Problem Sheet 9 [1]

Based on Chapters 7.1 - 7.4 of the course book.

Optional Revision Problems

Exercise 1. Let $X \sim Expo(\lambda)$. Find the median and mode of X.

Solution 1. By definition, the median for a continuous random variable X is the value c, for which $F_X(c) = P(X \leq c) = 1/2$ holds. The CDF of $X \sim Expo(\lambda)$ on its support is $F_X(x) = 1 - e^{-\lambda x}$, thus the median is

$$\begin{aligned} & 1 - e^{-\lambda c} = 1/2 \\ \iff & 1/2 = e^{-\lambda c} \\ \iff & \log(1/2) = -\lambda c \\ \iff & \frac{\log(2)}{\lambda} = c. \end{aligned}$$

(Same as Exercise S8E7 part 1.)

The mode of a distribution is defined as the point at which the probability density function (PDF) achieves its maximum value. The general approach to finding the maximizer of a function f involves taking its first derivative, setting it equal to zero, and then verifying that these critical points are indeed maxima by checking the second derivative at each solution. For the sake of completeness, we can proceed with this approach. The PDF of $X \sim Expo(\lambda)$ is $f_X(x) = \lambda e^{-\lambda x}$, so $f_X'(x) = -\lambda^2 e^{-\lambda x}$. On the support of X (i.e. $[0, \infty)$), this function is non-zero everywhere, and in fact, negative. Therefore, the maximizer of $f_X(x)$ is at the left boundary point of its support, which implies that the mode is equal to 0.

An equally good solution is just noticing that the PDF is strictly decreasing in x, without doing the integral, and then arguing that it is maximized at the left boundary point, at 0.

Exercise 2. Let $X \sim Bin(n, p)$.

1. For $n = 5$, $p = 1/3$, find all medians and all modes of X. How do they compare to the mean?
2. For $n = 6$, $p = 1/3$, find all medians and all modes of X. How do they compare to the mean?

Solution 2. Contrary to last week and exercise 1, **we cannot use the simplification $\mathbf{P(X \leq c) = 1/2}$** for finding the median, as the Binomial is discrete. Rather, the more general definition given in the lecture should be used: The median of X is the value c for which $P(X \leq c) \geq 1/2$

[1]Exercises are based on the coursebook Statistics 110: Probability by Joe Blitzstein

and $P(X \geq c) \geq 1/2$ simultaneously. Note that for a random variable X whose support is a subset of the integers, we have

$$P(X \geq k) = 1 - P(X < k) = 1 - P(X \leq (k-1)),$$

for $k \in \mathbb{N}$. Thus we can reformulate the condition of the median as follows: The median of a Binomial is the integer k for which $P(X \leq k) \geq 1/2$ and $P(X \leq (k-1)) \leq 1/2$ holds simultaneously.

1. For finding either the median or the mode, first we have to find the PMF of X. That is defined generally as

$$P(X = k) = \binom{5}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{5-k},$$

for $k \in \{0, \ldots 5\}$, so substituting in for the different k-s, we get

$$\begin{aligned}
P(X = 0) &= \binom{5}{0} \left(\frac{1}{3}\right)^0 \left(\frac{2}{3}\right)^5 = 1 \cdot 1 \cdot \frac{32}{243} = \frac{32}{243} \\
P(X = 1) &= \binom{5}{1} \left(\frac{1}{3}\right)^1 \left(\frac{2}{3}\right)^4 = 5 \cdot \frac{1}{3} \cdot \frac{16}{81} = \frac{80}{243} \\
P(X = 2) &= \binom{5}{2} \left(\frac{1}{3}\right)^2 \left(\frac{2}{3}\right)^3 = 10 \cdot \frac{1}{9} \cdot \frac{8}{27} = \frac{80}{243} \\
P(X = 3) &= \binom{5}{3} \left(\frac{1}{3}\right)^3 \left(\frac{2}{3}\right)^2 = 10 \cdot \frac{1}{27} \cdot \frac{4}{9} = \frac{40}{243} \\
P(X = 4) &= \binom{5}{4} \left(\frac{1}{3}\right)^4 \left(\frac{2}{3}\right)^1 = 5 \cdot \frac{1}{81} \cdot \frac{2}{3} = \frac{10}{243} \\
P(X = 5) &= \binom{5}{5} \left(\frac{1}{3}\right)^5 \left(\frac{2}{3}\right)^0 = 1 \cdot \frac{1}{243} \cdot 1 = \frac{1}{243}
\end{aligned}$$

, therefore the PMF is maximized at $k = 1$ and $k = 2$, so the modes are 1 and 2.

For finding the median we need to check conditions related to the CDF of the random variable. Since between integers, the CDF of X is constant, it is sufficient to look at CDF only at $k \in \{0, \ldots 5\}$.

$$\begin{aligned}
&F_X(0) = P(X \leq 0) = P(X = 0) = \frac{32}{243} \approx 0.132 \\
&F_X(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{32 + 80}{243} = \frac{112}{243} \approx 0.461 \\
&F_X(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{32 + 80 + 80}{243} = \frac{192}{243} \approx 0.79,
\end{aligned}$$

therefore, for $P(X \leq 2) \geq 1/2$ holds.

