# MATH 323 - Tutorial 6 Questions

Tyrel Stokes

October 2020

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

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:

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

c) Find the mean and variance of $X_1$ and $X_2$.

I will define the support for each r.v, $\mathcal{X}_1$ and $\mathcal{X}_2$ respectively. This is just the values that take positive probability so I know which values to sum over for the expectatations.

$$\mathcal{X}_1 = \{0, 4, 7\}$$
$$\mathcal{X}_2 = \{1, 5\}$$

Then

$$\begin{aligned} E[X_1] &= \sum_{x_1 \in \mathcal{X}_1} x_1 P(X_1 = x_1) \\ &= 0 * P(X_1 = 0) + 4 * P(X_1 = 4) + 7 * P(X_1 = 7) \\ &= 4(\frac{1}{6}) + 7(\frac{1}{3}) \\ &= \frac{2}{3} + \frac{7}{3} \\ &= 3 \end{aligned}$$

$$\begin{aligned} E[X_2] &= \sum_{x_2 \in \mathcal{X}_2} x_2 P(X_2 = x_2) \\ &= 1 * P(X_2 = 1) + 5 * P(X_2 = 5) \\ &= 1 * (\frac{1}{2}) + 5(\frac{1}{2}) \\ &= 3 \end{aligned}$$

The work for the variances is ommitted, but the answers are:

$$Var(X_1) = 10$$
$$Var(X_2) = 4$$

d)Now which stock would you buy? Why? Again, there is no absolute right answer.

**Solution**

Again, this is not a real question, just a prompt to get you thinking about probability in a different way.

In part c we found that on average, the two stocks will have the same value, but their variances are quite different. We expect at any given moment that the value of $X_1$ is farther away from this average value than $X_2$. This is related to risk. The first stock is riskier. People that are risk averse usually want to minimize their risk. One way to think about this is that they are willing to trade value on average, for more certainty. In that case, someone who is risk averse would definitely prefer $X_2$ to $X_1$ because on average they have the same value but $X_2$ is much more certain and huddled closer around the mean. It's okay to not share these preferences though.

This extra information about risk-aversion is not part of the course and you are not expected to know this. There reason it is included here is because I want you to start thinking about the kind of information that moments, like the mean and variance can tell us about random variables. This will become more important later in the course when in lecture you learn about things like the moment generating function which is another way or characterizing, or telling the story of, certain random variables.

e) Now consider just $X_2$ and following states, $A_1 = S_b$, $A_2 = S_m \cup S_g$. Suppose we have an index, $Y$ that gives us information about whether the economy will be in state $A_1$ or $A_2$ in the near future.

Let
$a* = \alpha + y + \epsilon$, where $y$ is a particular value of the index $Y$, $\alpha$ is some constant, and $\epsilon \sim N(0, 1)$ is a standard normal variable.

Let $P(A_2|Y = y) = P(a^* > 0)$. Find $P(X_2 = x_2|Y = y)$ for all admissible values $x_2$ in terms of $y, \alpha$ and the normal CDF.

**Solution**

We know that $X_2$ can only take the values 1 or 5, so we only need to solve this expression for those two values, everywhere else the probability is zero. The trick here again is that we really only know how $X_2$ relates to $Y$ through the state $A_2$. Again, we have a situation where we wish we knew the state, but we do not. The solution is to condition on it using the conditional total law of probability:

$$P(X_2 = x_2|Y = y) \overset{CLTP}{=} P(X_2 = x_2|Y = y, A_2)P(A_2|Y = y) + P(X_2 = x_2|Y = y, A_2^c)P(A_2^c|Y = y)$$

Now let's find $P(A_2|Y=y)$.

$$
\begin{aligned}
P(A_2|Y=y) &= P(a^* > 0) \\
&= P(\alpha + y + \epsilon > 0) \\
&= P(\epsilon > -(\alpha + y)) \\
&= 1 - P(\epsilon \leq (-(\alpha + y)) \\
&\overset{\epsilon \text{ std. normal}}{=} 1 - \Phi(-(\alpha + y))
\end{aligned}
$$

where $\Phi(x) = P(X \leq x)$ for a standard normal random variable $X$.
Therefore

$$
\begin{aligned}
P(A_2^c|Y=y) &= 1 - P(A_2|Y=y) \\
&= \Phi(-(\alpha + y))
\end{aligned}
$$

