MATH 323 - Tutorial 9 Solutions

Tyrel Stokes

October 2020

Email tyrel.stokes@mail.mcgill.ca for any questions about these problems.

I'm going to attempt something slightly differently with these solutions. In some cases, I will try and separate the barebones solutions and a more detailed explanation. The idea is there are many different use cases for solutions like these. Some students may be using them mostly to check their work, others may be trying to study the concepts in more detail. Hopefully this arrangement will make it easier to do both of those things. In cases where I do this, the barebones solution will be in a box. Bear in mind that what is in the box is not necessarily sufficient for full marks on an exam, where you might be expected to show your work, the work (or at least the keys steps) will be shown in the detailed explanation which follows.

Question 1 is slightly modified from what was covered in the actual recorded tutorial. Question 2 has additional parts.

1. Consider the following joint pmf, for the random variables X_1, X_2.

		X_1		
		0	1	2
	0	0.1	0.05	0,2
X_2	1	0	0.1	0.2
	2	0.3	0.05	0

a) Find the marginal support (values x_j for which $P(X_i = x_j) > 0$ for $i = 1, 2$) of each of the random variables X_1, X_2. Find the marginal distributions for X_1 and X_2.

Solution:

$$P(X_1 = x_1) = \begin{cases} 0.4, & x_1 = 0 \\ 0.2, & x_1 = 1 \\ 0.4, & x_1 = 2 \end{cases}$$

$$P(X_2 = x_2) = \begin{cases} 0.35, & x_2 = 0 \\ 0.3, & x_2 = 1 \\ 0.35, & x_2 = 2 \end{cases}$$

To find the marginal supports we simply need to scan the rows and columns. Start with X_1 and thus the columns. If the column has at least one entry in it with positive probability, it much have marginal support on that value. Thus:

Support of $X_1 = \{0, 1, 2\}$
Support of $X_2 = \{0, 1, 2\}$

Similarly, to find the marginal distributions we simply need to sum up the columns and rows. This may be a fact known to you, but let's try to connect why we are able to do this to things we studied earlier in the course. For example

$$\begin{aligned} P(X_1 = 1) &= P(\{(X_1 = 1) \cap (X_2 = 0)\} \cup \{(X_1 = 1) \cap (X_2 = 1)\} \cup \{(X_1 = 1) \cap (X_2 = 2)\}) \\ &\overset{disjoint}{=} P(\{(X_1 = 1) \cap (X_2 = 0)\}) + P(\{(X_1 = 1) \cap (X_2 = 1)\}) + P(\{(X_1 = 1) \cap (X_2 = 2)\}) \\ &= 0.05 + 0.1 + 0.05 \\ &= 0.2 \end{aligned}$$

Above we used the fact that all the values that a random variable can take (i.e the support) form a partition to break $P(X_1 = 1)$ into 3 disjoint parts which we already have information about.

Notice that we get back simply the entries in the second column. In fancier mathematical terms we can get back the marginal distribution by summing over the joint pmf with respect to the other variables, i.e $P(X_1 = x_1) = \sum_{x_2 \in \mathcal{X}_\in} P(X_1 = x_1, X_2 = x_2)$, where $x_2 \in \mathcal{X}_\in$ are the values that X_2 takes with positive probability (the support). This follows directly from the kinds of steps we did above to find $P(X_1 = x_1)$.

I'm not just needlessly trying to make what seems like a simple question harder, the goal is to set up question 2. In question 2 we are dealing with a slightly funky joint pdf instead of a simple pmf. Students, in my experience, often have good intuition about joint pmfs, but much less so about pdfs. The goal is to try and make it clear how much the two share in common and to hopefully allow your intuition to jump from pmfs to pdfs. Notice in particular that summing over and integrating over the other variables (which we do in question 2) are accomplishing the same thing.

b) Fjnd the marginal variance of X_1 and X_2.