Applying to our example the reformulation of the condition for the median stated at the top, we need to check whether $P(X \leq (2-1)) \leq 1/2$ holds, and as it does, it means that 2 is *a* median. Since $P(X \leq (3-1)) \nleq 1/2$, and the CDF is increasing in k, we can conclude that there are no other medians.

The mean of a $Binomial(n, p)$ distribution is, by definition, np. In our case, this is $5/3$. Therefore, the median is greater than the mean, while one of the modes is greater than the mean and the other is less than the mean.

2. By identical reasoning, we first need to find the PMF based on

$$P(X = k) = \binom{6}{k}\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{6-k},$$

for $k \in \{0, \ldots 6\}$. That is

$$\begin{aligned}
P(X = 0) &= \binom{6}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^6 = 1 \cdot 1 \cdot \frac{64}{729} = \frac{64}{729} \\
P(X = 1) &= \binom{6}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^5 = 6 \cdot \frac{1}{3} \cdot \frac{32}{729} = \frac{192}{729} \\
P(X = 2) &= \binom{6}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^4 = 15 \cdot \frac{1}{9} \cdot \frac{16}{81} = \frac{240}{729} \\
P(X = 3) &= \binom{6}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^3 = 20 \cdot \frac{1}{27} \cdot \frac{8}{27} = \frac{160}{729} \\
P(X = 4) &= \binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2 = 15 \cdot \frac{1}{81} \cdot \frac{4}{9} = \frac{60}{729} \\
P(X = 5) &= \binom{6}{5}\left(\frac{1}{3}\right)^5\left(\frac{2}{3}\right)^1 = 6 \cdot \frac{1}{729} \cdot \frac{2}{3} = \frac{12}{729} \\
P(X = 6) &= \binom{6}{6}\left(\frac{1}{3}\right)^6\left(\frac{2}{3}\right)^0 = 1 \cdot \frac{1}{729} \cdot 1 = \frac{1}{729}.
\end{aligned}$$

Therefore the PMF is maximized at 2, so the single mode is 2.

The CDF at $k \in \{0, \ldots 6\}$, is

$$\begin{aligned}
F_X(0) &= P(X \leq 0) = P(X = 0) = \frac{64}{729} \approx 0.088 \\
F_X(1) &= P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{64 + 192}{729} = \frac{256}{729} \approx 0.351 \\
F_X(2) &= P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = \frac{32 + 192 + 240}{729} = \frac{496}{729} \approx 0.68,
\end{aligned}$$

therefore as $P(X \leq 2) \geq 1/2$, and $P(X \leq (2-1)) \leq 1/2$, we can conclude that 2 is the single median of this distribution.

By definition, the mean is $6 \cdot \frac{1}{3} = 2$, so the median, the mode and the mean are all equal to 2.

Week 9 Exercises

Exercise 3. Alice, Bob, and Carl arrange to meet for lunch on a certain day. They arrive independently at uniformly distributed times between 1 pm and 1:30 pm on that day.

1. What is the probability that Carl arrives first?

   For the rest of this problem, assume that Carl arrives first at 1:10 pm, and condition on this fact.

2. What is the probability that Carl will be waiting alone for more than 10 minutes?
3. What is the probability that Carl will have to wait more than 10 minutes until his party is complete?
4. What is the probability that the person who arrives second will have to wait more than 5 minutes for the third person to show up?

Solution 3.

Solution 1 Except for part 1., you might find the geometric arguments more straightforward. Since all variables are Uniform, the different probabilities coincide with the proportion of the area corresponding to the event of interest, divided by the total area of the sample space. (Since in part 1. we have three random variables, we should be talking about volumes, instead of areas.) Try to convince yourself that the figures below show the events of interest, if in doubt, check Solution 2, for a more mathematical derivation.