Knowing the state $A_2$ or $A_2^c$ is enough to know the exact value of $X_2$. So putting everything together:

$$
\begin{aligned}
P(X_2 = 1|Y=y) &\overset{CLTP}{=} P(X_2 = 1|Y=y, A_2)P(A_2|Y=y) + P(X_2 = 1|Y=y, A_2^c)P(A_2^c|Y=y) \\
&= 0 \times (1 - \Phi(-(\alpha + y)) + 1 \times \Phi(-(\alpha + y)) \\
&= \Phi(-(\alpha + y))
\end{aligned}
$$

and

$$
\begin{aligned}
P(X_2 = 5|Y=y) &\overset{CLTP}{=} P(X_2 = 5|Y=y, A_2)P(A_2|Y=y) + P(X_2 = 5|Y=y, A_2^c)P(A_2^c|Y=y) \\
&= 1 \times (1 - \Phi(-(\alpha + y)) + 0 \times \Phi(-(\alpha + y)) \\
&= 1 - \Phi(-(\alpha + y))
\end{aligned}
$$

f) Bonus Question (Less directly related to lecture material, but may be of interest to some of you). Sometimes in statistical decision theory we use utility function (or alternatively loss functions or cost functions) to try and map outcomes to their value. This can take into account our preferences around uncertainty in a given application. An example of a utility function is $U : \mathcal{R}^+ \to \mathcal{R}^+$, $U(x) = \frac{x^{1-\alpha}}{1-\alpha}$ where $\alpha > 0, \alpha \neq 1$ is a parameter indicating risk tolerance. If $\alpha$ is close to 0, this indicates that a person does not negatively weight uncertainty very much, they make decisions based on what happens on average. If $\alpha$ is very large, say $\alpha > 3$ this indicates that they strongly prefer low amounts of variation to high variation and they are willing to sacrifice gains on average for extra certainty. $E[U(X_1)]$ is the expected utility of stock 1 and $E[U(X_2)]$ is the expected utility of stock two. Suppose $\alpha = \frac{1}{2}$, which stock would someone with preferences represented by the above utility curve prefer? I.e, which stock has higher expected utility? Think about how this relates to the information we learned about the stocks through their first two moments.

**Solution**

Proof Ommitted: $E[U(X_1)] < E[U(X_2)]$.

The reason why is related to the explanation in part d). The utility curve represents a risk averse preference, so such a person will get high expected utility from the less risky option $X_2$.

2. Consider a telephone operator who, in expectation, handles 5 calls every three minutes. Assume that the number of calls in disjoint intervals of time is independent and that as a time interval gets infinitely small the probability of more than one call (simultaneous calls) goes to 0. (Modified question from Cassella and Berger)

a) What distribution would be a good modelling choice for the number of calls in a minute?

**Solution**

Poisson Distribution. The assumptions stated are directly from the "Poisson Postulates" which are a way of deriving the poisson distribution. This wasn't seen in class directly, but many of these ideas are present in the proof of the binomial approximation to the poisson which is seen in class.

Directly one can rule out any continuous random variable, since this clearly takes values in $\{0, 1, 2, \dots\}$. We can further use this to rule out many other discrete random variables like the Bernoulli, Binomial. Geometric does share this property but is motivate as a number of trials until failure. Similarly, the negative binomial is not appropriate. This should give us a pretty good idea already that Poisson is at least a reasonable candidate.

The independent and identical counts between intervals of time is often the assumption of the Poisson that is checked in real life applications and is one that you should familiarize yourself with. A good example of when a seemingly Poisson random variable fails this assumption is found in modelling the number of scores in sports. Take European football (Soccer) as an example. One might think that the number of goals scored in a game for each team could be well approximated by a Poisson. If it were true, we would expect the number of goals scored on average in first half of games to be the same as the number in the second half. Empirically, this is not true. Many goals are scored in the last 10 minutes of the game in fact. Thus this fails the identical part of the assumption.

Also knowing the score can help you predict the number of goals in many sports. In hockey for example, teams that are leading slow down their scoring. If goals were Poisson (and thus scoring independent across time periods) knowing the score, i.e the information about counts from the previous time periods, would not help us predict future scoring once we know the average goal scoring rates.

