40.017 PROBABILITY & STATISTICS

Homework 2

Michael Hoon
1006617
Section 2

February 20, 2024

Question 1

Y has the following pmf:

$$\mathbb{P}(Y=-1)=0.4, \quad \mathbb{P}(Y=0)=0.25, \quad \mathbb{P}(Y=1)=0.35$$

In order to simulate Y, we use a basic conditional structure in Python:

Question 2

We have the individual probabilities $\mathbb{P}(E_1) = \frac{6}{36} = \frac{1}{6}$, $\mathbb{P}(E_2) = \frac{1}{6}$, and $\mathbb{P}(E_3) = \frac{1}{6}$, and we need to show that E_1 and E_2 are not conditionally independent given E_3. Starting with the definition in Week 3 Class 2, we have the result:

$$\mathbb{P}(B_2|B_1 \cap A) = \mathbb{P}(B_2|A) \tag{1}$$

In the context of the problem, using Equation 1:

$$\begin{aligned}\mathbb{P}(E_2|E_1 \cap E_3) &= \mathbb{P}(\text{'first dice shows 4'}|\text{'sum of two dice is 7' and 'second dice shows 3'}) \\ &= 1\end{aligned}$$

whereas

$$\mathbb{P}(E_2|E_3) = \frac{1}{6}$$

since they are not independent. As we can see,

$$\mathbb{P}(E_2|E_1 \cap E_3) \neq \mathbb{P}(E_2|E_3)$$

and thus we conclude that E_1 and E_2 are not conditionally independent given E_3.

Question 3

(a)

Since we have an unloaded dice, the probability of each dice roll $\mathbb{P}(X = i)$, $\forall i \in \{1, 2, 3, 4\}$ is $\frac{1}{4}$. Let x be the result of the first dice roll, and y be the result of the second dice roll. There are three cases for this event, namely when $x = y$, $x > y$ and otherwise. In the case where $x = y$, $\mathbb{P}(\max\ = x \text{ and } \min\ = x) = \mathbb{P}(\{(x, x)\}) = (1/4)^2$. However, if $x > y$, we have $\mathbb{P}(\max\ = x \text{ and } \min\ = y) = \mathbb{P}(\{(x, y)\}, \{(y, x)\}) = 2 \times (1/4)^2$. Otherwise, it is not possible. The piecewise function for the joint pmf is thus:

$$f(x, y) = \begin{cases} \frac{1}{16}, & \forall x = y \\ \frac{1}{8}, & \forall x > y \\ 0, & \text{otherwise} \end{cases}$$

We now construct the joint pmf table of X and Y in table 1 below.

x \ y	1	2	3	4	f_X
1	$\frac{1}{16}$	0	0	0	$\frac{1}{16}$
2	$\frac{1}{8}$	$\frac{1}{16}$	0	0	$\frac{3}{16}$
3	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	0	$\frac{5}{16}$
4	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{7}{16}$
f_Y	$\frac{7}{16}$	$\frac{5}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	1

Table 1: Joint pmf of X and Y

(b)

The conditional pmf of X given that $Y = 2$ is:

$$f_{X|Y}(x|2) = \mathbb{P}((X = x)(Y = 2)) = \frac{f(x, 2)}{f_Y(2)}$$

where $f_Y(2) = \frac{5}{16}$ from table 1. $f(x, 2)$ is given by the 2nd column of table 1, thus we have the conditional pmf table:

x	1	2	3	4
$f_{X\|Y}(x\|2)$	0	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{2}{5}$

Table 2: Conditional pmf of $X|Y = 2$

Question 4

(a)

Let $Z \sim \mathcal{N}(0,1)$, and $X = Z^2$. From MU Week 8,

$$\begin{aligned}
F_X(x) &= \mathbb{P}(X \leq x) \\
&= \mathbb{P}(Z^2 \leq x) \\
&= \mathbb{P}(-\sqrt{x} \leq Z \leq \sqrt{x}) \\
&= \Phi(\sqrt{x}) - \Phi(-\sqrt{x}) \\
f_X(x) &= \frac{1}{2\sqrt{x}}\phi(\sqrt{x}) + \frac{1}{\sqrt{x}}\phi(-\sqrt{x}) \quad \text{(differentiating both sides)} \\
&= \frac{1}{2\sqrt{x}}\left(\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x}{2}\right)\right) + \frac{1}{2\sqrt{x}}\left(\frac{1}{2\sqrt{\pi}}\exp\left(-\frac{x}{2}\right)\right) \\
&= \frac{1}{\sqrt{2\pi x}}\exp\left(-\frac{x}{2}\right)
\end{aligned}$$

Thus the pdf of X is given by

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}}\exp\left(-\frac{x}{2}\right), & x > 0 \\ 0, & x \leq 0 \end{cases} \tag{2}$$

which is interestingly also the pdf of a χ^2 distribution with 1 degree of freedom (defined as 0 at $x = 0$).