1. Figure 1

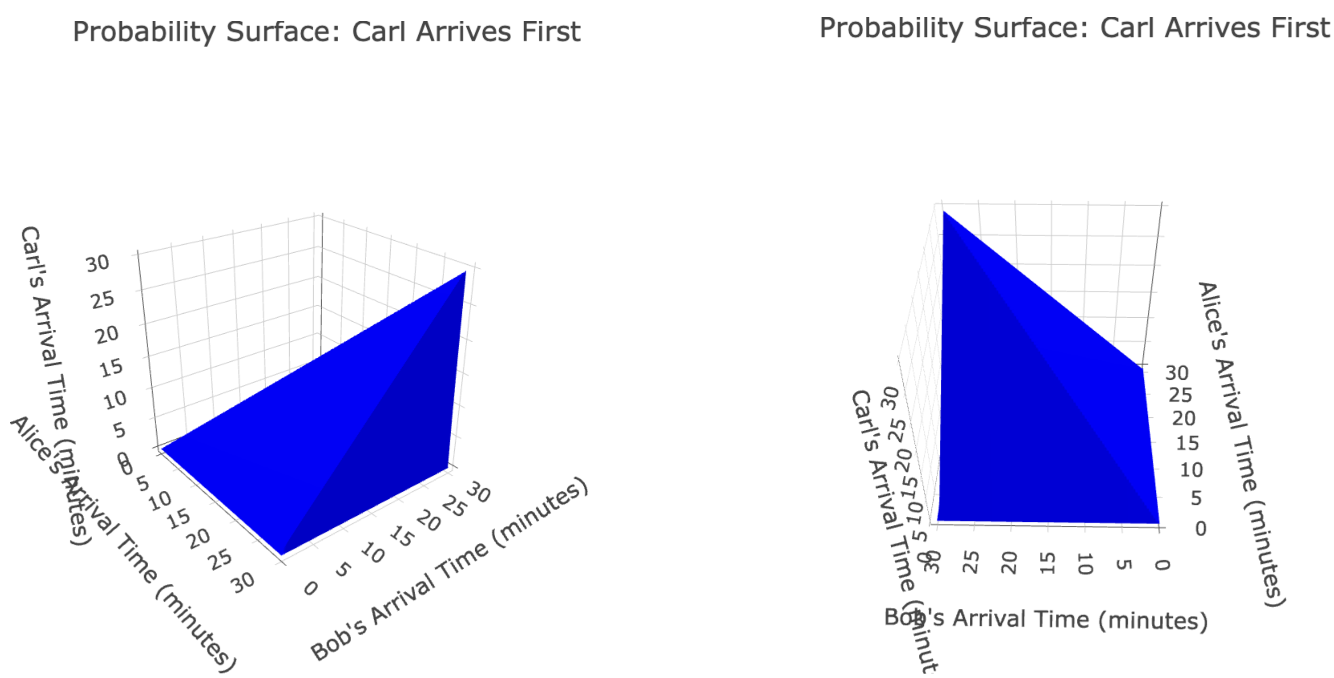

Figure 1: 3D visualization of Carl arriving first. The two planes correspond to Carl and Alice arriving at the same time, and Carl and Bob arriving at the same time. These two planes are intersected with each other, as Carl should arrive before both of them. The volume of interest is everything below both of these planes (as he should arrive before), and that corresponds to the part of the sample space where Carl arrives first. The two figures are showing the same object from different angles.

2. Figure 2

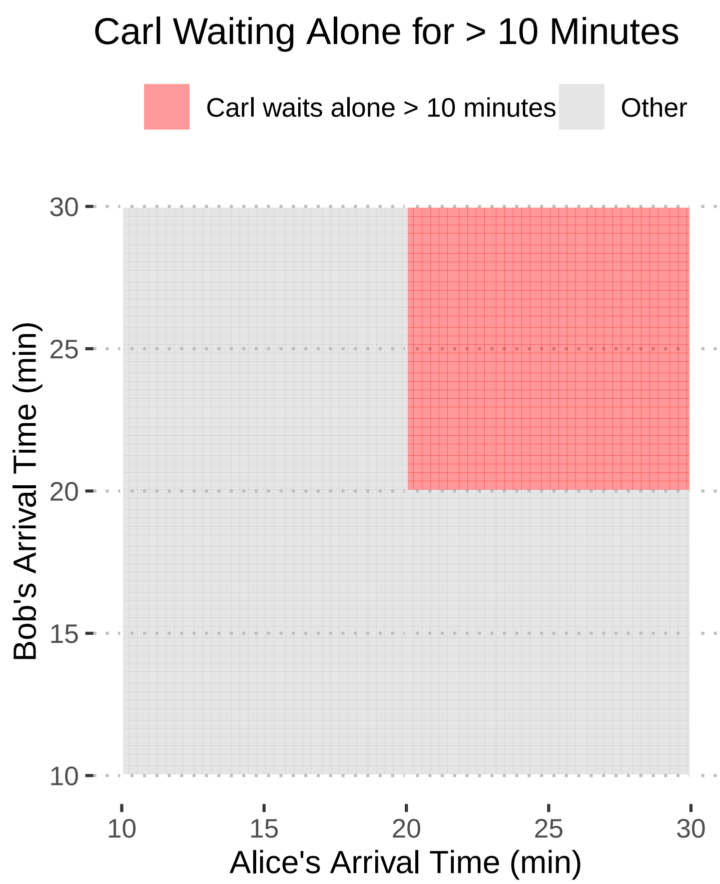

Figure 2: Carl waiting alone for more than 10 minutes. Conditional on that he arrived first at 1:10 pm, this just means that both Alice and Bob arrived after 1:20 pm, so this is the intersection of Alice arriving after 1:20 pm, with Bob arriving after 1:20 pm. As we condition on Carl arriving first at 1:10 pm, the sample space reduces to 2D, and also the arrival time of Alice and Bob must be between 1:10 and 1:30 pm.

