STAT2003 Assignment 1

Ryan White s4499039

March 29, 2022

Question 1

a. Choose:

- $A =$ event that a 5 is given
- $B =$ event that a number ≥ 4 is given

Naturally, $\mathbb{P}(A) = 1/6$ and $\mathbb{P}(B) = |B|/|\Omega| = 3/6 = 1/2$. And so,

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{1/6}{1/2} \\ &= 2/6 = 1/3\end{aligned}$$

Therefore, $\mathbb{P}(A|B) > \mathbb{P}(A)$ since $1/3 > 1/6$.

b. Choose:

- $A =$ event that the given number is even $\Rightarrow A = \{2, 4, 6\} \Rightarrow \mathbb{P}(A) = 3/6 = 1/2$
- $B =$ event that the number is a factor of 6 $\Rightarrow B = \{1, 2, 3, 6\} \Rightarrow \mathbb{P}(B) = 4/6 = 2/3$

Also note that $\mathbb{P}(A \cap B) = 2/6 = 1/3$. Then,

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{1/3}{2/3} \\ &= 3/6 = 1/2 = \mathbb{P}(A)\end{aligned}$$

And so $\mathbb{P}(A|B) = \mathbb{P}(A)$.

c. Choose:

- $A =$ event that the given number is $\geq 4 \Rightarrow \mathbb{P}(A) = 3/6 = 1/2$
- $B =$ event that the given number is prime $\Rightarrow B = \{2, 3, 5\} \Rightarrow \mathbb{P}(B) = 3/6 = 1/2$

Note that $\mathbb{P}(A \cap B) = 1/6$ since they share only one comment element (5). And so

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \\ &= \frac{1/6}{1/2} \\ &= 2/6 = 1/3\end{aligned}$$

And so $\mathbb{P}(A|B) < \mathbb{P}(A)$ since $1/3 < 1/2$.

Question 2

a. Let:

- M = event that a bird has feather mites $\Rightarrow \mathbb{P}(M) = 0.2$
- L = event that a bird tests positive for feather mites $\Rightarrow \mathbb{P}(L|M) = 0.9$
- K = event that a bird tests negative for feather mites

The above assumes that a bird cannot test positive for feather mites if the condition is present (no false positives). And so,

$$\begin{aligned}\mathbb{P}(K) &= 1 - \mathbb{P}(L) \\ &= 1 - \mathbb{P}(M)\mathbb{P}(L|M) \\ &= 1 - 0.2 \cdot 0.9 \\ &= 0.82\end{aligned}$$

And so approximately 82% of sparrows tested will test negative.

b. We now want to calculate the probability of a false negative, $\mathbb{P}(M|K)$:

$$\begin{aligned}\mathbb{P}(M|K) &= \frac{\mathbb{P}(M \cap K)}{\mathbb{P}(K)} \\ &= \frac{\mathbb{P}(K|M)\mathbb{P}(M)}{\mathbb{P}(K)} \\ &= \frac{\frac{\mathbb{P}(K \cap M)}{\mathbb{P}(M)} \cdot \mathbb{P}(M)}{\mathbb{P}(K)} \\ &= \frac{\mathbb{P}(K) + \mathbb{P}(M) - \mathbb{P}(K \cup M)}{\mathbb{P}(K)} \\ &= 1 + \frac{0.2}{0.82} - \frac{1}{0.82} \\ \mathbb{P}(M|K) &\simeq 0.0244\end{aligned}$$

where the above $\mathbb{P}(K \cup M) = 1$ as those two conditions encapsulate the whole population. As shown above, there is about a 0.0244 probability of a false negative.

Question 3

Since each consecutive 6 is independent of the previous, there is a $1/6^{10}$ chance of a string of 10 consecutive 6's. However, a particular string of 10 numbers cannot occur in the last 9 positions of some sample space of random numbers, and so the *valid* sample space in this context is $6^8 - 9$. As such, the probability of getting 10 consecutive (and equilikely) 6's in the valid sample space is

$$\begin{aligned}\mathbb{P}(10 \text{ Consecutive 6's}) &= \frac{6^8 - 9}{6^{10}} \\ &\approx 0.027776 < 1/36 = 0.0277777\ldots\end{aligned}$$

Of course, there may be some string of 6's longer than 10 but these are subsets of the 10 string length case and so don't contribute. As such, it was shown that the probability of there being a sequence of at least 10 consecutive in a sample space of length 6^8 was less than $1/36$.