The standard Gamma distribution is given by

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{3}$$

where $X \sim \text{gamma}(\alpha, \lambda)$. To show that X is a gamma random variable, we must transform equation 2 into the form in equation 3. By inspection, it is immediately obvious that $\lambda = \frac{1}{2}$. From equation 2,

$$\begin{aligned}
f_X(x) &= \frac{1}{\sqrt{2\pi x}}\exp(-\frac{x}{2}) \\
&= \frac{\left(\frac{1}{2}\right)^{\frac{1}{2}}}{\sqrt{\pi}}x^{\frac{1}{2}-1}e^{-\frac{1}{2}x}
\end{aligned}$$

where $\alpha = \frac{1}{2}$ by inspection. To confirm that the value of $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, we can find the integral:

$$\begin{aligned}
\Gamma\left(\frac{1}{2}\right) &= \int_0^\infty x^{-\frac{1}{2}} e^{-x}\,\mathrm{d}x \\
&= \int_0^\infty \left(\frac{t^2}{2}\right)^{-\frac{1}{2}} e^{-\frac{t^2}{2}} \frac{\mathrm{d}x}{\mathrm{d}t}\,\mathrm{d}t \\
&= \int_0^\infty \frac{\sqrt{2}}{t} e^{-\frac{t^2}{2}} t\,\mathrm{d}t \\
&= 2\sqrt{\pi} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,\mathrm{d}t \\
&= \frac{1}{2}(2\sqrt{\pi}) \\
&= \sqrt{\pi}
\end{aligned}$$

With this, we can conclude that $\boxed{X \sim \text{gamma}\left(\frac{1}{2}, \frac{1}{2}\right)}$.

(b)

If we have two independent standard normal r.v.s Z_1 and Z_2, the sum of squares is given by $Z_1^2 + Z_2^2 = X_1 + X_2$. In Week 3 Lecture 1, we proved the MGF of the Gamma distribution:

$$\begin{aligned}
M_X(t) &= \mathbb{E}(e^{tX}) \\
&= \dots \\
&= \frac{\lambda^\alpha (\lambda - t)^{-\alpha})}{\Gamma(a)} \underbrace{\int_0^\infty y^{\alpha-1} e^{-y}\,\mathrm{d}y}_{\Gamma(\alpha)} \\
&= \lambda^\alpha (\lambda - t)^{-\alpha}
\end{aligned}$$

Using the property for the MGF of the sum of two independent random variables (where $X \sim \text{gamma}(1/2, 1/2)$),

$$\begin{aligned}
M_{X_1+X_2}(t) &= M_{X_1}(t) M_{X_2}(t) \\
&= \left(\frac{1}{2}\right)^{1/2} \left(\frac{1}{2} - t\right)^{-1/2} \left(\frac{1}{2}\right)^{1/2} \left(\frac{1}{2} - t\right)^{-1/2} \\
&= \frac{1/2}{1/2 - t}
\end{aligned}$$

The MGF of an exponential random variable is given by

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

By inspection, we can see that this is actually the MGF of an exponential r.v. with $\lambda = \frac{1}{2}$, thus $X_1 + X_2 \sim \text{exponential}(\frac{1}{2})$.

Question 5

Let X be the number of tosses required to reach our goal of 3 consecutive $\texttt{H}$'s appearing for the first time. We start from the hint, where there are 4 cases for any sequence to start: $\texttt{T}$, $\texttt{HT}$, $\texttt{HHT}$, $\texttt{HHH}$. Let $Y = i$ for each of the 4 cases accordingly, and $\mathbb{P}(\texttt{H}) = \frac{1}{2}$ (since we have a fair coin), so that:

$$\begin{aligned}
\mathbb{P}(Y = 1) &= \mathbb{P}(\texttt{T}) = \frac{1}{2} \\
\mathbb{P}(Y = 2) &= \mathbb{P}(\texttt{HT}) = \left(\frac{1}{2}\right)^2 \\
\mathbb{P}(Y = 3) &= \mathbb{P}(\texttt{HHT}) = \left(\frac{1}{2}\right)^3 \\
\mathbb{P}(Y = 4) &= \mathbb{P}(\texttt{HHH}) = \left(\frac{1}{2}\right)^3
\end{aligned}$$