3. Figure 3

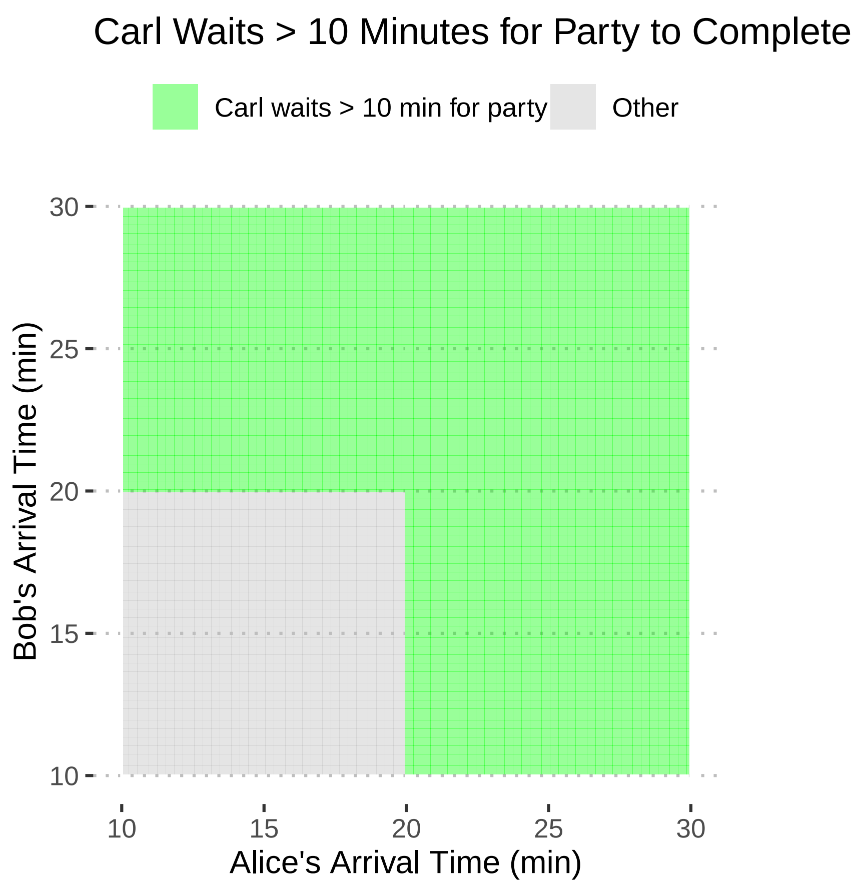

Figure 3: Carl waiting more than 10 minutes for his party to be complete. This means that either Alice arrives after 1:20 pm or Bob arrives after 1:20 pm, or both of them, i.e. the union of Alice arriving after 1:20 pm and Bob arriving after 1:20 pm. The sample space is reduced analogously by conditioning.

4. Figure 4

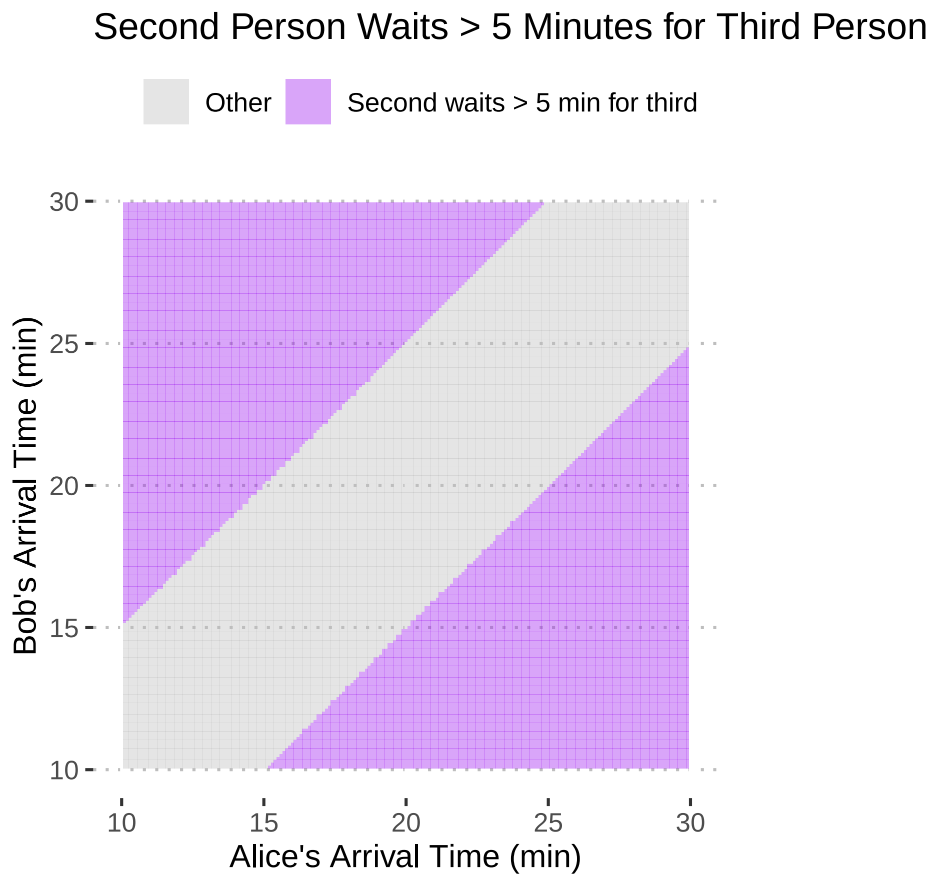

Figure 4: The second person waiting more than 5 minutes. This is everything in the sample space, except for the diagonal $A = B$, when Alice and Bob arrive at the same time, and a 5 minute margin added to both sides, as both Alice and Bob can be the second or third person respectively. The sample space is reduced analogously by conditioning.