Question 4

a. Since the chance of b_i being the largest element of the set b so far is equally likely for all values of $\{1, \ldots, i\}$, according to the equilikely principle the probability that some b_r (for values of $2 \leq r \leq n$) is greater than all of the values before it is

$$\mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\}) = 1/r$$

Intuitively, take $r = n$ in which case the proof is analogous to that of the equilikely principle.

b. Now, choose:

- $A =$ event that $b_r = n \Rightarrow \mathbb{P}(A) = 1/n$ since there are n elements in the set and each of them equally likely.
- $B =$ event that $b_r > \max\{b_1, \ldots, b_{r+1}\} \Rightarrow \mathbb{P}(B) = 1/r$ as in part a. of the question.

Of course, $\mathbb{P}(B|A) = 1$ since $b_r = n$ is the largest element in the set and so $(b_r > \max\{b_1, \ldots, b_{r+1}\}|\, b_r = n)$ is a certainty.
Then,

$$\begin{aligned}
\mathbb{P}(b_r = n|b_r > \max\{b_1, \ldots, b_{r-1}\}) &= \frac{\mathbb{P}(b_r = n \cap b_r > \max\{b_1, \ldots, b_{r-1}\})}{\mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\})} \\
&= \frac{\mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\}|b_r = n) \cdot \mathbb{P}(b_r = n)}{\mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\})} \\
&= \frac{1 \cdot 1/n}{1/r} \\
&= \frac{r}{n}
\end{aligned}$$

And so the probability that $b_r = n$, given that b_r is greater than the elements in the set before it, is equal to r/n.

c. If $b_r < \max\{b_1, \ldots, b_{r-1}\}$, then $b_r^* = b_{r+1}^*$, since the set b^* takes the minimum value of the remaining elements in the set of C, which is comprised of the ***ordered*** maximal values of the set b, stored as index of the set. As such, for some $r \geq 2$, if $b_r < \max\{b_1, \ldots, b_{r+1}\}$ then C_r and b_r^* are the smallest maximum values remaining in the sets. Since $b_r < \max\{b_1, \ldots, b_{r+1}\}$, even if $b_{r+1} > \max\{b_1, \ldots, b_{r+1}\}$, b_r^* will have the same value as b_{r+1}^*. Therefore,

$$\mathbb{P}(b_r^* = n|b_r < \max\{b_1, \ldots, b_{r-1}\}) = \mathbb{P}\left(b_{r+1}^* = n\right)$$

d. Notice that, in the event that $b_r > \max\{b_1, \ldots, b_{r-1}\}$, b_r is the next largest element, and so $b_r^* = b_r$. As such,

$$\mathbb{P}(b_r^* = n|b_r > \max\{b_1, \ldots, b_{r-1}\}) = \mathbb{P}(b_r = n|b_r < \max\{b_1, \ldots, b_{r-1}\}) = \frac{r}{n}$$

Therefore, using the law of total probability,

$$\begin{aligned}
\mathbb{P}(b_r^* = n) &= \mathbb{P}(b_r^* = n|b_r > \max\{b_1, \ldots, b_{r-1}\}) \cdot \mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\}) \\
&\quad + \mathbb{P}(b_r^* = n|b_r < \max\{b_1, \ldots, b_{r-1}\})\,(1 - \mathbb{P}(b_r > \max\{b_1, \ldots, b_{r-1}\})) \\
&= \frac{r}{n} \cdot \frac{1}{r} + \mathbb{P}\left(b_{r+1}^* = n\right)\left(1 - \frac{1}{r}\right) \\
&= \frac{1}{n} + \left(1 - \frac{1}{r}\right)\mathbb{P}\left(b_{r+1}^* = n\right)
\end{aligned}$$

Question 5

a. i.

1. We require that $ae^x \leq b + ce^{-3x} \quad \forall x \in \mathbb{R}$ (as per the third property of distribution functions).
2. Due to the first three properties, we require that both ae^x and $b + ce^{-3x}$ be between 0 and 1 inclusive for all values of x. Therefore, $0 \leq a \leq 1$ since, as x approaches 0 from below, e^x approaches 1. Also, *at* $x = 0$, $b + ce^{-3x} = b + c$ and so $0 \leq b + c \leq 1$.
3. Since the distribution function must be increasing continuously towards the right, c must be negative. This is because, for increases x, e^{-3x} gets smaller and so ce^{-3x} gets smaller.
4. Closely entwined with the previous point, b must equal 1. Since the function is always approaching $F_X(x) = 1$ as $x \to \infty$, $b = 1$ because as $x \to \infty$, $ce^{-3x} \to 0$ and so $b + ce^{-3x} \to 1 = b$. Because of this, this also means that $-1 \leq c \leq 0$.

