

1.

(a)

We can find c by starting with the fact that the total integral of the density must equal 1, as it represents a probability, i.e.:

$$\int_0^1 \int_0^x cxy dy dx = 1$$

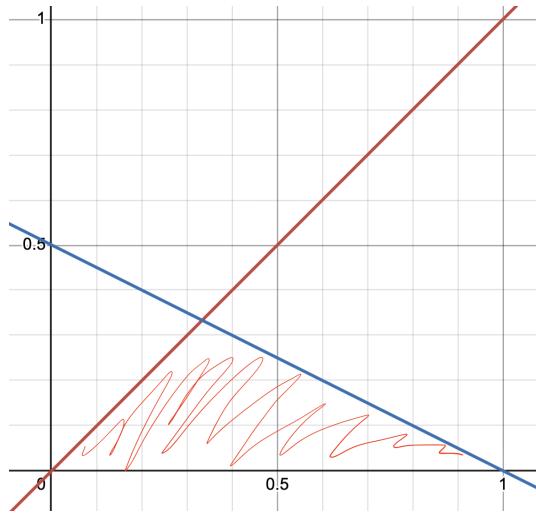
And we can integrate and solve from here:

$$\begin{aligned} \int_0^1 cx \cdot \frac{y^2}{2} \Big|_0^x dx &= 1 \\ \int_0^1 cx \cdot \left(\frac{x^2}{2} - 0\right) dx &= 1 \\ \int_0^1 \frac{cx^3}{2} dx &= 1 \\ \frac{c}{2} \cdot \frac{x^4}{4} \Big|_0^1 &= 1 \\ \frac{c}{2} \cdot \left(\frac{1}{4} - 0\right) &= 1 \\ \frac{c}{8} &= 1 \\ \boxed{c = 8} \end{aligned}$$

(b)

To determine $P(X + 2Y \leq 1)$, we intersect the line of $y = -\frac{x}{2} + \frac{1}{2}$ with $y = x$ from the original problem statement:

$$\begin{aligned} x &= -\frac{x}{2} + \frac{1}{2} \\ \frac{3x}{2} &= \frac{1}{2} \\ x &= \frac{1}{3} \end{aligned}$$



So we have regions from $0 \leq x \leq \frac{1}{3}$ and $\frac{1}{3} \leq x \leq 1$. For the first region, we see that $y = x$ is the binding equation, while we see that for the second region, $y = \frac{-x}{2} + \frac{1}{2}$ is the binding equation. Let's evaluate accordingly. Region 1:

$$\begin{aligned}
& \int_0^{\frac{1}{3}} \int_0^x 8xy dy dx \\
&= \int_0^{\frac{1}{3}} 8x \cdot \frac{y^2}{2} \Big|_0^x dx \\
&= \int_0^{\frac{1}{3}} 8x \cdot \left(\frac{x^2}{2} - 0\right) dx = \int_0^{\frac{1}{3}} 4x^3 dx \\
&= x^4 \Big|_0^{\frac{1}{3}} \\
&= \frac{1}{81}
\end{aligned}$$

Region 2:

$$\begin{aligned}
& \int_{\frac{1}{3}}^1 \int_0^{\frac{-x+1}{2}} 8xy dy dx \\
&= \int_{\frac{1}{3}}^1 8x \cdot \frac{y^2}{2} \Big|_0^{\frac{-x+1}{2}} dx \\
&= \int_{\frac{1}{3}}^1 8x \cdot \frac{(-x+1)^2}{4 \cdot 2} dx = \int_{\frac{1}{3}}^1 8x \cdot \frac{x^2 + 1 - 2x}{8} dx = \int_{\frac{1}{3}}^1 x^3 + x - 2x^2 dx \\
&= \frac{x^4}{4} + \frac{x^2}{2} - \frac{2x^3}{3} \Big|_{\frac{1}{3}}^1 \\
&= \left(\frac{1}{4} + \frac{1}{2} - \frac{2}{3}\right) - \left(\frac{1}{324} + \frac{1}{18} - \frac{2}{81}\right) = \frac{1}{12} - \frac{11}{324}
\end{aligned}$$

$$= \frac{27}{324} - \frac{11}{324} = \frac{16}{324} = \frac{4}{81}$$

So in total, we have $\frac{1}{81} + \frac{4}{81} = \boxed{\frac{5}{81}}$

(c)