Solution 2 Denote the time elapsed between 1 pm and the arrival of Alice, Bob, and Carl with A, B, and C, respectively. For example, the event that Alice arrived between 1:10 pm and 1:25 pm would be denoted as $\{10 < A < 25\}$. Choosing minutes as our scale, we can define these r.v.-s as

$$A, B, C \overset{\text{iid}}{\sim} Unif(0, 30).$$

1. (In this context, we mean by the symmetry of the random variables A, B, and C that they are independent and identically distributed). The event of interest is $\{(C < A) \cap (C < B)\}$. Using the symmetry of the random variables, we have that $P((C < A) \cap (C < B)) = P((B < A) \cap (B < C)) = P((A < B) \cap (A < C))$, i.e. each of the persons have the same probability arriving last. Finally, observe that *someone* has to arrive last, so $P((C < A) \cap (C < B)) + P((B < A) \cap (B < C)) + P((A < B) \cap (A < C)) = 1$, so $P((C < A) \cap (C < B)) = 1/3$.

   Alternatively, you can solve the integral corresponding to this probability (and to the volume in Figure 1), using that they are independent random variables, hence $f_{A,B,C}(a, b, c) = f_A(a)f_B(b)f_C(c)$:

$$\begin{aligned} P((C < A) \cap (C < B)) &= \int_0^{30} \int_c^{30} \int_c^{30} f_A(a)f_B(b)f_C(c)\,\mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c \\ &= \left(\frac{1}{30}\right)^3 \int_0^{30} \int_c^{30} \int_c^{30} \mathrm{d}a\,\mathrm{d}b\,\mathrm{d}c, \end{aligned}$$

   but for the sake of brevity this is left as a practice exercise.

   Important to point out, that even though A, B, C are iid, the events $A > C$ and $B > C$ are not independent, so $P((C < A) \cap (C < B)) \neq P(C < A)P(C < B)$.

2. The event that we condition on has two pieces of information: On the one hand, Carl arrived at 1:10 pm, and on the other, Alice and Bob arrived after Carl. Rephrasing this, this is just equal to conditioning on that Carl arrived at 1:10 pm, and both Alice and Bob arrived after 1:10. Using our notation $\{(A > 10) \cap (B > 10) \cap (C = 10)\}$. Even though this is a probability zero event (as these are continuous r.v.-s), remember from the lecture (, or from technical point 7.1.17. from the book), that we are technically conditioning on the event $\{(A > 10) \cap (B > 10) \cap (C \text{ is "close" to } 10)\}$, so the conditional probabilities will be well defined. In this sense, you can think about $P(C \text{ is "close" to } 10)$, as $P(C \in [10 - \epsilon, 10 + \epsilon]) = \int_{10-\epsilon}^{10+\epsilon} f_C(c)\mathrm{d}c$, and for the joint probabilities accordingly. However, for notational ease, I will simply write $P(C = 10)$ in the following.

   Carl will be alone for more than 10 minutes if both Alice and Bob arrive after 1:20, so the probability of interest is $P((A > 20) \cap (B > 20) | (A > 10) \cap (B > 10) \cap (C = 10))$.

Then

$$
\begin{aligned}
&P((A>20)\cap(B>20)|(A>10)\cap(B>10)\cap(C=10))\\
=&\frac{P((A>20)\cap(B>20)\cap(A>10)\cap(B>10)\cap(C=10))}{P((A>10)\cap(B>10)\cap(C=10))}\\
=&\frac{P((A>20)\cap(B>20))\cap(C=10))}{P((A>10)\cap(B>10)\cap(C=10))}\\
=&\frac{P(A>20)P(B>20)P(C=10)}{P(A>10)P(B>10)P(C=10)}\\
=&\frac{P(A>20)P(B>20)}{P(A>10)P(B>10)}\\
=&\frac{(1-P(A\le 20))(1-P(B\le 20))}{(1-P(A\le 10))(1-P(B\le 10))}\\
=&\frac{(1-20/30)(1-20/30)}{(1-10/30)(1-10/30)}=\frac{1}{4},
\end{aligned}
$$

where in the first step we used the definition of conditional probability, in the second step we simplified as one event was implied by the other, in the third step we used independence, in the fourth we simplified by $P(C=10)$ (that can be thought of as a small ϵ, not 0), and finally we rewrote in a form that involves the CDF of $Uniform(0,30)$, and calculated the result.

There are multiple other solutions to this exercise (and to the remaining two, as well), using visual/geometrical arguments based on the uniform law of the random variables, or via integration of the PDF-s on the correct sub-intervals defined by the events. If you used any other method, but got the right solution, probably your method is correct. If in doubt, reach out on the forum.