All this to say that we have to be careful when modelling things in real life. It is not enough to just know which values a random variable can take to model it correctly. Although that information can potentially rule out distributions that would not make good candidates.

b) Supposing the distribution in a, what is the probability of getting exactly 0 calls in a minute?

**Solution**

Using the fact that for $X \sim Poisson(\lambda)$ that $P(X = x) = \frac{\exp(-\lambda)\lambda^x}{x!}$ we can find this probability. We only need $\lambda$. From class we know that $E[X] = \lambda$ for a Poisson random variable. In the question we are told that on average there are 5 calls in 3 minutes. This if $X$" ="number of calls in a minute", we have that $E[X] = \frac{5}{3} = \lambda$.

$$\begin{aligned}
P(X = 0) &= \frac{\exp(-\frac{5}{3})\frac{5}{3}^0}{0!} \\
&= \exp(-\frac{5}{3}) \\
&\approx .189
\end{aligned}$$

c) What is the probability of at least 2 calls in a minute?

**Solution**

$$\begin{aligned}
P(X \leq 2) &= 1 - P(X < 2) \\
&= 1 - P(X = 0) - P(X = 1) \\
&= 1 - \exp(-\frac{5}{3}) - \exp(-\frac{5}{3})\frac{5}{3} \\
&\approx .496
\end{aligned}$$

d) Suppose I knew in a particular minute that there would be at least 2 phone calls, what is the probability that there will be at least four?

**Solution**

$$\begin{aligned}
P(X \geq 4 | X \geq 2) &= \frac{P(\{X \geq 4\} \cap \{X \geq 2\})}{P(X \geq 2)} \\
&\stackrel{subset}{=} \frac{P(X \geq 4)}{P(X \geq 2)} \\
P(X \geq 4) &= 1 - P(X < 4) \\
&= 1 - P(X < 2) - P(X = 2) - P(X = 3) \\
&= 1 - \exp(-\frac{5}{3}) - \exp(-\frac{5}{3})\frac{5}{3} - \frac{\exp(-\frac{5}{3})\frac{5}{3}^2}{2} - \frac{\exp(-\frac{5}{3})\frac{5}{3}^3}{6} \\
&\approx .08817 \\
P(X \geq 4 | X \geq 2) &\approx \frac{0.08817}{.496} \\
&= .1778
\end{aligned}$$

e) Challenge Question: Again assuming that the number of calls from minute to minute are independent, consider the following experiment. The boss in the next room feels like most minutes there aren't any calls (doesn't know much about probability, just has a gut feeling). Using the calculations from part c, the telephone operator says that almost half of the time there at least 2 calls in any given minute, but the boss isn't convinced by "fancy maths". The boss and the telephone operator make a bet. They monitor the phones and count the number of calls in each minute. If there is a minute with exactly 0 calls before a minute with at least 2 calls then the boss wins and the telephone operator agrees to do an extra overtime shift. If there is a minute with at least 2 calls before a minute with exactly zero calls then the operator wins the bet, then the operator gets to go home early and the boss with handle the incoming calls. What is the probability that the operator wins the bet?

**Solution**

We need to define some events and random variables. Let $X_i$ be the number of calls in the ith minute. We know that these are identical poissons ($\lambda = \frac{5}{3}$) and independent of each other.

Let $W_i :=$ "The telephone operator wins in exactly the ith minute"

In order to the telephone operator to win, they must win in one of the minutes. In other words we can partition the event $W :=$ "The event the telephone operator wins the bet" into the event they win in the 1st minute, or the 2nd minute, or third and so on. Formally $W = \cup_{i=1}^{n} W_i$, where the $W_i$s are a partition (and thus disjoint).

$$\begin{aligned} P(W) &= P(\cup_{i=1}^{\infty} W_i) \\ &\overset{disjoint}{=} \sum_{i=1}^{\infty} P(W_i) \end{aligned}$$

Now we just need to figure out the probabilities $P(W_i)$ and evaluate the infinite sum. What must happen in order for the telephone operator to win in exactly the ith minute? Well, neither of the two players can have won in all of the previous i-1 minutes.

