# MATH 323 - Tutorial 4 Solutions

Tyrel Stokes

October 2020

1. Stonks!

Consider two different stocks, $X_1$ and $X_2$. Suppose that there are three different states of the world, say for example there is a state where the economy crashes ($S_b$), another where status quo is maintained ($S_m$), and the state of the world where the economy booms $S_g$. The stocks will have different values depending on the state of the world. You can assume that the value for both stocks will be lowest in $S_b$, at least as great as the bad state in $S_m$, and at least as great as that in the boom economy $S_b$. It is known that $P(S_b) = \frac{1}{2}$, $P(S_m) = \frac{1}{6}$, $P(S_g) = \frac{1}{3}$.

Suppose we know the CDF of $X_1$ is:

$$P(X_1 \leq x_1) = \begin{cases} = 0, x_1 < 0 \\ = \frac{1}{2}, 0 \leq x_1 < 4 \\ = \frac{2}{3}, 4 \leq x_1 < 7 \\ = 1, x \geq 7 \end{cases}$$

For $X_2$ we know that if the economy is bad then it will be worth 1 and it will be worth 5 if the economy either stays the same or booms.

a) This is not a real probability question, but imagine you were forced to buy one of the two stocks (without being able to predict the state of the economy, i.e only knowing the marginal probabilities above). Which would you prefer?

**Solution**

Not a real question, just something to think about. No answer.
b) Find the PMF of $X_1$ and the CDF of $X_2$ (draw it).

**Solution**

To translate a CDF to PMF, we can look at the jumps.

$$P(X_1 = x_1) = \begin{cases} \frac{1}{2}, & x_1 = 0 \\ \frac{1}{6}, & x_1 = 4 \\ \frac{1}{3}, & x_1 = 7 \end{cases}$$

Translating the information we get from the question, we were given the pmf of $X_2$

$$P(X_2 = x_2) = \begin{cases} \frac{1}{2}, & x_2 = 1 \\ \frac{1}{2}, & x_2 = 5 \end{cases}$$

Which translated into a CDF gives us:

$$P(X_2 \leq x_2) = \begin{cases} 0, & x_2 < 1 \\ \frac{1}{2}, & x_2 \in [1, 5) \\ 1, & x_2 \geq 5 \end{cases}$$

We will return to this question in a few weeks and explore more of the Stocks properties.

2. When coded messages are sent, there are sometimes errors in transmission. In particular, Morse code uses dots and dashes, which are know to be sent in the proportions $3dots : 4dashes$. Suppose there is interference in the signal and potentially human error in properly distinguishing a dot from a dash, so that sometimes a dot which is sent is received as a dash and vice versa. Suppose that it is known that with probability $\frac{1}{4}$ a dot is mistakely encoded as a dash and with probability $\frac{1}{3}$ a dash is encoded as a dot.

a) If a dash is recieved, what is the probability that a dash was actually sent?

b) Now consider the random variable $X$, where $X :=$ "The number of dots recieved" is the sample space generated from the experiment of sending and receiving $n$ independent messages.

i) What is the support of $X$?
ii) what is the pmf of $X$ written in terms of $p_{dot_R} = P(DOT\ Received)$?

**Solutions:**

Attached separately!

3. Consider a sequence of random variables $Y_1, Y_2, \ldots Y_n$ generated from the experiment of flipping a coin n times independently.

Let

$$Y_i = \begin{cases} 1 \text{ if the ith flip is heads} \\ 0 \text{ if the ith flip is tails} \end{cases}$$

Suppose that the probability of a heads is $P(Y_i = 1) = p_y, p_y \in (0,1), \forall i = 1, 2, \ldots, n$. These are bernouilli variables.

Now consider a random variable, which is a function of the sequence of variables.