Solution:

$$\begin{aligned}
E[X_1] &= \sum_{x_1 \in \mathcal{X}_\infty} x_1 P(X_1 = x_1) \\
&= 0 * 0.4 + 1 * 0.2 + 2 * 0.4 \\
&= 1 \\
E[X_2] &= \sum_{x_2 \in \mathcal{X}_\in} x_2 P(X_2 = x_2) \\
&= 0 * 0.35 + 1 * 0.3 + 2 * 0.35 \\
&= 1
\end{aligned}$$

$$\begin{aligned}
Var(X_1) &= \sum_{x_1 \in \mathcal{X}_\infty} (x_1 - E[X_1])^2 P(X_1 = x_1) \\
&= 0.4 * (0-1)^2 + 0.2 * (1-1)^2 + 0.4(2-1)^2 \\
&= 0.4 + 0.4 \\
&= 0.8 \\
Var(X_2) &= \sum_{x_2 \in \mathcal{X}_\in} (x_2 - E[X_2])^2 P(X_2 = x_2) \\
&= 0.35 * (0-1)^2 + 0.3 * (1-1)^2 + 0.35 * (2-1)^2 \\
&= 0.35 + 0.35 \\
&= 0.7
\end{aligned}$$

c) Find the covariance between X_1 and X_2.

Solution:

$$Cov(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2]$$

From part b, we have that $E[X_1]E[X_2] = 1 * 1 = 1$, so it remains only to find $E[X_1X_2]$.

$$\begin{aligned}
E[X_1X_2] &= \sum_{x_1 \in \mathcal{X}_\infty} \sum_{x_2 \in \mathcal{X}_\in} x_1x_2P(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1=0}^{2} \sum_{x_2=0}^{2} x_1x_2P(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1=1}^{2} \sum_{x_2=1}^{2} x_1x_2P(X_1 = x_1, X_2 = x_2) \\
&= \sum_{x_1=1}^{2} (x_1 * 1 * P(X_1 = x_1, X_2 = 1) + x_1 * 2 * P(X_1 = x_1, X_2 = 1) \\
&= (1 * 1 * P(X_1 = 1, X_2 = 1) + 1 * 2 * P(X_1 = 1, X_2 = 1) + (2 * 1 * P(X_1 = 2, X_2 = 1) + 2 * 2 * P(X_1 = 2, X_2 \\
&= (1 * 1 * (0.1) + 1 * 2 * (0.05) + 2 * 1 * (0.2) + 2 * 2 * (0) \\
&= 0.1 + 0.1 + 0.4 + 0 \\
&= 0.6
\end{aligned}$$

Where the third line follows from noticing that whenever either $x_1 = 0$ or $x_2 = 0$ the summand will be zero. This is just to save us from having to work out a larger sum.

$$\begin{aligned}
Cov(X_1, X_2) &= E[X_1X_2] - EX_1E[X_2] \\
&= 0.6 - 1 \\
&= -0.4
\end{aligned}$$

This tells us there is a negative linear relationship between X_1 and X_2. Look again at the joint pmf and see if you notice this pattern there directly.

d) Using the information in the previous parts, find the correlation between X_1 and X_2.

Solution:

$$\begin{aligned} Cor(X_1, X_2) = \rho(X_1, X_2) &= \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)}\sqrt{Var(X_2)}} \\ &= \frac{-0.4}{\sqrt{0.8}\sqrt{0.7}} \\ &= -0.53452248382 \\ &\approx -0.53 \end{aligned}$$

Remember that this quantity must be between -1 and 1. This can help weed out some numerical errors by checking that your solution adheres to this bound.

e) Are X_1 and X_2 independent, justify your answer using information from the previous questions.

Solution:

There are many ways that we could see they are in fact dependent. First notice that $Cov(X_1, X_2) \neq 0$ which implies dependency. Remember that the converse is NOT TRUE, i.e showing that the covariance is 0 is not enough to prove independence except in a few very special cases. The same holds for correlation (notice that correlation is just a scaling of the covariance. If correlation is 0 then so is covariance and vice versa).

A few other things we could notice that would also show dependence. Notice that the conditional support of the variables depends on each other. For example, knowing that $X_2 = 2$ means that it is impossible for $X_2 = 2$. Another way to say this is that the support depends on each other. Similar to covariance, showing the support is depedent is enough for dependence, but the converse showing the supports do not depend on each other is not enough to prove independence.

