally omitting a problem I don't wish to complete. The solutions will of course be written by myself, so if that counts as a change, let it be known that I declared that change in advance.

To MIT: In the unlikely event (no pun intended) that I have violated the creative commons license, please contact me at `traviswantstolearn@gmail.com` and I will happily make the necessary measures to right my wrong.

CORRECTNESS STATEMENT

While there is always a chance of errors in any body of the work, everything I achieve here will be self taught—which means the risk of errors is perhaps higher than usual. If, for some reason, one finds themself on my page tempted to use my work, please first check that it is indeed correct. If you find an error, email me and I'll fix it :)

Note: The majority of the below problems are from *A First Course in Probability* 8th ed. by Sheldon Ross.

Chapter 4.

1. **Problem 70.** At time 0 a coin that comes up heads with probability p is flipped and falls to the ground. Suppose it lands on heads. At times chosen according to a Poisson process with rate λ, the coin is picked up and flipped. (Between these times the coin remains on the ground.) What is the probability that the coin is on its head side at time t?

   Hint: What would be the conditional probability if there were no additional flips by time t, and what would it be if there were additional flips by time t?

   Solution.

   Let's start with some definitions. Let N_t be the random variable obeying a Poisson process that denotes the number of flips between time 0 and time t. Since the coin is already heads, the conditional probability that the coin is heads up at time t given that there have been 0 flips is just

   $$P(N_t = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}.$$

   We now handle the case that $N_t \neq 0$. Now, it isn't clear to me that conditional probability is actually the way to go here because the probability that the coin lands on heads is independent of how many flips there are (because only the last flip matters). That is, $P(H|N_t = k) = P(H) = p$. However, this completely loses information about the number of flips, which...well, it seems silly. If we didn't care about flips, why ask about the Poisson prcoess at all?

   Notice, we're trying to find the probability that the coin is on heads in the end. Either there was one flip, or there was more than one. Both events are desireable, so we want their union. When there is more than one flip, there are arbitrarily many desirable outcomes (that is 2 flips and heads is desireable, as is 3 and heads, 4 and heads, and so on). Thus, the probability that the coin lands on heads *and* there was more than one flip is given by $P(H \cap N_t \geq 1)$. Since these events are independent, we have

   $$P(H \cap N_t \geq 1) = P(H)P(N_t \geq 1).$$

   Observe that $P(N_t \geq 1)$ is given by

   $$1 - P(N_t = 0) = 1 - e^{-\lambda t}.$$

   Therefore, the probability that the coin is heads up at time t (regardless of the number of flips) is

   $$P(N_t = 0) + P(H)P(N_t \geq 1) = e^{-\lambda t} + p(1 - e^{-\lambda t}).$$

   Remark. With some more pondering, I'm even more confident that conditional probability isn't the right way to go. Recall in chapter 3 (which ironically, is called "Conditional Probability and Indpendence") that we said that $P(E) = P(E \cap F) + P(E \cap F^c)$ where $F \cup F^c = \Omega$. In this case, $N_t = 0$ corresponds to F, $N_t \geq 1$ corresponds to F^c, and E corresponds to H. From there, our above result follows identically.

   ■

2. **Problem 84.** Suppose that 10 balls are put into 5 boxes, with each ball independently being put in box i with probability p_i where

$$\sum_{i=1}^{5} p_i = 1.$$

a) Find the expected number of boxes that do not have any balls.

b) Find the expected number of boxes that have exactly 1 ball.

Solution.

a) I think it's possible to solve this problem via cases and nested sums, but the math would be disgusting. I tried to make it work on scratch paper, and I quickly determined that I must be missing an easier method. Looking back on my notes lead me to the linearity of expectation (which is a good trick to keep in mind for any expectation problem, and admittedly, is the first thing I should have thought of). Recall that we can calculate the expecation of a binomial random variable by using indicator variables. We'll use the same trick here.

First, notice that for any given ball, the probability that that ball didn't enter box i is given by $(1-p_i)$. Therefore, the probability that all 10 balls didn't enter box i is $(1-p_i)^{10}$. We then have an indicator variable for each box

$$P(I_i = 1) = (1-p_i)^{10}$$

where $I_i = 1$ denotes that box i is empty.

We can then form a sum of indicator variables $S = \sum_i I_i$ to denote the number of empty boxes. Our goal, then, is to find $E[S]$, which, due to linearity of expecation, is just

$$E[X] = \sum_{i=1}^{5} E[I_i] = \sum_{i=1}^{5} (1-p_i)^{10}.$$

b) We can still use indicator variables and linearity of expectation. In this case, suppose that the first ball enters box i, which occurs with probability p_i. Then, all 9 remaining balls have to enter any of the other boxes, which occurs with probability $(1-p_i)^9$ due to complementation (is that even a word?) and indpendence. Finally, this is only 1 of 10 disjoint cases (1 case for each ball), so the indicator variable for box i is

$$P(I_i = 1) = 10p_i(1-p_i)^9$$

where $I_i = 1$ denotes that box i has exactly 1 ball and $I_i = 0$ indicates that box i has any other number of balls. We then consider all boxes by considering the sum S of indicator variables, and get

$$E[S] = \sum_{i=1}^{5} E[I_i] = \sum_{i=1}^{5} 10p_i(1-p_i)^9.$$

■

3. **Theoretical Exercise 16.** Let X be a Poisson random variable with parameter λ. Show that $P(X = i)$ increases monotonically and then decreases monotonically as i increases, reaching its maximum when i is the largest integer not exceeding λ.

   Solution.