We can split up each minute into 3 events: Either the boss wins ($\{X_i = 0\}$) or the telephone operator wins $\{X_i \geq 2\}$ or neither of them win and we try again the next minute ($\{X_i = 0\} \cup \{X_i \geq 2\})^c = \{X = 1\}$). Thus in order to win in the ith minute, it must have been that $\{X_j = 1\}$ for all $j \in \{1, 2, \ldots, i-1\}$. Therefore

$$
\begin{aligned}
P(W_i) &= P(X_1 = 1, X_2 = 1, \ldots, X_{i-1} = 0, X_i \geq 2) \\
&= P((\cap_{j=1}^{i-1}\{X_j = 1\}) \cap \{X_i \geq 2\}) \\
&\overset{ind.}{=} (\prod_{j=1}^{i-1} P(X_j = 1))P(X_i \geq 2) \\
&\overset{ident.}{=} (P(X_1 = 1))^{i-1} P(X_1 \geq 2) \\
&\approx (\exp{(\frac{-5}{3})}\frac{5}{3})^{i-1}(.496)
\end{aligned}
$$

Which means that

$$
\begin{aligned}
P(W) &= \sum_{i=1}^{\infty}(\exp{(\frac{-5}{3})}\frac{5}{3})^{i-1}(.496) \\
&= \sum_{k=0}^{\infty}(\exp{(\frac{-5}{3})}\frac{5}{3})^{k}(.496) \\
&\overset{\text{geometric series}}{=} \frac{.496}{1-\exp{(\frac{-5}{3})}\frac{5}{3}} \\
&\approx \frac{.496}{.685} \\
&= .724
\end{aligned}
$$

Where the second step followed by setting $k = i - 1$. $1 \leq i \leq \infty \implies 0 \leq k \leq \infty$. We are able to use the geometric in this form since $0 < (\exp{(\frac{-5}{3})}\frac{5}{3}) < 1$. We wouldn't need a calculator to check this since it is in fact equal to $P(X = 1)$ of a Poisson random variable which must be bounded between 0 and 1.

Heartwarming that the probability loving telephone operator is likely going to best their boss and go home early.

**The remaining problems will not have solutions posted, but feel free to talk to me about them in office hours or over email**

f) Taking calls is tiring, the more calls the operator takes the more tired they get. Suppose $U : \mathcal{R}^+ \to [0, 1]$ is a function which maps the number of calls in a particular minute $X$ to the interval $[0, 1]$ which represents how rested or tired the operator is (1 being rested, 0 being exhausted) due to that minute of calls.

Let $U(x) = e^{-\alpha x}, \alpha > 0$.

What is the expected level of rest in a minute, i.e $E[U(X)]$?

**Solution**

$$E[U(X)] = \sum_{x=0}^{\infty} \frac{\exp(-\alpha x)\exp(-\lambda)\lambda^x}{x!}$$

This is a strange looking infinite sum. The secret here is we want to rewrite the above term in terms of some poisson pmf because we know all valid pmfs must sum to 1.

$$\begin{aligned}
E[U(X)] &= \sum_{x=0}^{\infty} \frac{\exp(-\alpha x)\exp(-\lambda)\lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} \frac{\exp(-\alpha)^x \lambda^x \exp(-\lambda)}{x!} \\
&= (\exp(-\lambda)) \sum_{x=0}^{\infty} \frac{(\exp(-\alpha)\lambda)^x}{x!}
\end{aligned}$$

Now the part inside the sum looks almost like a poisson. It's some number $(\exp(-\alpha)\lambda) > 0$ to the power of $x$. Notice that since this is a positive number we could call this term $\lambda'$, which is the parameter for some poisson distribution. We want the infinite sum part to go away and we know that all poisson pmfs must sum to 1. Therefore we will multiply by $1 = \frac{\exp(-\lambda')}{\exp(-\lambda')}$ and rearrange appropritately.

$$\begin{aligned}
(\exp(-\lambda)) \sum_{x=0}^{\infty} \frac{(\exp(-\alpha)\lambda)^x}{x!} &= \frac{\exp(-\lambda')}{\exp(-\lambda')}(\exp(-\lambda)) \sum_{x=0}^{\infty} \frac{(\exp(-\alpha)\lambda)^x}{x!} \\
&= \frac{\exp(-\lambda')}{\exp(-\lambda')}(\exp(-\lambda)) \sum_{x=0}^{\infty} \frac{(\lambda')^x}{x!} \\
&= \frac{\exp(-\lambda)}{\exp(-\lambda\prime)} \sum_{x=0}^{\infty} \frac{(\lambda')^x \exp(-\lambda')}{x!}
\end{aligned}$$

