Lecture 11 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.

Before we get started, please remember (it is crucial to differentiate these two notions): we refer to X and Y (upper case, by convention) as two random variables, whereas x and y (lower case) are values they take.

Activity 1: Jointly distributed discrete random variables

In the next four questions, we focus on a pair of *discrete* random variables X and Y, distributed with joint probability mass function

$$f_{XY}(x,y) = c \cdot \frac{x+1}{y}.$$

At this point, it is useful to remind ourselves that the probability mass function values *are actual probabilities*, since we are discussing about discrete random variables.

Problem 1: Joint probability mass functions

Assume that $X = \{1, 2, 3\}$ and $Y = \{10, 20\}$, that is X is allowed to take values 1, 2, or 3, and Y is allowed to be equal to either 10 or 20. What should c be in order for this to be a valid joint pmf? [1]

Answer to Problem 1.

[1] Remember the first axiom of probability mass functions for discrete random variables: summing over all possible values should give us 1. In math terms:

$$\sum_x \sum_y f_{XY}(x,y) = 1 \implies \ldots \implies c = \ldots$$

Problem 2: Table form

We may construct a table! Based on your answer in Problem 1, we may collect the different probability mass function values in tabular form. Fill the table below with the actual probabilities for each pair of values.

Answer to Problem 2.

		Y	
		10	20
X	1		
	2		
	3		

Verify once again that indeed the summation all of them is equal to 1, just to be sure.

Problem 3: Marginal probabilities

Where may we find the *marginal probabilities* for X or Y alone? In the table form we saw earlier, they are found by summing over a row or a column, depending on which one we are looking for! In this case, what is:

a) $P(X = 1)$?

b) $P(Y = 20)$?

Answer to Problem 3.

$P(X = 1) =$

$P(Y = 20) =$

Problem 4: Conditional probabilities

Where may we find the *conditional probabilities* for X given $Y = y$ or for Y given $X = x$? In the table form we saw earlier, they are found by focusing on one column or row, and then dividing the probability we are looking for over the summation of all elements in that column or row. [2] In our example, what is:

- $P(X = 1|Y = 20)$?
- $P(Y = 20|X = 1)$?

Answer to Problem 4.

$P(X = 1|Y = 20) =$

$P(Y = 20|X = 1) =$

Of course, this table form is valuable; but it is also limited to smaller sample spaces. What happens when we are dealing with a huge number of cases? In that case, we need to resort to the actual formulations for each and every one of our probability calculations. We'll see how that algebraic way of dealing with probabilities works in Activity 3 (in Page 6).

First, though, let's take a walk in the realm of continuous random variables in the next activity.

[2] Without a table, we would calculate a conditional probability by dividing appropriately. For example, the probability of getting $P(X = x|Y = y)$ could be found by $\frac{f_{XY}(x,y)}{f_Y(y)}$.

Activity 2: Jointly distributed continuous random variables

Let X and Y be two continuous random variables that are allowed to take any value between 0 and 1: that is, $0 \leq X \leq 1$ and $0 \leq Y \leq 1$. We further assume that they are jointly distributed with probability density function:

$$f_{XY}(x,y) = \frac{12}{11}\left(x^2 + y^2 + xy\right).$$

Problem 5: Calculating a probability

Recall that with continuous random variables, we need to integrate properly to calculate a probability. With that in mind, what is the probability that $0.3 \leq X \leq 0.7$ and $Y \geq 0.75$? [3]

Answer to Problem 5.

[3] You'll need to do a double integration. I'll get you started:

$$\int_{0.3}^{0.7}\int_{0.75}^{1} \cdots\cdots$$

Problem 6: Deriving a marginal distribution

Again, like you did earlier, derive the two marginal distributions. However, remember, that we are no longer summing! In the continuous space, we integrate. With that in mind, what is the marginal distribution of X? What is the marginal distribution of Y? [4]

Answer to Problem 6.

[4] In our case, because X and Y are only allowed to be between 0 and 1, we have (for Y, but it is very similar for X):

$$f_Y(y) = \int_0^1 f_{XY}(x,y)dx.$$

Problem 7: Deriving the conditional distribution

What is $P(X \leq 0.5|Y = 1)$? [5]

Answer to Problem 7.

[5] For the conditional distribution, apart from the fact we are integrating instead of summing, we follow exactly the same logic as for discrete random variables. Remember to divide appropriately!

Think about your approach. Could you have calculated the probability $P(X \leq 0.5|Y = y)$ for any value y?

Problem 8: Final touch

This is a little tougher. What is $P(X \leq Y)$? To help you get started, we have begun the solution approach, but do take a look at the hint for another explanation.. [6]

[6] Well, if we knew that $Y = y$, we could then calculate $P(X \leq Y|Y = y) = P(X \leq y)$, right? And, if we need to do that for any value that Y is allowed to take, can we use the **law of total probability for continuous random variables**? How does the law of total probability for continuous random variables look like again?

Answer to Problem 8.

From the law of total probability for continuous random variables, we have:

$$P(X \leq Y) = \int_0^1 P(X \leq Y|Y = y) f_Y(y) dy = \int_0^1 P(X \leq y) f_Y(y) dy = \ldots$$

Activity 3: Discrete, but infinite

Consider two discrete random variables X and Y that take on integer values $X \geq 0$ and $1 \leq Y \leq 5$. That is, X could be 3, 107, or 0, and Y could be equal to 1, 2, 3, 4, or 5. Their joint probability mass function is given by:

$$f_{XY}(x,y) = e^{-y} \cdot \frac{y^x}{5 \cdot x!}.$$

As a side note, remember that

$$\boxed{\sum_{i=0}^{\infty} e^{-\alpha} \cdot \frac{\alpha^i}{i!} = 1 \text{ for any } \alpha > 0.}$$

Problem 9: Using the joint pmf

Let's start easy. What is the probability that both X and Y are equal to 1? That is, what is $P(X = 1 \cap Y = 1)$? [7]

[7] This can also be written as $P(X = 1, Y = 1)$. Recall that (because this is a pmf for a discrete random variable) this can be found as simply the value of the pmf for the given X and Y.

Answer to Problem 9.

Problem 10: Deriving a marginal distribution

What is the probability that $Y = 1$, regardless of what X is? That is, what is $P(Y = 1)$? [8]

[8] Remember! To find the marginal distribution of one discrete random variable, sum over the other variable! In our case:

$$f_Y(y) = P(Y = y) = \sum_{x=0}^{\infty} f_{XY}(x,y).$$

Answer to Problem 10.

After answering this, what is the probability that $Y = y$, for any value y? Is it always $\frac{1}{5} = 20\%$?

Problem 11: Deriving a conditional distribution

What is the probability that $X \geq 1$ given that $Y = 1$? That is, what is $P(X \geq 1 | Y = 1)$? [9]

Answer to Problem 11.

[9] You will probably get something similar to what we had seen earlier in the semester...