3. Carl has to wait more than 10 minutes until everyone is there, if *either* Alice *or* Bob arrives after 10:20 pm. In our notation, we want to get the probability $P((A>20)\cup(B>20)|(A>10)\cap(B>10)\cap(C=10))$. By the De Morgan's laws this is equal to

$$
\begin{aligned}
&P((A>20)\cup(B>20)|(A>10)\cap(B>10)\cap(C=10))\\
=&1-P(((A>20)\cup(B>20))^c|(A>10)\cap(B>10)\cap(C=10))\\
=&1-P((A\le 20)\cap(B\le 20)|(A>10)\cap(B>10)\cap(C=10)).
\end{aligned}
$$

We can calculate this probability similarly to the previous exercise as

$$\begin{aligned}
&P((A \leq 20) \cap (B \leq 20)|(A > 10) \cap (B > 10) \cap (C = 10))\\
=&\frac{P((A \leq 20) \cap (B \leq 20) \cap (A > 10) \cap (B > 10) \cap (C = 10))}{P((A > 10) \cap (B > 10) \cap (C = 10))}\\
=&\frac{P((10 < A \leq 20) \cap (10 < B \leq 20) \cap (C = 10))}{P((A > 10) \cap (B > 10) \cap (C = 10))}\\
=&\frac{P(10 < A \leq 20)P(10 < B \leq 20)P(C = 10)}{P(A > 10)P(B > 10)P(C = 10)}\\
=&\frac{P(10 < A \leq 20)P(10 < B \leq 20)}{P(A > 10)P(B > 10)}\\
=&\frac{(10/30)^2}{(20/30)^2} = \frac{1}{4}
\end{aligned}$$

where in the first step we used the definition of conditional probability, in the second step we simplified the intersection of the events for A and B respectively, in the third step we used independence, in the fourth we simplified by $P(C = 10)$, and finally we used that for uniform random variables, the probabilities are proportional to the interval (or you can see this by using $P(10 < A \leq 20) = F_A(20) - F_A(10)$). Therefore, the probability that Carl has to wait more than 10 minutes for his party to be completed is $1 - 1/4 = 3/4$.

4. The event that the second person to arrive has to wait more than 5 minutes for the third person to arrive, corresponds to either Bob arriving 5 minutes after Alice, or Alice arriving 5 minutes after Bob. In our notation this is $P((A > B+5) \cup (B > A+5)|(A > 10) \cap (B > 10) \cap (C = 10))$. As the two events are mutually exclusive

$$\begin{aligned}
&P((A > B+5) \cup (B > A+5)|(A > 10) \cap (B > 10) \cap (C = 10))\\
=&P(A > B+5|(A > 10) \cap (B > 10) \cap (C = 10))\\
+&P(B > A+5|(A > 10) \cap (B > 10) \cap (C = 10)),
\end{aligned}$$

and moreover, by the symmetry of the random variables

$$\begin{aligned}
&P(B > A+5|(A > 10) \cap (B > 10) \cap (C = 10)\\
=&P(A > B+5|(A > 10) \cap (B > 10) \cap (C = 10)).
\end{aligned}$$

Then,

$$\begin{aligned}
&P(A > B+5|(A > 10) \cap (B > 10) \cap (C = 10))\\
=&\frac{P((A > B+5) \cap (A > 10) \cap (B > 10) \cap (C = 10))}{P((A > 10) \cap (B > 10) \cap (C = 10))}\\
=&\frac{P((A > B+5) \cap (25 > B > 10) \cap (C = 10))}{P((A > 10) \cap (B > 10) \cap (C = 10))}\\
=&\frac{P((A > B+5) \cap (25 > B > 10))P(C = 10)}{P((A > 10))P(B > 10)P(C = 10)}\\
&\frac{P((A > B+5) \cap (25 > B > 10))}{P((A > 10))P(B > 10)}
\end{aligned}$$

where in the first step we used the definition. In the second step, we noticed that $(B > 10) \cap (A > B + 5)$ implies $A > 10$, so we simplified by that, and that $A > B + 5$ implies $B < 25$, as otherwise, A would be greater than 30, i.e. it would be outside of the support. In the last two steps we used independence, and simplified by $P(C = 10)$.

The denominator is 4/9, similarly to the previous two sub-exercises, but for the numerator, we cannot just use the CDF of the marginal Uniform random variables, as in the inequality $A > B + 5$ we have both A and B. In this case, we must calculate the probability as the integral of the joint PDF, on the intervals defined by the inequalities:

$$
\begin{aligned}
&\int_{10}^{25}\int_{b+5}^{30} f_{A,B}(a,b)\,\mathrm{d}a\,\mathrm{d}b \\
&= \int_{10}^{25}\int_{b+5}^{30} f_A(a)f_B(b)\,\mathrm{d}a\,\mathrm{d}b \\
&= \left(\frac{1}{30}\right)^2 \int_{10}^{25}\int_{b+5}^{30} \mathrm{d}a\,\mathrm{d}b \\
&= \left(\frac{1}{30}\right)^2 \int_{10}^{25} (25-b)\,\mathrm{d}b \\
&= \left(\frac{1}{30}\right)^2 \left[25b - \frac{b^2}{2}\right]_{b=10}^{25} \\
&= \left(\frac{1}{30}\right)^2 (625 - \frac{625}{2} - 250 + \frac{100}{2}) = \frac{225}{2\cdot 900},
\end{aligned}
$$

where in the first step we used independence, in the second step the per definition PDF of a $Unif(0, 30)$, and then solved the integral.

Plugging this back as the numerator, and taking $\frac{4}{9}$ as the denominator (see exercises above), $P(A > B + 5|(A > 10) \cap (B > 10) \cap (C = 10)) = \frac{225}{2\cdot900} \cdot \frac{9}{4} = \frac{225}{2\cdot400}$.