To find $E(X|Y = y)$, we can start with calculating the conditional density $f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$. We have $f(x,y)$, so we need to find $f_Y(y)$. We can derive this from $f(x,y)$:

$$\begin{aligned} f_Y(y) &= \int_y^1 8xy dx \\ f_Y(y) &= 8y \cdot \frac{x^2}{2} \Big|_y^1 \\ f_Y(y) &= 8y \cdot \left(\frac{1}{2} - \frac{y^2}{2}\right) \\ f_Y(y) &= 4y - 4y^3 \end{aligned}$$

And plug in accordingly:

$$\begin{aligned} f_{X|Y}(x|y) &= \frac{f(x,y)}{f_Y(y)} \\ f_{X|Y}(x|y) &= \frac{8xy}{4y - 4y^3} \\ f_{X|Y}(x|y) &= \frac{2x \cdot 4y}{4y(1 - y^2)} = \frac{2x}{1 - y^2} \end{aligned}$$

We can now finally find the expectation:

$$\begin{aligned} E(X|Y = y) &= \int_y^1 x \cdot f_{X|Y}(x|y) dx \\ E(X|Y = y) &= \int_y^1 x \cdot \frac{2x}{1 - y^2} dx \\ E(X|Y = y) &= \int_y^1 \frac{2x^2}{1 - y^2} dx \\ E(X|Y = y) &= \frac{2}{1 - y^2} \cdot \frac{x^3}{3} \Big|_y^1 \\ E(X|Y = y) &= \frac{2}{1 - y^2} \cdot \frac{1 - y^3}{3} \\ E(X|Y = y) &= \frac{2}{(1 + y)(1 - y)} \cdot \frac{(1 - y)(y^2 + y + 1)}{3} = \boxed{\frac{2(y^2 + y + 1)}{3(y + 1)}} \end{aligned}$$

(d)

To do this, we simply need to find $f_X(x)$:

$$\begin{aligned} f_X(x) &= \int_0^x 8xy dy \\ f_X(x) &= 8x \cdot \frac{y^2}{2} \Big|_0^x \\ f_X(x) &= 8x \cdot \left(\frac{x^2}{2} - 0\right) = 4x^3 \end{aligned}$$

and thusly find $E[X]$:

$$\begin{aligned} E[X] &= \int_0^1 x \cdot f_X(x) dx \\ E[X] &= \int_0^1 x \cdot 4x^3 dx = \int_0^1 4x^4 dx \\ E[X] &= \frac{4x^5}{5} \Big|_0^1 \\ E[X] &= \frac{4}{5} - 0 = \boxed{\frac{4}{5}} \end{aligned}$$

2.

(a)

We are given $\mu = (2, -3, 1)^T$, and $X \sim N(\mu, \Sigma)$, with $X = (X_1, X_2, X_3)^T$. We can use this fact to determine that $Y = AX \sim N(A\mu, A\Sigma A^T)$. We therefore have $A = (3, -2, 1)$ as given in the problem statement. Let's calculate the mean first:

$$\begin{aligned} E(Y) &= A\mu \\ E(Y) &= \begin{pmatrix} 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} \\ E(Y) &= \begin{pmatrix} 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} \end{aligned}$$

$$E(Y) = 3(2) + -2(-3) + 1(1) = 6 + 6 + 1 = 13$$

and now the variance:

$$\text{Var}(Y) = A\Sigma A^T$$

$$\text{Var}(Y) = \begin{pmatrix} 3 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

$$\text{Var}(Y) = \begin{pmatrix} 3(1) - 2(1) + 1(1) & 3(1) - 2(3) + 1(2) & 3(1) - 2(2) + 1(2) \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

$$\text{Var}(Y) = \begin{pmatrix} 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$$

$$\text{Var}(Y) = 2(3) - 1(-2) + 1(1) = 9$$

$$\text{So therefore } Y = AX = \boxed{3X_1 - 2X_2 + X_3 \sim N(13, 9)}$$

(b)