In this same line of thinking we could show that $P(X_1 = x_1, X_2 = x_2) \neq P(X_1 = x_1)P(X_2 = x_2)$ for at least one set of constants. For example $P(X_1 = 1, X_2 = 1) = 0.1 \neq P(X_1 = 1)P(X_2 = 1) = 0.2 * 0.3 = 0.06$.

Any of these answers are sufficient to show dependence and thus answer the question. Support dependence was not mentioned in class to my knowledge so I would focus on the correlation/covariance or joint and marginal routes to answer this question for this course.

2. Consider the following joint density function $f_{Y_1,Y_2}(y_1, y_2) = c, 0 < y_1 < 1, y_1 < y_2 < y_2 + 1$, where c is some constant. (This question has a non-rectangular domain of integration and thus a trickier question than what is likely to show up on the final. I would suggest being comfortable with rectangular domains of integration first. For example, there is a question in my tutorial 10 like this. I believe that Professor Wolfson has also posted one. If those feel easy then this is a great challenge question to solidify your understanding of the material)

a) Find the constant c which makes the above a valid joint density function.

Solution:

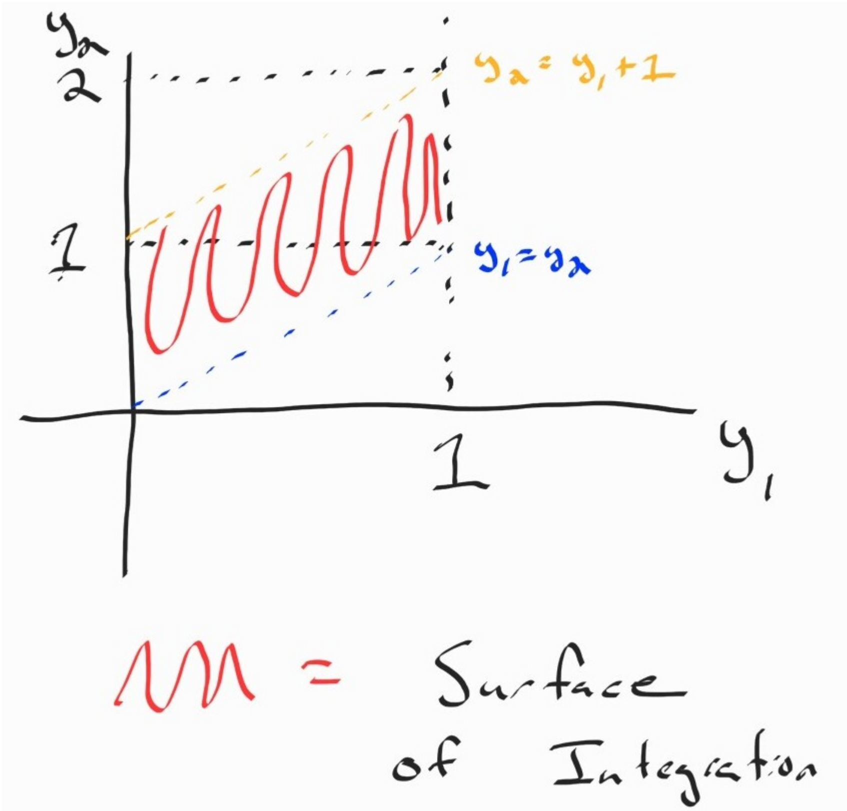
$$\boxed{\text{c} = 1}$$

To find this solution, we exploit the fact that any valid joint density must integrate to 1.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2) dy_1 dy_2 \overset{\text{set to}}{=} 1$$

The tricky part of this is that the joint support is not a nice rectangle, so we have to remember a little bit from calc 2 or 3. Below I have drawn out the surface of integration. The trick is to turn the inequalities into equalities and to draw the lines.

For example we can re-express $y_1 < y_2 < y_2 + 1$ into two separate inequalities, $y_1 < y_2$ and $y_2 < y_1 + 1$. Now we can draw the line $y_1 = y_2$ (45 degree line) and $y_2 = y_1 + 1$ (45 degree line shift 1 up from the origin.



Notice that there are two ways to integrate over this, we can integrate over y_1 or y_2 first (by fubini's theorem). In this case, our life is much easier if we choose to integrate over y_2 first since the bounds of integration do not change over the entire integral.