Applying the law of total expectation, we have:

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{\text{all } y} \mathbb{E}(X|Y = y)\mathbb{P}(Y = y) \\
&= \left(\frac{1}{2}\right)\mathbb{E}(X|Y = 1) + \left(\frac{1}{2}\right)^2 \mathbb{E}(X|Y = 2) + \left(\frac{1}{2}\right)^3 \mathbb{E}(X|Y = 3) + \left(\frac{1}{2}\right)^3 \mathbb{E}(X|Y = 4)
\end{aligned}$$

To evaluate the conditional expectations, notice that each case involves recursion, where for $Y = 1$, tossing a $\texttt{T}$ at first yields no contribution to our goal of 3 consecutive $\texttt{H}$'s, so we are now back to the beginning with $1 + \mathbb{E}(X)$. For $\mathbb{E}(X|Y = 2)$ and $\mathbb{E}(X|Y = 3)$, a similar reasoning holds and we are back to the beginning after 2 and 3 'useless' tosses. For $\mathbb{E}(X|Y = 4)$ however, we have the case for the least number of tosses required to reach our goal, which is 3 tosses. Thus, the expected number of tosses $\mathbb{E}(X|Y = 4)$ is trivially 3. We now have:

$$\begin{aligned}
\mathbb{E}(X) &= \left(\frac{1}{2}\right)[1 + \mathbb{E}(X)] + \left(\frac{1}{2}\right)^2 [2 + \mathbb{E}(X)] + \left(\frac{1}{2}\right)^3 [3 + \mathbb{E}(X)] + \left(\frac{1}{2}\right)^3 (3) \\
\mathbb{E}(X) &= \frac{7}{4} + \frac{7}{8}\mathbb{E}(X) \\
\mathbb{E}(X) &= \boxed{14}
\end{aligned}$$

Question 6

(a)

Using the hint, let X be the amount paid, while Y be the number of claims filed. Since there can only be at most 1 claim filed, $Y = 1$ or $Y = 0$, and $\mathbb{P}(Y = 1) = 0.2$, $\mathbb{P}(Y = 0) = 1 - 0.2 = 0.8$. If a claim is filed, then the mean amount paid is \$1000: $\mathbb{E}(X|Y = 1) = 1000$, and if no claim is filed, then nothing is paid: $\mathbb{E}(X|Y = 0) = 0$. Using the alternate formulation of the law of total expectation,

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(\mathbb{E}(X|Y)) \\
&= \sum_{i=0}^{1} \mathbb{E}(X|Y = i)\mathbb{P}(Y = i) \\
&= \mathbb{E}(X|Y = 0)\mathbb{P}(Y = 0) + \mathbb{E}(X|Y = 1)\mathbb{P}(Y = 1) \\
&= 0 \times 0.8 + 1000 \times 0.2 \\
&= \boxed{200}
\end{aligned}$$

(b)

Using the hint, we have the pmf tables for $\mathbb{E}(X|Y)$ and $\text{Var}(X|Y)$ below. Since X is exponentially distributed when $Y = 1$, for an exponential distribution with parameter λ, the mean is given by $\frac{1}{\lambda}$ and the variance is $\left(\frac{1}{\lambda}\right)^2$. Thus, $\text{Var}(X|Y) = [\mathbb{E}(X|Y)]^2 = 1000^2$.

Y	0	1
$\mathbb{E}(X\|Y = i)$	0	1000
$\mathbb{P}(Y = i)$	0.8	0.2

Table 3: pmf table of $\mathbb{E}(X|Y)$

Y	0	1
$\text{Var}(X\|Y = i)$	0	1000000
$\mathbb{P}(Y = i)$	0.8	0.2

Table 4: pmf table of $\text{Var}(X|Y)$

Now to find $\text{Var}(X)$, we use the law of total variance:

$$\begin{aligned}
\text{Var}(X) &= \mathbb{E}(\text{Var}(X|Y)) + \text{Var}(\mathbb{E}(X|Y)) \\
&= \mathbb{E}(\text{Var}(X|Y)) + \mathbb{E}(\mathbb{E}(X|Y)^2) - (\mathbb{E}(X))^2 \\
&= \mathbb{E}(\text{Var}(X|Y)) + \mathbb{E}(\text{Var}(X|Y)) - (\mathbb{E}(X))^2 \\
&= 2 \cdot \mathbb{E}(\text{Var}(X|Y)) - (\mathbb{E}(X))^2 \\
&= 2\,(0 \times 0.8 + 1000000 \times 0.2) - 200^2 \\
&= \boxed{360000}
\end{aligned}$$