To find $a^T = (a_1 \ a_2)$ that satisfies the condition of X_2 and $X_2 - a^T \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ are independent, we can use the definition of independence for the jointly normal variables as the covariance being 0, i.e.

$$\text{Cov}(X_2, X_2 - a^T \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}) = 0$$

$$\text{Cov}(X_2, X_2 - (a_1 \ a_2) \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}) = 0$$

$$\text{Cov}(X_2, X_2 - a_1 X_1 - a_2 X_2) = 0$$

and using linearity of covariance:

$$\text{Cov}(X_2, X_2 - a_1 X_1 - a_2 X_2) = \text{Cov}(X_2, X_2) - a_1 \text{Cov}(X_2, X_1) - a_2 \text{Cov}(X_2, X_2)$$

Looking at Σ for these values, we see that $\text{Cov}(X_2, X_2) = \text{Var}(X_2) = \Sigma_{2,2} = 3$ and $\text{Cov}(X_2, X_1) = \Sigma_{1,2} = 1$. Plugging these in:

$$\begin{aligned}\text{Cov}(X_2, X_2) - a_1\text{Cov}(X_2, X_1) - a_2\text{Cov}(X_2, X_2) \\ = 3 - a_1(1) - a_2(3) \\ = 3 - a_1 - 3a_2 \\ a_1 + 3a_2 = 3\end{aligned}$$

One sample answer given this equation is setting $a_1 = 3$, with $a_2 = 0$ as a result. So

one example solution is
$$a = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

3.

Find $P(X^2 < Y < X)$ where $f(x, y) = 2x$ for $0 \leq x, y \leq 1$ [1].

We first establish the bounds – given that we need $X^2 < Y < X$, it is obvious that we must have $0 < x < 1$. Given these two bounds, we can now establish the double integral for representing the joint pdf and evaluate:

$$\begin{aligned}P(X^2 < Y < X) &= \int_0^1 \int_{x^2}^x 2xy dy dx \\ P(X^2 < Y < X) &= \int_0^1 2x|_{x^2}^x dx \\ P(X^2 < Y < X) &= \int_0^1 2x^2 - 2x^3 dx \\ P(X^2 < Y < X) &= \frac{2x^3}{3} - \frac{2x^4}{4}|_0^1 \\ P(X^2 < Y < X) &= \left(\frac{2}{3} - \frac{2}{4}\right) - (0 - 0) \\ P(X^2 < Y < X) &= \frac{1}{6}\end{aligned}$$

4.

(a)

Let's begin by calculating the marginal distributions from the table.

$$P(X = 1) = \frac{1}{12} + \frac{1}{6} + 0 = \frac{1}{4}$$

$$P(X = 2) = \frac{1}{6} + 0 + \frac{1}{3} = \frac{1}{2}$$

$$P(X = 3) = \frac{1}{12} + \frac{1}{6} + 0 = \frac{1}{4}$$

$$P(Y = 2) = \frac{1}{12} + \frac{1}{6} + \frac{1}{12} = \frac{1}{3}$$

$$P(Y = 3) = \frac{1}{6} + 0 + \frac{1}{6} = \frac{1}{3}$$

$$P(Y = 4) = 0 + \frac{1}{3} + 0 = \frac{1}{3}$$

Now, we can simply check for independence. We know that X and Y are independent if $P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$. Most values fulfill the condition, but we find that for $Y = 3$ and $X = 1$:

$$\begin{aligned} P(X = 1, Y = 3) &= P(X = 1) \cdot P(Y = 3) \\ \frac{1}{6} &= \frac{1}{4} \cdot \frac{1}{3} \end{aligned}$$

$$\frac{1}{6} \neq \frac{1}{12}$$

So X and Y are not independent, and therefore dependent.