Let $W_1 = min(Y_1, Y_2, \ldots, Y_n)$.

a) i) What is the support of the random variable $W_1$?
(You may not have seen this word support in class and that is okay. All it is the set of numbers that the random variable can take with positive probability, $\{x : P(X = x) > 0\}$ for some random variable $X$)
ii) Write down the PMF of $W_1$ in terms of $p_y$

b) Consider the random variable $W_2 = max(Y_1, Y_2, \ldots, Y_n)$
i) What is it's support
ii) Write down the PMF in terms of $p_y$

c) Assume now that $p_y = 0.5$ and that $n = 5$.
i) Find the expectations of $W_1$ and $W_2$ respectively.
ii) Challenge: Find the expectation: $E[|W_1 - W_2|]$? [There is a way to simplify this, do you see it?]

**Solutions**

a.i) $\{w_1 : P(W_1 = w_1) > 0\} = \{0, 1\}$

a.ii) From above, we know that we simply have to find $P(W_1 = 1)$ and $P(W_1 = 0)$, since all other probabilities are 0. Further we know that all probability mass functions must sum to probability 1, and that the events $\{W_1 = 1\}$ and $\{W_1 = 0\}$ are disjoint, thus forming a partition of the sample space. In general, we can always use the support of a discrete random variable to partition the sample space (and this is used often implicitly or explicitly in probability). Thus, we only need to find one of these probabilities and using the complement rule, we get the other. In general it pays in these kinds of questions to think about which will be easiest. In this case, when $W_1 = 0$, all it means is that at least one of the bernouilli variables $Y_1, \ldots, Y_n$ is 0. There are many ways for this to be the case, making this problem not necessarily easy. On the other hand, if $W_1 = 1$ if means that every single bernouilli value was 1, this is an easy probability to calculate.

$$
\begin{aligned}
P(W_1 = 1) &= P(\cap_{i=1}^{n} \{Y_i = 1\}) && (1) \\
&= \prod_{i=1}^{n} P(Y_i = 1) \quad [\text{ Independence }] && (2) \\
&= (p_y)^n \quad [\text{ Identical }] && (3)
\end{aligned}
$$

$$
\begin{aligned}
P(W_1 = 0) &= 1 - P(W_1 = 1) &(4)\\
&= 1 - (p_y)^n &(5)
\end{aligned}
$$

$$
P(W_1 = w_1) = \begin{cases} 1 - p_y^n, w_1 = 0 \\ p_y^n, w_1 = 1 \\ 0, \quad \text{else} \end{cases} \tag{6}
$$

Here are two alternative ways to get $P(W_1 = 0)$

$$
\begin{aligned}
P(W_1 = 0) &= P(\cup_{i=1}^n \{Y_i = 0\}) &(7)\\
&= 1 - P((\cup_{i=1}^n \{Y_i = 0\})^c) &(8)\\
&= 1 - P(\cap_{i=1}^n \{Y_i = 0\}^c) \quad \text{[DeMorgan's]} &(9)\\
&= 1 - P(\cap_{i=1}^n \{Y_i = 1\}) &(10)\\
&= 1 - p_y^n &(11)
\end{aligned}
$$

Where the last step follows from the steps and assumptions invoked in our derivation of $P(W_1 = 1)$. A third way involves partitioning the event into all of the ways that you could get at least 1 variable $Y_i$ with the value 0. Let's define a variable $Z = \sum_{i=1}^n (1 - Y_i)$, which simply sums up the number of zeros. (The savvy reader, will notice that $Z$ is a binomial ($1 - Y_i$ are identical and independently drawn bernoulli variables). If this was not recognized, it could still be easily derived from first principles.

$$
\begin{aligned}
P(W_1 = 0) &= P(Z \geq 1) &(12)\\
&= \sum_{z=1}^{n} P(Z = z) &(13)\\
&= \sum_{z=1}^{n} \binom{n}{z} (1 - p_y)^z p_y^{n-z} \quad \text{[ Binomial ]} &(14)\\
&= 1 - \binom{n}{0} p_y^n &(15)\\
&= 1 - p_y^n &(16)
\end{aligned}
$$

Where the second last line foll0ws from the from the fact that p.m.fs sum to one, $1 = \sum_{z=0}^n \binom{n}{z}(1-p_y)^z p_y^{n-z} = \binom{n}{0} p_y^n + \sum_{z=1}^n \binom{n}{z}(1 - p_y)^z p_y^{n-z} = 1$. Equivalently we could have just as easily noted that $P(Z \geq 1) = 1 - P(Z = 0)$ using the same reasoning that the support of the discrete random variable forms a partition.