ii. Provided that X is a continuous random variable, this adds the constraint that, at $x = 0$, $ae^x = b + ce^{-3x}$ for continuity. Therefore, $a = b + c$.

b. i. The probability that $X \leq 0$ is calculated by:

$$\begin{aligned}\mathbb{P}(X \leq 0) &= \int_{-\infty}^{0} ae^x \, dx \\ &= \int_{-\infty}^{0} \frac{1}{3}e^x \, dx \\ &= \left[\frac{1}{3}e^x\right]_{-\infty}^{0} \\ &= \frac{1}{3} - 0 = \frac{1}{3}\end{aligned}$$

ii. The piecewise distribution function is represented as

$$F_X(x) = \begin{cases} 1/3e^x & x < 0 \\ 1 - 2/3e^{-3x} & x \geq 0 \end{cases}$$

Since the probability density function of X is given by: $f_X(x) = {F_X}'(x)$, the negative x-axis pdf is

$$\begin{aligned}f_X(x) &= \frac{d}{dx}\left(\frac{1}{3}e^x\right) \\ &= \frac{1}{3}e^x\end{aligned}$$

Similarly, the piecewise function for the positive x-axis is

$$\begin{aligned}f_X(x) &= \frac{d}{dx}\left(1 - \frac{2}{3}e^{-3x}\right) \\ &= 2e^{-3x}\end{aligned}$$

As such, the piecewise probability density function is

$$f_X(x) = \begin{cases} 1/3e^x & x < 0 \\ 2e^{-3x} & x \geq 0 \end{cases}$$

iii. The moment generating function is given by

$$\begin{aligned}M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f_X(x) \, dx \\ &= \int_{-\infty}^{0} e^{tx} \cdot \frac{1}{3}e^x \, dx + \int_{0}^{\infty} e^{tx} \cdot 2e^{-3x} \, dx \\ &= \frac{1}{3}\int_{-\infty}^{0} e^{(t+1)x} dx + 2\int_{0}^{\infty} e^{(t-3)x} dx \\ &= \frac{1}{3}\left[\frac{e^{(t+1)x}}{t+1}\right]_{-\infty}^{0} + 2\left[\frac{e^{(t-3)x}}{t-3}\right]_{0}^{\infty} \\ &= \frac{1}{3}\left(\frac{1}{t+1} - 0\right) + 2\left(0 - \frac{1}{t-3}\right) \quad [\forall t < 3] \\ &= \frac{1}{3(t+1)} - \frac{2}{t-3}\end{aligned}$$

where this solution is valid for t in the range $-1 < t < 3$, where the upper bound of 3 is from the 2nd term integral and the lower bound is from the 1st term denominator.

iv. To calculate the expectation value and the variance, first the first and second derivatives of the moment generating function must be calculated:

$$
\begin{aligned}
M_X{}'(t) &= \frac{d}{dt}\left(\frac{1}{3(t+1)} - \frac{2}{t-3}\right) \\
&= \frac{2}{(t-3)^2} - \frac{1}{3(t+1)^2} \\
M_X{}''(t) &= \frac{d}{dt}\left(\frac{2}{(t-3)^2} - \frac{1}{3(t+1)^2}\right) \\
&= \frac{2}{3(t+1)^3} - \frac{4}{(t-3)^3}
\end{aligned}
$$

And so,

$$
\begin{aligned}
\mathbb{E}\left(X\right) &= M_X{}'(0) \\
&= \frac{2}{9} - \frac{1}{3} \\
&= -\frac{1}{9}
\end{aligned}
$$

And,

$$
\begin{aligned}
M_X{}''(t) &= \frac{2}{3} - \frac{4}{-27} \\
&= \frac{18}{27} + \frac{4}{27} \\
&= \frac{22}{27} \\
\Rightarrow \operatorname{Var}\left(X\right) &= M_X{}''(0) - (M_X{}'(0))^2 \\
&= \frac{22}{27} - \left(-\frac{1}{9}\right)^2 \\
&= \frac{66}{81} - \frac{1}{81} = \frac{65}{81}
\end{aligned}
$$

Therefore, the function has an expectation value of $\mathbb{E}\left(X\right) = -1/9$ with a variance $\operatorname{Var}\left(X\right) = 65/81$.