(b)

We can do this by simply recreating the table, except multiply the marginals together. This ensures that they are independent, taking $X = U$ and $Y = V$:

	$U = 1$	$U = 2$	$U = 3$
$V = 2$	$\frac{1}{4} \cdot \frac{1}{3}$	$\frac{1}{2} \cdot \frac{1}{3}$	$\frac{1}{4} \cdot \frac{1}{3}$
$V = 3$	$\frac{1}{4} \cdot \frac{1}{3}$	$\frac{1}{2} \cdot \frac{1}{3}$	$\frac{1}{4} \cdot \frac{1}{3}$
$V = 4$	$\frac{1}{4} \cdot \frac{1}{3}$	$\frac{1}{2} \cdot \frac{1}{3}$	$\frac{1}{4} \cdot \frac{1}{3}$

	$U = 1$	$U = 2$	$U = 3$
$V = 2$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
$V = 3$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
$V = 4$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$

5.

We aim to find $P(S \leq 10)$ with $S = \sum_{i=1}^{20} X_i$ each variable independent and in the range $[0, 1]$ with $f(x) = 2x$. We can determine this by using the Central Limit Theorem, for which we need to find the mean and variance of S . We start by thus finding the mean and variance for one X_i :

$$E[X] = \int_0^1 x \cdot f(x) dx$$

$$E[X] = \int_0^1 x \cdot 2x dx = \int_0^1 2x^2 dx$$

$$E[X] = \frac{2x^3}{3} \Big|_0^1$$

$$E[X] = \frac{2}{3} - 0 = \frac{2}{3}$$

For variance, let's first find $E[X^2]$:

$$\begin{aligned} E[X^2] &= \int_0^1 x^2 \cdot f(x) dx \\ E[X^2] &= \int_0^1 x^2 \cdot 2x dx = \int_0^1 2x^3 dx \\ E[X^2] &= \frac{x^4}{2} \Big|_0^1 \\ E[X^2] &= \frac{1}{2} - 0 = \frac{1}{2} \end{aligned}$$

And now let's find the corresponding variance:

$$\begin{aligned} \text{Var}(X) &= E[X^2] - (E[X])^2 \\ \text{Var}(X) &= \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18} \end{aligned}$$

Since these variables are all independent, we can use linearity of expectation and variance here to find the same values for S :

$$\begin{aligned} E[S] &= 20 \cdot E[X] \\ E[S] &= 20 \cdot \frac{2}{3} = \frac{40}{3} \\ \text{Var}(S) &= 20 \cdot \text{Var}(X) \\ \text{Var}(S) &= 20 \cdot \frac{1}{18} = \frac{20}{18} = \frac{10}{9} \end{aligned}$$

Finally, let's apply the Central Limit Theorem to find $P(S \leq 10)$. We first find the corresponding Z-value to look up against, in terms of S :

$$\begin{aligned} Z &= \frac{S - E[S]}{\text{Std}(S)} \\ Z &= \frac{10 - \frac{40}{3}}{\sqrt{\frac{10}{9}}} = \frac{-\frac{10}{3}}{\frac{\sqrt{10}}{3}} = \frac{-10}{\sqrt{10}} = -3.16227766017 \end{aligned}$$

So therefore

$$P(S \leq 10) = P(Z \leq -3.16227766017) \approx \boxed{0.00079}$$

which we determine by looking up in a table.

6.

We are given the distribution of the average of the independent measurements as $\bar{X} \sim N(\mu, \frac{25}{n})$. We apply the Central Limit Theorem after standardizing the probability expression to $N(0, 1)$ with $Z = \frac{\bar{X} - \mu}{\sigma}$:

$$\begin{aligned} P(|\bar{X} - \mu| < 1) \\ P\left(\frac{-1}{\sqrt{\frac{25}{n}}} < Z < \frac{1}{\sqrt{\frac{25}{n}}}\right) \\ P\left(-\frac{\sqrt{n}}{5} < Z < \frac{\sqrt{n}}{5}\right) \end{aligned}$$

Since we are looking for tails on either side, we look for 2 symmetrical regions that give us the remaining $1 - 0.95 = 0.05$:

$$\begin{aligned} 2\phi\left(\frac{\sqrt{n}}{5}\right) &= 0.05 \\ \phi\left(\frac{\sqrt{n}}{5}\right) &= 0.025 \\ \frac{\sqrt{n}}{5} &= z_{0.025} \end{aligned}$$

We now do a lookup for the corresponding z -value, and substitute:

$$\begin{aligned} \frac{\sqrt{n}}{5} &= 1.96 \\ \sqrt{n} &= 9.8 \\ n &= 96.04 \end{aligned}$$

We round up as we need a whole number, which gives us the final answer that we need the number of measurements $n \geq 97$.