The lesson here is that in a sense all (valid) roads lead to Rome, and that there are many ways to attack the same problem. But most of the ways will come down to the same kind of reasoning. In every approach, one way or another, we end up tranforming the problem of finding or evaluating the "hard" probability ($P(W_1 = 0)$) into the easier problem of finding it's complement.

b) Here I will spare you the sermon and simply write one approach, but note that the same kind of approaches as above could be used.
i) $\{w_2 : P(W_2 = w_2) > 0\} = \{0, 1\}$
ii) Here the easy probability to calculate is $P(W_2 = 0)$ since it means all $Y_i$'s are zero.

$$
\begin{aligned}
P(W_2 = 0) &= P(\cap_{i=1}\{Y_i = 0\}) && (17)\\
&\overset{independence}{=} \prod_{i=1}^{n} P(Y_i = 0) && (18)\\
&= (1 - p_y)^n && (19)
\end{aligned}
$$

$$
\begin{aligned}
P(W_2 = 1) &= 1 - P(W_2 = 0) && (20)\\
&= 1 - (1 - p_y)^n && (21)
\end{aligned}
$$

$$
P(W_2 = w_2) = \begin{cases} (1 - p_y)^n, w_2 = 0 \\ 1 - (1 - p_y)^n, w_2 = 1 \\ 0, \quad \text{else} \end{cases} \tag{22}
$$

c) Let $p_y = 0.5$ and $n = 5$

i)

$$
\begin{aligned}
E[W_1] &= \sum_{w_1=0}^{1} w_1 P(W_1 = w_1) && (23)\\
&= 0 \times P(W_1 = 0) + 1 \times P(w_1 = 1) && (24)\\
&= P(W_1 = 1) && (25)\\
&= p_y^n && (26)
\end{aligned}
$$

$$
\begin{aligned}
E[W_2] &= \sum_{w_2=0}^{1} w_2 P(W_2 = w_2) && (27)\\
&= 0 \times P(W_2 = 0) + 1 \times P(w_2 = 1) && (28)\\
&= P(W_2 = 1) && (29)\\
&= 1 - (1 - p_y)^n && (30)
\end{aligned}
$$

The expectation of a bernoulli random variable is always just the probability that the bernoulli random variable take the value 1! This is a useful fact that is used often in probability theory.
ii) This question in hindsight, would be better saved for later in the course, but I will show you that you can still solve it with the tools you have. One of the best and most useful properties of the Expectation operator is that it is in fact a linear operator. We have already calculated the expectations $E[W_1]$ and $E[W_2]$ so we need to find a way to be able to invoke the linearity assumption. The problem is the frustrating absolute value sign, which is messing us up. $E[|W_1 - W_2|] \neq E[|W_1|] - E[|W_2|]$ in general since this is not a linear function. However in this specific case:

$$\begin{aligned}
E[|W_1 - W_2|] &= E[|min(Y_1, \ldots, Y_n) - max(Y_1, \ldots, Y_n)|] && (31)\\
&= E[-(min(Y_1, \ldots, Y_n) - max(Y_1, \ldots, Y_n))] \quad \text{Since } min(Y_1, \ldots, Y_n) \leq max(Y_1, \ldots, Y_n) && (32)\\
&= E[-W_1 + W_2] && (33)\\
&= -E[W_1] + E[W_2] && (34)\\
&= -(p_y)^n + (1 - (1 - p_y)^n) && (35)\\
&= -(.5)^5 + 1 - (0.5)^5 && (36)\\
&= \frac{-1}{32} + \frac{31}{32} && (37)\\
&= \frac{15}{16} && (38)
\end{aligned}$$

So here we were able to transform the non-linear function into a nice linear one we know how to deal with by using the fact that the min must be less than the max. We will learn more general techniques for finding the expectations of non-linear functions of more than one variable later in the course. Another approach that only requires the tools we have developed thusfar would be to find the distribution of $|W_1 - W_2|$ directly.