1. X is the mixture of two exponential random variables with parameters 1 and 3, which are selected with probability 1/3 and 2/3, respectively. Hence, the PDF of X is

\[ f_X(x) = \begin{cases} \frac{1}{3} \cdot e^{-x} + \frac{2}{3} \cdot 3e^{-3x} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases} \]

2. X is a mixture of two exponential random variables, one with parameter \( \lambda \) and one with parameter \( \mu \). We select the exponential with parameter \( \lambda \) with probability \( p \), so the transform is

\[ M_X(s) = p \frac{\lambda}{\lambda - s} + (1 - p) \frac{\mu}{\mu - s}. \]

Note that the transform only exists for \( s < \min\{\lambda, \mu\} \).

3. (a) The definition of the transform is

\[ M_Z(s) = \mathbb{E}[e^{sz}] \]

Therefore, we know the following must be true:

\[ M_Z(0) = \mathbb{E}[e^{0Z}] = \mathbb{E}[1] = 1. \]

So in our case

\[ M_Z(0) = \frac{a}{8} = 1 \]

and

\[ a = 8. \]

(b) We approach this problem by first finding the PDF of \( Z \) using partial fraction expansion:

\[ M_Z(s) = \frac{8 - 3s}{s^2 - 6s + 8} = \frac{A}{s - 4} + \frac{B}{s - 2} \]

\[ A = (s - 4)M_Z(s) \bigg|_{s=4} = \frac{8 - 3s}{s - 2} \bigg|_{s=4} = -2 \]

\[ B = (s - 2)M_Z(s) \bigg|_{s=2} = \frac{8 - 3s}{s - 4} \bigg|_{s=2} = -1. \]

Thus,

\[ M_Z(s) = \frac{-2}{s - 4} + \frac{-1}{s - 2} = \frac{1}{2} \left( \frac{4}{4 - s} + \frac{2}{2 - s} \right) \]

and

\[ f_Z(z) = \begin{cases} \frac{1}{2} (4e^{-4z} + 2e^{-2z}) & \text{for } z \geq 0, \\ 0 & \text{otherwise.} \end{cases} \]

From this we get

\[ P(Z \geq 0.5) = \int_{0.5}^{\infty} \frac{1}{2} (4e^{-4z} + 2e^{-2z}) dz = \frac{e^{-2} + e^{-1}}{2}. \]

(c) \( \mathbb{E}[Z] = \int_{0}^{\infty} \frac{1}{2} (4e^{-4z} + 2e^{-2z}) dz = \frac{1}{2} \left( \int_{0}^{\infty} 4ze^{-4z} dz + \int_{0}^{\infty} 2ze^{-2z} dz \right) = \frac{1}{2} \left( \frac{1}{4} + \frac{1}{2} \right) = \frac{3}{8} \)

(d) \( \mathbb{E}[Z] = \left. \frac{d}{ds} M_Z(s) \right|_{s=0} = \left. \frac{d}{ds} \left( \frac{2}{4-s} + \frac{1}{2-s} \right) \right|_{s=0} = \frac{2}{(4-s)^2} + \frac{1}{(2-s)^2} \bigg|_{s=0} = \frac{3}{8} \)
(e) $\text{var}(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2$

$\mathbb{E}[Z^2] = \int_0^\infty \frac{2}{z^2} (4e^{-4z} + 2e^{-2z})dz = \frac{1}{2} (\int_0^\infty 4e^{-4z}dz + \int_0^\infty 2e^{-2z}dz) = \frac{1}{2} (\frac{2}{4} + \frac{2}{2}) = \frac{5}{16}$

$\text{var}(Z) = \frac{5}{16} - (\frac{3}{8})^2 = \frac{11}{64}$

(f) $\mathbb{E}[Z^2] = \frac{d^2}{ds^2} \mathbb{E}[Z(s)] \bigg|_{s=0} = \frac{d^2}{ds^2} \left( \frac{2}{4-s} + \frac{1}{2-s} \right) \bigg|_{s=0} = \frac{4}{(4-s)^2} + \frac{2}{(2-s)^2} \bigg|_{s=0} = \frac{5}{16}$