7.

(a)

We run the code a few times, reducing n and keeping ϵ constant:

```

> limitTheorem(100, 0.5)
[1] 0.9874
> limitTheorem(50, 0.5)
[1] 0.928
> limitTheorem(20, 0.5)
[1] 0.7522
> limitTheorem(5, 0.5)
[1] 0.4088

```

Here, we see that law of large numbers applies with higher n , with a higher fraction of numbers landing within the $\epsilon = 0.5$ of the mean $\mu = 4.2$.

We now run reducing ϵ and keeping n constant:

```

> limitTheorem(50, 1)
[1] 0.9992
> limitTheorem(50, 0.5)
[1] 0.9179
> limitTheorem(50, 0.1)
[1] 0.2635

```

We see that the fraction reduces as we expect it to.

(b)

While we already showed this previously, something we can do to further show this is to further standardize the averages, and we see that the fraction further decreases and the histogram becomes more normalized:

```

> limitTheorem = function(n,epsilon) {
+
+   vecOfAverages = c(rep(0,10000))
+   for (i in 1:10000) {
+     smpl = rpois(n,4.2)

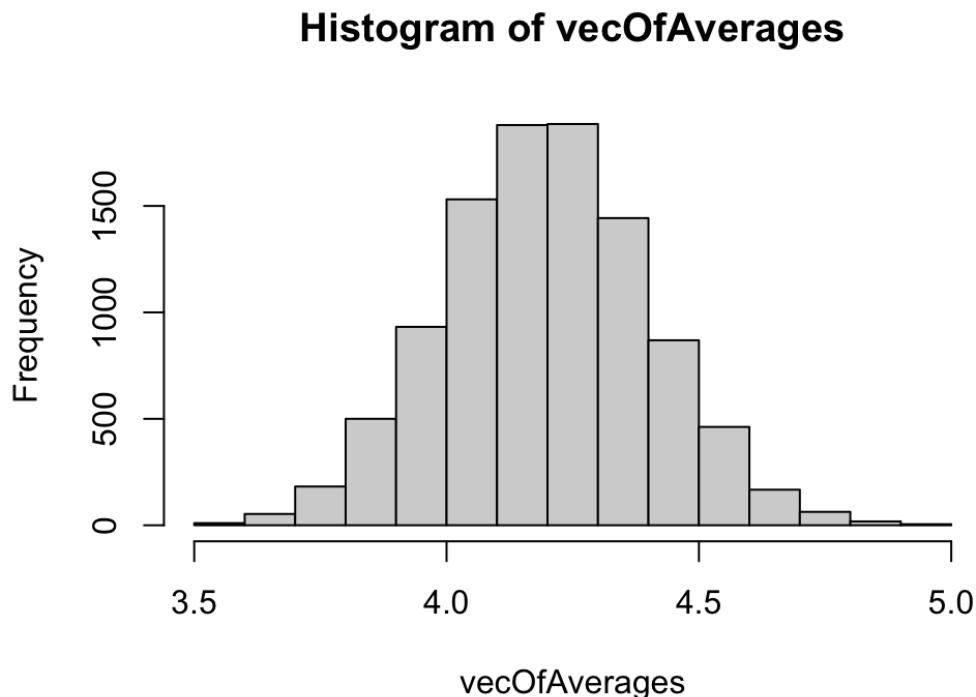
```

```

+   avg = mean(smpl);
+   vecOfAverages[i] = avg;
+ }
+ standardizedAverages = (vecOfAverages - 4.2) / sqrt(4.2/n)
+ hist(standardizedAverages);
+ vecOfDifferences = standardizedAverages - 4.2*c(rep(1,10000));
+ nmbCloseToTruth = length(vecOfDifferences[abs(vecOfDifferences) <= epsilon]);
+ print(nmbCloseToTruth/10000);
+ }
>
> limitTheorem(100, 0.5)
[1] 1e-04

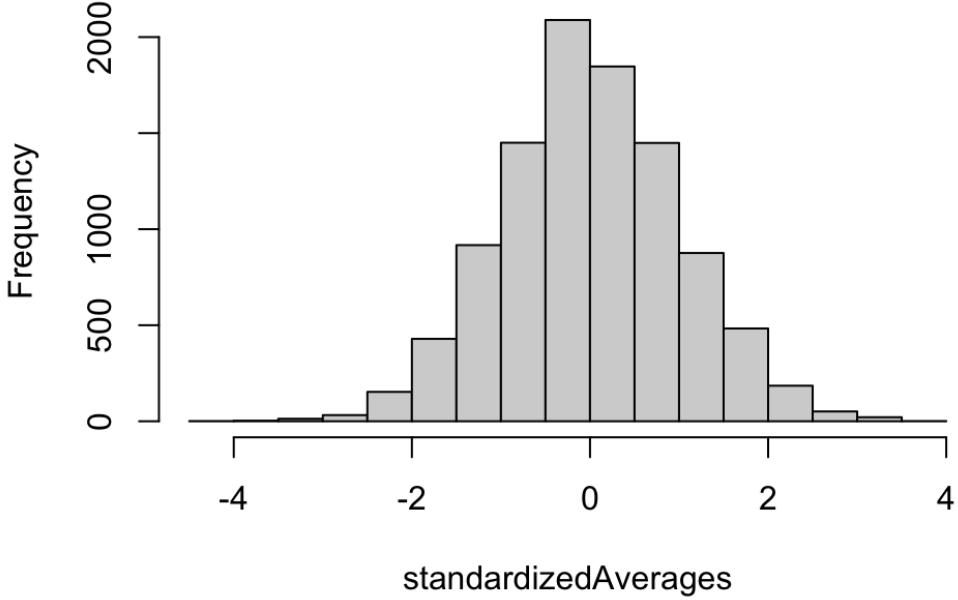
```