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)dy_1dy_2 &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)dy_2dy_1 \\
&= \int_0^1 [\int_{y_1}^{y_1+1} cdy_2]dy_1 \\
&= \int_0^1 [cy_2]_{y_1}^{y_1+1} d_y 1 \\
&= \int_0^1 c(y_1+1-y_1)dy_1 \\
&= \int_0^1 c * 1dy_1 \\
&= [cy_1]_0^1 \\
&= c
\end{aligned}$$

Thus $c = 1$

b) Find the marginal pdfs for Y_1 and Y_2 (Hint: If you did step a correctly you already showed the work for one of these two cases.)

Solution:

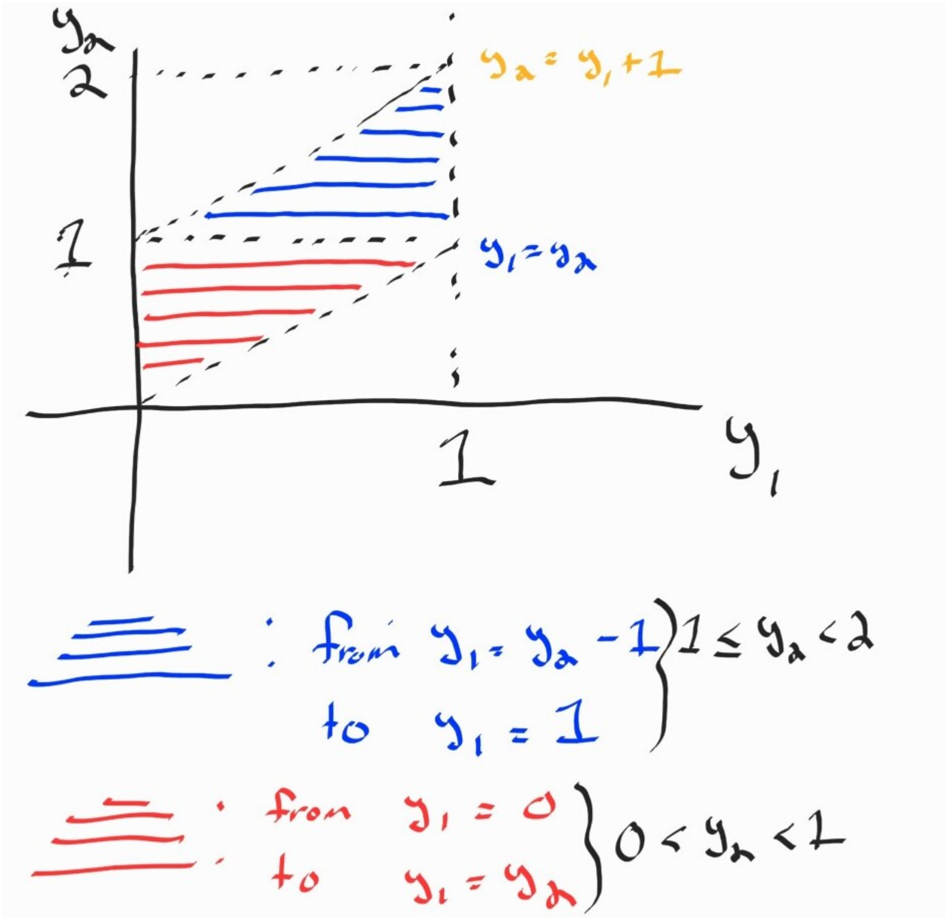
Notice that in part a, we integrated the pdf over dy_2 first, which is exactly what we need to do to get the marginal distribution of y_1.

$$\begin{aligned} f_{Y_1}(y_1) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1, y_2) dy_2 \\ &= \int_{y_1}^{y_1+1} c dy_2 \\ &= 1 \end{aligned}$$

We solved the integration, but remember we have to specify where the function is valid. In this case, the support comes directly from the joint. Thus:

$\boxed{\text{f}_{Y_1}(y_1) = 1, 0 < y_1 < 1}$

For this course, you should be able to get the marginal to something like $f_{Y_1}(y_1)$. $f_{Y_2}(y_2)$ is a slightly more difficult question requiring that you remember a little bit more from your calc classes. Below I have split the surface of integration into two parts.