   We start with the ratio of sucessive probabilities $P(X = i)$ and $P(X = i-1)$. We have

$$\begin{aligned}\frac{P(X = i-1)}{P(X = i)} &= \frac{\frac{\lambda^{i-1}e^{-\lambda}}{(i-1)!}}{\frac{\lambda^{i}e^{-\lambda}}{i!}} \\ &= \frac{i}{\lambda}.\end{aligned}$$

   When $i < \lambda$, the ratio above is less than 1, implying

$$P(X = i) > P(X = i-1).$$

   Conversely, when $i > \lambda$ the ratio is greater than 1, implying

$$P(X = i) < P(X = i-1).$$

   As for the maximum, we have just shown that the sequence is increasing until $i \geq \lambda$. Thus, the greatest value within the sequence must occur at the last possible i such that $i < \lambda$.

   ■

4. Omitted.

Chapter 5.

1. **Problem 8.** The lifetime in hours of an electronic tube is a random variable having a probability density function given by

$$f(x) = xe^{-x} \quad x \geq 0.$$

   Compute the expected lifetime of such a tube.

   Edit: I started this problem set as soon as I started reading chapter 5. At that time, I wasn't aware of the gamma distribution. Using the gamma distribution would have reduced this problem to a few lines, but oh well. Integration review is probably healthy anyway.

   Solution. Using the definition of expectation of a CRV, we have

$$\begin{aligned}E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx \\ &= \int_{0}^{\infty} x^2 e^{-x} dx\end{aligned}$$

where the change of bounds is justified by the fact that the integral will be zero for $x < 0$. To solve the integral, we use integration by parts. Recall that in integration by parts for definite integrals,, we have

$$\int_a^b u dv = uv\Big|_a^b - \int_a^b v du.$$

Source: *Calculus* by James Stewart [1]

In this problem, we let

$$u = x^2 \quad \Rightarrow \quad du = 2x dx \qquad\qquad v = -e^{-x} \quad \Rightarrow \quad dv = e^{-x} dx$$

so

$$\int_0^\infty x^2 e^{-x} dx = -e^{-x}x^2\Big|_0^\infty + 2\int_0^\infty xe^{-x} dx.$$

In order to evaluate the term outside the integral, we could apply L'Hopital twice, but I leave that to the reader. Ultimately, as $x \to \infty$, we have $-e^{-x}x^2 \to 0$. Hence, we have

$$\int_0^\infty x^2 e^{-x} dx = 2\int_0^\infty xe^{-x} dx.$$

To integrate the right side, we actually apply integration by parts again. Just to avoid mistakes, we write it all out:

$$u = x \quad \Rightarrow \quad du = dx \qquad\qquad v = -e^{-x} \quad \Rightarrow \quad dv = e^{-x} dx$$

so

$$\begin{aligned}\int_0^\infty x^2 e^{-x} dx &= 2\Big[-e^{-x}x\Big|_0^\infty + \int_0^\infty e^{-x} dx\Big] \\ &= -2\Big[e^{-x}\Big]_0^\infty \\ &= -2 \cdot 0 + 2 \cdot 1 = 2.\end{aligned}$$

We have now shown that $\boxed{\text{the expected lifetime is 2 hours.}}$

■

---

[1] The reader may (or may not) object to the use of Stewart's book. Admittedly, I'm not a fan of Stewart's book becuase those that use it seem to heavily emphasize formulas, not proofs. (Though there are exceptions; my high school calculus teacher used Stewart's book and he's one of the finest instructors I've ever had.) Anyway, yes, there are other more rigorous books out there, but at this moment, rigor wasn't what I needed or wanted. It's good to be flexible and use different tools for different problems.

2. **Problem 11.** A point is chosen at random on a line segment of length L. Interpret this statement, and find the probability that the ratio of the shorter to the longer segment is less than $1/4$.

   Solution. I'm not entirely sure what there is to interpret, as the wording is quite clear. I'll assume that we are to define a uniform random variable that fits the problem statement. In order for the random variable to have a total probability of 1, it's pdf must be

$$f_X(x) = \begin{cases} \frac{1}{L} & 0 \le x \le L \\ 0 & \text{otherwise.} \end{cases}$$

   As for the ratio, we first observe that if the ratio between the smaller section and the larger section is $1:4$, then the larger section is 4 times as long as the shorter section; that is, the line has been divided into fifths. If we make the short section even smaller, then the larger section becomes even larger, and the ratio decreases. There are two points at which we achieve a shorter section and a longer section with ratio $1/4$: $L/5$ and $4L/5$. Hence, if the point is somewhere before $L/5$ or somewhere after $4L/5$, we get the desired result. Written rigorously, we have

$$\begin{aligned} P(X < L/5 \cup X > 4L/5) &= P(X < L/5) + P(X > 4L/5) \\ &= \int_0^{L/5} \frac{1}{L} dx + \int_{4L/5}^{L} \frac{1}{L} dx \\ &= \frac{2}{5}. \end{aligned}$$

■

MIT Problems.

1. Compute the expectation of X^n where n is a positive integer and X is a uniform random variable on the interval $[0, 1]$.

   Using the definition of expectation of a function of random variable, we have

$$\begin{aligned} E[X^n] &= \int_{-\infty}^{\infty} x^n f_X(x) dx \\ &= \int_0^1 x^n dx \\ &= \frac{x^{n+1}}{n+1}\Big|_0^1 \\ &= \frac{1}{n+1}. \end{aligned}$$

2. How does the answer change if the random variable is instead taken to be uniform on $[0, L]$ for some constant L.

   The function $f_X(x)$ changes, but other than that, the work is the same:

$$\begin{aligned} E[X^n] &= \int_{-\infty}^{\infty} x^n f_X(x) dx \\ &= \int_0^L \frac{x^n}{L} dx \\ &= \frac{x^{n+1}}{L(n+1)}\Big|_0^1 \\ &= \frac{1}{L(n+1)}. \end{aligned}$$

■