Lecture 5 Worksheet

Chrysafis Vogiatzis

Every worksheet will work as follows.

1. You will be asked to form a group with other students in the class: you can make this as big or as small as you'd like, but groups of 4-5 work best.
2. Read through the worksheet, discussing any questions with the other members of your group.
   - You can call me at any time for help!
   - I will also be interrupting you for general guidance and announcements at random points during the class time.
3. Answer each question (preferably in the order provided) to the best of your knowledge.
4. While collaboration between students is highly encouraged and expected, each student has to submit their own version.
5. You will have 24 hours (see gradescope) to submit your work.

Activity 1: Basic discrete random variable questions

A movie magazine gives each movie they review a score from 0 to 4 stars – no movie is allowed to have a fractional number of stars though. Over the years, the magazine has observed that the movies get a score that is distributed as a discrete random variable X with probability mass function

$$p(x) = \begin{cases} \frac{2x+2}{c}, & x = 0, 1, 2, 3, 4 \\ 0, & \text{otherwise.} \end{cases}$$

Problem 1: Probability mass functions

What is the value of c for which $p(x)$ is a valid probability mass function?

Answer to Problem 1.

This can prove very useful. When we have an idea for the probability mass function of a *discrete* random variable, but we are missing part of it (e.g., missing a coefficient's value), then we can sum over all possible values of the pmf and it should equate to 1. This summation only holds for *discrete random variables*. More on what we anticipate for *continuous random variables* that are allowed to take any real value within a range in Lecture 7.

Problem 2: Constructing cumulative distribution functions

Using the same probability mass function $p(x)$ that you were given in Problem 1 (replace the value for c that you calculated), what is the cumulative distribution function (cdf) of discrete random variable X? We could write it as a summation! That is,

$$F(x) = \sum_{y \leq x} p(y).$$

Use the cdf to calculate the probability that a movie gets up to 2 stars (including 2 stars) out of 4.

Answer to Problem 2.

$P(X \leq 2) =$

Problem 3: Calculating probabilities

For the same distribution, calculate the following probabilities. [1]

Answer to Problem 3.

$P(1 \leq X \leq 3) =$

$P(1 < X \leq 3) =$

[1] For discrete random variables, we can calculate the probability of X being between a and b in many different ways:

- $P(a \leq X \leq b) = \sum_{x=a}^{b} p(x)$.
- $P(a \leq X \leq b) = P(a - 1 < X \leq b) = F(b) - F(a - 1)$.

Be careful with $\leq$ vs. $<$.

Activity 2: Binomial, geometric, or hypergeometric?

Problem 4

A foundry has received an order for **5 castings**, made from precious metals. Each casting produced is of high quality (and hence can be sold to the customer) with probability 0.97. All castings are produced independently. You decide to schedule 6 castings for production, knowing full well that the customer only wants 5. What is the probability you get more than or equal to 5 high quality castings? [2]

Answer to Problem 4.

[2] Let X be the number of high quality castings produced out of the 6 you tried to produce.

1. What is X distributed as?
2. What is $P(X \geq 5)$? Can we make the claim that

$$P(X \geq 5) = P(X = 5) + P(X = 6)?$$

Problem 5

The foundry from earlier has received the same order for **5 castings**. However, instead of producing new ones, they use a batch of older castings already produced. The batch contains 100 castings, 97 of whom are of high quality. They decide to pick 6 of them at random and give them to the customer. What is the probability that the customer gets more than or equal to 5 high quality castings in the sample of 6 they receive? [3]

Answer to Problem 5.

[3] Once again, let X be the number of high quality castings you get in the group of 6 castings you pick from the batch.

1. What is X distributed as now?
2. Which formula should we use to calculate $P(X = 5)$ and $P(X = 6)$?

Problem 6

5% of all bits (a signal of 0 or 1) transmitted are sent in error (a 0 is sent instead of a 1, or vice versa). The message stops transmitting when the first bit is transmitted in error. Let X be the length of the message. What is the probability that $X = 5$?

Answer to Problem 6.

Problem 7

For the length of the message from Problem 6, what is the probability that $X \leq 5$?

Answer to Problem 7.

Activity 3: Overbooking

An airline has found that 10% of people buying a first class ticket do not show up to travel on the day of their flight (and independent of one another). This is why the company has decided to **sell 32 first class tickets for a flight that contains 30 first class seats**. If more than 30 first class ticket holders show up, then the company has to pay a penalty.

Problem 8

What is the probability the company does not pay penalty for a flight?

Answer to Problem 8.

Problem 9

A flight just departed without having to pay a penalty. What is the probability the flight departed with at least one empty first class seat?

Answer to Problem 9.

By the way, congratulations for answering this! This was a question in last year's midterm exam!