Using the above image we get the following

$$
\begin{aligned}
f_{Y_2}(y_2) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y_1,y_2)dy_1 \\
&= \begin{cases} \int_0^{y_2} f_{Y_1,Y_2}(y_1,y_2)dy_1, & 0 < y_2 < 1 \\ \int_{y_2-1}^{1} f_{Y_1,Y_2}(y_1,y_2)dy_1, & 1 \leq y_2 < 2 \end{cases} \\
&= \begin{cases} \int_0^{y_2} 1dy_1, & 0 < y_2 < 1 \\ \int_{y_2-1}^{1} 1dy_1, & 1 \leq y_2 < 2 \end{cases} \\
&= \begin{cases} y_2, & 0 < y_2 < 1 \\ 2 - y_2, & 1 \leq y_2 < 2 \end{cases}
\end{aligned}
$$

Try to connect the integration process to summing across the rows and columns we did in question 1. A joint density of two variables lives in 3 dimensions. We have already visualized the two dimensional surface of integration. The third dimension, the height is the value of the joint density. In this case, the joint density has a constant value, which is just 1 over the entire surface of integration. In this case, the whole joint density looks like a loaf of bread. The value of the marginal density at some specific value $y_2^\star$ ($f_{Y_2}(y_2^\star)$) can be thought of as the size of the infinitely think slice of bread at $y_2 = y_2^\star$. To figure this out, we break the infinitely thin slice into infinitely thin strips and "sum" up the size of each of these strips. This is a bit of a simplification and makes the most sense from the reimannian integral perspective (if you go on in probability, stats or analysis, you will see the Lebegue integral is the more natural in these contexts), but can hopefully help you relate the process of integrating over the joint to summing over the joint in the discrete case.

One important caveat is that it is important to remember that joint densities do not have a probability interpretation like joint mass functions. The rest of the mechanics are quite similar.

c) Find the covariance between Y_1 and Y_2.

Solution:

$$\begin{aligned}
E[Y_1] &= \int_{-\infty}^{\infty} y_1 f_{Y_1}(y_1) dy_1 \\
&= \int_0^1 y_1 dy_1 \\
&= [\frac{y_1^2}{2}]_0^1 \\
&= \frac{1}{2} \\
E[Y_2] &= \int_{-\infty}^{\infty} y_2 f_{Y_2}(y_2) dy_2 \\
&= \int_0^1 y_2^2 dy_2 + \int_1^2 y_2(2-y_2) dy_2 \\
&= [\frac{y_2^3}{3}]_0^1 + [y_2^2 - \frac{y_2^3}{3}]_1^2 \\
&= \frac{1}{3} + (3 - \frac{8}{3}) + (1 - \frac{1}{3}) \\
&= \frac{4}{3}
\end{aligned}$$

$$
\begin{aligned}
E[Y_1Y_2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_1y_2f_{Y_1,Y_2}(y_1,y_2)dy_2dy_1 \\
&= \int_0^1\int_{y_1}^{y_1+1} y_1y_2dy_2dy_1 \\
&= \int_0^1 [y_1\frac{y_2^2}{2}]_{y_1}^{y_1+1}dy_1 \\
&= \int_0^1 [\frac{(y_1(y_1+1)^2)}{2} - \frac{y_1^3}{2}]dy_1 \\
&= \int_0^1 \frac{(y_1(y_1^2+2y_1+1)}{2} - \frac{y_1^3}{2})dy_1 \\
&= \int_0^1 \frac{(y_1^3+2y_1^2+y_1)}{2} - \frac{y_1^3}{2}dy_1 \\
&= \int_0^1 (y_1^2+\frac{y_1}{2})dy_1 \\
&= [\frac{y_1^3}{3}+y_1^2]_0^1 \\
&= \frac{4}{3}
\end{aligned}
$$

Once again we chose to integrate over y_2 first so that we can avoid splitting the integral into two.

$$
\boxed{
\begin{aligned}
Cov(X_1,X_2) &= E[X_1X_2] - E[X_1]E[X_2] \\
&= \frac{4}{3} - \frac{4}{3}\frac{1}{2} \\
&= \frac{4}{3}(1-\frac{1}{2}) \\
&= \frac{2}{3}
\end{aligned}
}
$$