Now the sum goes away since $\sum_{x=0}^{\infty} \frac{(\lambda')^x \exp(-\lambda')}{x!} = 1$ since it is a sum over the entire pmf of a valid poisson with parameter $\lambda'$.

$$\begin{aligned}
\frac{\exp(-\lambda)}{\exp(-\lambda\prime)} \sum_{x=0}^{\infty} \frac{(\lambda')^x \exp(-\lambda')}{x!} &= \frac{\exp(-\lambda)}{\exp(-\lambda\prime)}(1) \\
&= \exp(-\lambda)\exp(\lambda') \\
&= \exp(-\lambda)\exp(\exp(-\alpha)\lambda)) \\
&= \exp(-\lambda + \exp(-\alpha)\lambda) \\
&= \exp(\lambda(\exp(-\alpha) - 1))
\end{aligned}$$

This question is intimately related to the problem of finding the Moment generating function for the Poisson distribution. If you can solve this problem, you can find the MGF of a poisson and vice versa. Moment Generating functions are not yet covered but will be seen in class in the next few weeks.

g) Suppose the operator has a rule that determines how they take breaks. They take a break as soon as they have had 10 tiring minutes, that is 10 minutes where $U(X) < .5$.

i) What is the probability of a given minute being tiring according to the above definition when $\alpha = \frac{1}{2}$.

ii) What is the probability that the break occurs exactly on the 20th minute?

iii) What is the probability that it has been 30 minutes and they still haven't taken a break?
**Solution**

i)

$$\begin{aligned}
P(U(X) \leq 0.5) &= P(\exp(-\frac{1}{2}X) < 0.5) \\
&= P(-\frac{1}{2}X < log(0.5)) \\
&= P(X > -\frac{log(0.5)}{\frac{1}{2}}) \\
&= P(X > 1.38...) \\
&= P(X \geq 2) \\
&= 0.496
\end{aligned}$$

Where the last line follows from the work in part c).

ii) What must happen to take a break on exactly the 20th minute? It must be that up to 19 minutes there must have been exactly 9 tiring minutes and then on the 20th minute, it must be a tiring minute.

Let $T_i :=$ "The ith minute is tiring"

$$\begin{aligned}
P(\text{ Break on 20th minute }) &= P(\text{ 9 of first 19 tired} \cap T_{20}) \\
&\overset{ind.}{=} P(\text{ 9 of first 19 tired})P(T_{20})
\end{aligned}$$

By the fact that being tired in each minute is independently and identically distributed, any sequence with exactly 9 tired minutes and 10 non-tired minutes would have the following probability:

$$\begin{aligned}
P(T_1, T_2, \ldots, T_9, T_{10}^c, T_{11}^c, \ldots, T_{19}^c) &= (P(T_1))^9(1 - P(T_1))^{10} \\
&= (0.496)^9(1 - 0.496)^{10}
\end{aligned}$$

There are $\binom{19}{10}$ such sequences, thus $P($ 9 of first 19 tired$) = \binom{19}{10}(0.496)^9(1-0.496)^{10}$. And then all together:

$$\begin{aligned}
P(\text{ Break on 20th minute }) &= \binom{19}{10}(0.496)^9(1-0.496)^{10}(0.496) \\
&= \binom{19}{10}(0.496)^{10}(1-0.496)^{10}
\end{aligned}$$

Notice that this is a negative binomial distribution.

iii) Let $Y :=$ "The minute they take a first break"

Notice that since it takes 10 tiring the minutes to take a break, so the values that $Y$ can take with positive probability are $\{10, 11, \ldots, \ldots\}$

Using this negative binomial random variable we defined above we can answer our question:

$$\begin{aligned}
P(\text{Haven't taken a break after 30 minutes}) &= P(Y \geq 31) \\
&= 1 - P(Y \leq 30) \\
&= \sum_{y=10}^{30} \binom{y-1}{9} p^{10}(1-p)^{y-10}
\end{aligned}$$

Where $p = 0.496$