Question 7

(a)

We first compute the $\mathbb{E}$ and Var of $2X + 3Y$ (linear combination of X and Y):

$$\begin{aligned}\mathbb{E}(2X+3Y) &= 2\mathbb{E}(X) + 3\mathbb{E}(Y) \\ &= 2\mu_X + 3\mu_Y \\ &= 2(2) + 3(-1) \\ &= 1\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(2X+3Y) &= \mathrm{Var}(2X) + \mathrm{Var}(3Y) + 2\cdot\mathrm{Cov}(2X, 3Y) \\ &= 4\cdot\mathrm{Var}(X) + 9\cdot\mathrm{Var}(Y) + 2\cdot(6)\cdot\mathrm{Cov}(X,Y) \\ &= 4 + 36 + 12\left(\rho\sqrt{\sigma_X^2\sigma_Y^2}\right) \\ &= 40 + 12(-\frac{1}{2})\left(\sqrt{1(4)}\right) \\ &= 28\end{aligned}$$

Thus, we have the normally distributed $2X+3Y$: $2X + 3Y \sim \mathcal{N}(1, 28)$. To obtain $\mathbb{P}(2X + 3Y \geq 6)$, we use a standard normal Z:

$$\begin{aligned}\mathbb{P}(2X+3Y \geq 6) &= \mathbb{P}\left(Z \geq \frac{6-1}{\sqrt{28}}\right) \\ &= 0.17235 \\ &\boxed{\approx 0.172}\end{aligned}$$

(b)

We first find the conditional pdf of $X|Y = 2$ using the joint pdf of the bivariate normal:

$$\begin{aligned}
f_{X|Y}(X|Y=2) &= \frac{f(X, Y=2)}{f_Y(Y=2)} \\
&= c_1 f(X, 2) \\
&= c_2 \exp\left[-\frac{1}{2(1-(\frac{1}{2})^2)}\left(\left(\frac{x-2}{1}\right)^2 - 2\left(-\frac{1}{2}\right)\frac{x-2}{1}\frac{2-(-1)}{2} + \left(\frac{2-(-1)}{2}\right)^2\right)\right] \\
&= c_2 \exp\left[-\frac{2}{3}\left((x-2)^2 + \frac{3}{2}(x-2) + \frac{9}{4}\right)\right] \\
&= c_2 \exp\left[-\frac{2}{3}\left(x^2 - 4x + 4 + \frac{3}{2}x - 3 + \frac{9}{4}\right)\right] \\
&= c_2 \exp\left[-\frac{2}{3}(x^2 - \frac{5}{2}x + \frac{13}{4})\right] \\
&= c_2 \exp\left[-\frac{2}{3}\left(\left(x - \frac{5}{4}\right)^2 + \frac{13}{4} - \left(\frac{5}{4}\right)^2\right)\right] \qquad \text{(Completing the square)} \\
&= c_3 \exp\left[-\frac{2}{3}(x - \frac{5}{4})^2\right] \qquad \text{(Constant absorbed to } c_3\text{)} \\
&= c_3 \exp\left[-\frac{(x-\frac{5}{4})^2}{2(\frac{3}{4})}\right]
\end{aligned}$$

Comparing the result with the pdf of a univariate normal, we can see that

$$\mathbb{E}(X|Y=2) = \frac{5}{4}, \quad \mathrm{Var}(X|Y=2) = \frac{3}{4}$$

where $c_3 \in \mathbb{R}$. Thus, to compute $\mathbb{P}(X < 1|Y = 2)$, we use a standard normal Z:

$$\begin{aligned}
\mathbb{P}(X < 1|Y=2) &= \mathbb{P}\left(Z < \frac{1-5/4}{\sqrt{3/4}}\right) \\
&= 0.386414 \\
&\boxed{\approx 0.386}
\end{aligned}$$