Therefore, the probability that the second person has to wait more than 5 minutes for the last person to arrive, is $2 \cdot \frac{225}{2\cdot400} = \frac{225}{400} = \frac{9}{16}$.

Exercise 4. A fair coin is flipped twice. Let X be the number of Heads in the two tosses, and Y be the indicator r.v for the tosses landing the same way.

1. Find the joint PMF of X and Y.
2. Find the marginal PMFs of X and Y.
3. Are X and Y independent?
4. Find the conditional PMFs of Y given $X = x$ and of X given $Y = y$.

Solution 4. 1. Denote the outcomes of the two tosses with H, and T, for Heads and Tails, respectively. The four possible outcomes of this experiment are HH, HT, TH, and TT. These outcomes are giving the the different values for X and Y as follows:

$$
\begin{aligned}
HH &\implies X = 2, Y = 1, \\
HT &\implies X = 1, Y = 0, \\
TH &\implies X = 1, Y = 0, \\
TT &\implies X = 0, Y = 1,
\end{aligned}
$$

so the joint PMF can be written as in Table 1

$Y \backslash X$	$X=0$	$X=1$	$X=2$
$Y=0$	0	$\frac{1}{2}$	0
$Y=1$	$\frac{1}{4}$	0	$\frac{1}{4}$

Table 1: Joint PMF of X and Y

2. By definition the marginal PMF of X is $P(X=x) = \sum_y P(X=x, Y=y)$, so in this example

$$P(X=0) = P(X=0, Y=0) + P(X=0, Y=1) = 0 + \frac{1}{4} = \frac{1}{4}$$
$$P(X=1) = P(X=1, Y=0) + P(X=1, Y=1) = \frac{1}{2} + 0 = \frac{1}{2}$$
$$P(X=2) = P(X=2, Y=0) + P(X=2, Y=1) = 0 + \frac{1}{4} = \frac{1}{4}.$$

Similarly, the marginal PMF of Y is $P(Y=y) = \sum_x P(X=x, Y=y)$, so

$$P(Y=0) = P(X=0, Y=0) + P(X=1, Y=0) + P(X=2, Y=0) = 0 + \frac{1}{2} + 0 = \frac{1}{2}$$
$$P(Y=1) = P(X=0, Y=1) + P(X=1, Y=1) + P(X=2, Y=1) = \frac{1}{4} + 0 + \frac{1}{4} = \frac{1}{2}.$$

3. No they are not independent. From Table 1 we can read that $P(X=1, Y=1) = 0$, but from part 2., $P(X=1) = \frac{1}{2}$, and $P(Y=1) = \frac{1}{2}$, so

$$P(X=1, Y=1) = 0 \neq \frac{1}{4} = P(X=1)P(Y=1).$$

4. For discrete random variables, the conditional PMF is the joint PMF divided by the marginal PMF, that is, $P(Y=y|X=x) = \frac{P(Y=y, X=x)}{P(X=x)}$. Then,

$P(Y=y\|X=x)$	$X=0$	$X=1$	$X=2$
$Y=0$	0	1	0
$Y=1$	1	0	1

Table 2: Conditional PMF of Y given X

That means, that by knowing the value of X, we can tell the value of Y with certainty, which sounds reasonable because if we know how many heads we have out of two flips, we instantly know, whether the flips were the same or not.

Exercise 5. Let X and Y have joint PDF $f_{X,Y}(x, y) = cxy$, for $0 < x < y < 1$.

1. Find c to make this a valid joint PDF.
2. Find the marginal PDFs of X and Y.
3. Are X and Y independent?
4. Find the conditional PDF of Y given $X = x$.

Solution 5. 1. To ensure that the PDF is valid, we have to find a c, such that the PDF is non-negative, and it integrates to 1. The PDF will be non-negative, as long as $c \geq 0$. Now, we have to integrate over the support of X and Y, and since $x < y$ must hold, that is

$$\int_0^1 \int_0^y cxy \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \frac{y^2}{2} cy \, \mathrm{d}y = \frac{1}{8}c,$$

so the integral is equal to one if $c = 8$.

2. The marginal of X and Y are defined respectively as $f_X(x) = \int_0^1 f_{X,Y}(x,y) \, \mathrm{d}y$ and $f_Y(y) = \int_0^1 f_{X,Y}(x,y) \, \mathrm{d}x$, therefore

$$f_X(x) = \int_x^1 8xy \, \mathrm{d}y = 4x - 4x^3$$

for $0 < x < 1$. and

$$f_Y(y) = \int_0^y 8xy \, \mathrm{d}x = 4y^3$$

for $0 < y < 1$.

3. Since $f_{X,Y}(x,y) = 8xy \neq (4x - 4x^3) \cdot 4y^3 = f_X(x) \cdot f_Y(y)$, for all x and y, the two random variables are not independent.

4. By definition $f_{Y|X}(y|x) = \frac{f_{Y,X}(y,x)}{f_X(x)} = \frac{8xy}{4x-4x^3} = \frac{2y}{1-x^2}$ for $0 < x < y < 1$.

Exercise 6. Two students, A and B, are working independently on homework assignments. Student A takes $Y_1 \sim \mathit{Expo}(\lambda_1)$ hours to finish their homework, while B takes $Y_2 \sim \mathit{Expo}(\lambda_2)$ hours.