Before:



After:

Histogram of standardizedAverages



8.

We are given the exponential distribution with probability density $f(x|\lambda) = \lambda e^{-\lambda x}$. We first take the proper expression for the likelihood function corresponding to the n observed values of X :

$$L(\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i}$$

which we can easily simplify by multiplying to get

$$L(\lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i}$$

We then proceed to take the log likelihood:

$$\begin{aligned} l(\lambda) &= \ln(\lambda^n e^{-\lambda \sum_{i=1}^n x_i}) \\ l(\lambda) &= n \ln(\lambda) - \lambda \sum_{i=1}^n x_i \end{aligned}$$

We then try to maximize this, so find a critical point by taking the derivative and setting equal to 0:

$$0 = \frac{dl}{d\lambda}$$

$$0 = \frac{d}{d\lambda} (n \ln(\lambda) - \lambda \sum_{i=1}^n x_i)$$

$$0 = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$

$$\sum_{i=1}^n x_i = \frac{n}{\lambda}$$

$$\lambda \sum_{i=1}^n x_i = n$$

$$\lambda = \frac{n}{\sum_{i=1}^n x_i}$$

We can also ensure that this is a maximum by taking the derivative again:

$$\frac{d^2 l}{d\lambda^2} = \frac{d}{d\lambda} \left(\frac{n}{\lambda} - \sum_{i=1}^n x_i \right)$$

$$\frac{d^2 l}{d\lambda^2} = -\frac{n}{\lambda^2} - 0 = -\frac{n}{\lambda^2}$$

The original probability density must be positive, so we know that $\lambda > 0$ always.

We see that this value of $\frac{d^2 l}{d\lambda^2} < 0$ for all $\lambda > 0$, which means that this is indeed a

maximum. We can now finally conclude that $\hat{\lambda}_{\text{MLE}} = \frac{n}{\sum_{i=1}^n X_i} = \boxed{\frac{1}{\bar{X}}}$