1. Find the CDF and PDF of Y_1/Y_2, the ratio of their problem-solving times.

2. Find the probability that A finishes their homework before B does.

Solution 6. 1. Let $t > 0$. The CDF of the ratio is

$$F(t) = P\left(\frac{Y_1}{Y_2} \leq t\right) = P(Y_1 \leq tY_2)$$

$$= \int_0^\infty \left(\int_0^{ty_2} \lambda_1 e^{-\lambda_1 y_1} \, dy_1\right) \lambda_2 e^{-\lambda_2 y_2} \, dy_2$$

$$= \int_0^\infty \left(1 - e^{-\lambda_1 t y_2}\right) \lambda_2 e^{-\lambda_2 y_2} \, dy_2$$

$$= 1 - \int_0^\infty \lambda_2 e^{-(\lambda_1 t + \lambda_2) y_2} \, dy_2$$

$$= 1 - \frac{\lambda_2}{t\lambda_1 + \lambda_2}$$

$$= \frac{t\lambda_1}{t\lambda_1 + \lambda_2}.$$

Of course, $F(t) = 0$ for $t \leq 0$. The PDF of the ratio is

$$f(t) = \frac{d}{dt}\left(\frac{t\lambda_1}{t\lambda_1 + \lambda_2}\right) = \frac{\lambda_1 \lambda_2}{(\lambda_1 t + \lambda_2)^2}, \quad \text{for } t > 0.$$

2. Plugging in $t = 1$ above, we have

$$P(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Alternatively, we can get the same result by applying Example 7.1.23. (The result can also be derived without using calculus by thinking about Poisson processes, as shown in Chapter 13.)

Exercise 7. Let X and Y be discrete r.v.s.

1. Use 2D LOTUS (without assuming linearity) to show that $E(X+Y) = E(X) + E(Y)$.

   Hint: If you are really stuck, check Example 7.2.4, but after that, this exercise reduces to copying.

2. Now suppose that X and Y are independent. Use 2D LOTUS to show that $E(XY) = E(X)E(Y)$.

Solution 7. 1. Let X and Y be discrete random variables. Using the 2D Law of the Unconscious Statistician (LOTUS), we have:

$$E(g(X,Y)) = \sum_x \sum_y g(x,y)\, P(X = x, Y = y),$$

where we can replace $g(x,y)$ with $x + y$, giving us

$$E(X+Y) = \sum_x \sum_y (x+y)\, P(X = x, Y = y).$$

Expanding $(x + y)$ inside the sum, and using that summation is linear, we get:

$$\begin{aligned} E(X+Y) &= \sum_x \sum_y x\, P(X = x, Y = y) + \sum_x \sum_y y\, P(X = x, Y = y) \\ &= \sum_x x \sum_y P(X = x, Y = y) + \sum_y y \sum_x P(X = x, Y = y) \end{aligned}$$

By definition, the marginal PMF-s are $\sum_y P(X = x, Y = y) = P(X = x)$ and $\sum_x P(X = x, Y = y) = P(Y = y)$. we have:

$$E(X+Y) = \sum_x x\, P(X = x) + \sum_y y\, P(Y = y).$$

Thus,

$$E(X+Y) = E(X) + E(Y).$$

2. By the 2D LOTUS, with $g(x,y) = x \cdot y$ we can rewrite $E(g(X,Y))$ as

$$E(XY) = \sum_x \sum_y xy\, P(X = x, Y = y).$$

By independence, $P(X = x, Y = y) = P(X = x)P(Y = y)$, so we can take out both x and $P(X = x)$ from the inner sum, giving

$$\sum_x \sum_y xy\, P(X = x, Y = y) = \sum_x x\, P(X = x) \sum_y y\, P(Y = y) = \sum_x x\, P(X = x)E(Y).$$

Using that $E(Y)$ is constant in x,

$$\sum_x x\, P(X = x)E(Y) = E(Y) \sum_x x\, P(X = x) = E(Y)E(X),$$

as desired.

Correlation Exercises

If during the lecture the definition of *covariance* and *correlation* are not covered, leave these exercises for next week.

Exercise 8. Two fair, six-sided dice are rolled (one green and one orange), with outcomes X and Y for the green die and the orange die, respectively.

1. Compute the covariance of $X + Y$ and $X - Y$.
2. Are $X + Y$ and $X - Y$ independent?

Solution 8.

1. We have $Cov(X + Y, X - Y) = Cov(X, X) - Cov(X, Y) + Cov(Y, X) - Cov(Y, Y) = 0$, as $Cov(X, X) = Var(X)$, and because, X and Y are identically distributed, $Var(X) = Var(Y)$.
2. They are not independent: information about $X + Y$ may give information about $X - Y$, as shown by considering an extreme example. Note that if $X + Y = 12$, then $X = Y = 6$, so $X - Y = 0$. Therefore,$P(X - Y = 0|X + Y = 12) = 1 \neq P(X - Y = 0)$, which shows that $X + Y$ and $X - Y$ are not independent. Alternatively, note that $X + Y$ and $X - Y$ are both even or both odd, since the sum $(X + Y) + (X - Y) = 2X$ is even.