$\text{var}(Z) = \mathbb{E}[Z^2] - (\mathbb{E}[Z])^2 = \frac{5}{16} - (\frac{3}{8})^2 = \frac{11}{64}$

4. (a) Since it is impossible to get a run of $n$ heads with fewer than $n$ tosses, it is clear that $p_T(k) = 0$ for $k < n$. In addition, the probability of getting $n$ heads in $n$ tosses is $q^n$ so $p_T(n) = q^n$. Lastly, for $k \geq n+1$, we have $T = k$ if there is no run of $n$ heads in the first $k-n-1$ tosses, followed by a tail, followed by a run of $n$ heads, so

$$p_T(k) = \mathbb{P}(T > k-n-1)(1-q)q^n = \left( \sum_{i=k-n}^\infty p_T(i) \right) (1-q)q^n.$$ 

(b) We use the PMF we obtained in the previous part to compute the moment generating function. Thus,

$$M_T(s) = \mathbb{E}[e^{sT}] = \sum_{k=-\infty}^{\infty} p_T(k)e^{sk} = q^n e^{sn} + (1-q)q^n \sum_{k=n+1}^{\infty} \sum_{i=k-n}^{\infty} p_T(i)e^{sk}.$$ 

We observe that the set of pairs $\{(i, k) \mid k \geq n+1, i \geq k-n\}$ is equal to the set of pairs $\{(i, k) \mid i \geq 1, n+1 \leq k \leq i+n\}$, so by reversing the order of the summations, we have

$$M_T(s) = q^n e^{sn} + (1-q)q^n \sum_{i=1}^{\infty} \sum_{k=n+1}^{i+n} p_T(i)e^{sk} = q^n e^{sn} \left( 1 + (1-q) \sum_{i=1}^{\infty} \sum_{k=1}^{i} p_T(i)e^{sk} \right) = q^n e^{sn} \left( 1 + (1-q) \sum_{i=1}^{\infty} p_T(i) \frac{e^s - e^{s(i+1)}}{1-e^s} \right) = q^n e^{sn} \left( 1 + \frac{(1-q)e^s}{1-e^s} \sum_{i=1}^{\infty} p_T(i)(1-e^{-si}) \right).$$

Now, since $\sum_{i=1}^{\infty} p_T(i) = 1$ and, by definition, $\sum_{i=1}^{\infty} p_T(i)e^{si} = M_T(s)$, it follows that

$$M_T(s) = q^n e^{sn} \left( 1 + \frac{(1-q)e^s}{1-e^s}(1-M_T(s)) \right).$$

Rearrangement yields

$$M_T(s) = \frac{1 + (1-q)e^s}{1-q^n e^{sn} + (1-q)e^s} = q^n e^{sn}(1-e^s) + (1-q)q^n \sum_{i=1}^{\infty} (1-e^{-si}) = q^n e^{sn}(1-qe^s) + (1-q)q^n e^{sn(n+1)}.$$ 

(c) We have

$$\mathbb{E}[T] = \frac{d}{ds} M_T(s) \bigg|_{s=0} = \left\{ \frac{[1-e^s+(1-q)q^n e^{sn(n+1)}][nq^n e^{sn}(1-qe^s) - qe^s q^n e^{sn}]}{(1-e^s+(1-q)q^n e^{sn(n+1)})^2} \right\}$$
5. We calculate $f_{X|Y}(x|y)$ using the definition of a conditional density. To find the density of $Y$, recall that $Y$ is normal, so the mean and variance completely specify $f_Y(y)$. $Y = X + N$, so $E[Y] = E[X] + E[N] = 0 + 0 = 0$. Because $X$ and $N$ are independent, $\text{var}(Y) = \text{var}(X) + \text{var}(N) = \sigma_x^2 + \sigma_n^2$. So,

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{f_X(x)f_N(y - x)}{f_Y(y)}$$

$$= \frac{1}{\sqrt{2\pi\sigma_x^2}} \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{x^2}{2\sigma_x^2}} e^{-\frac{(y-x)^2}{2\sigma_n^2}}$$

$$= \frac{1}{\sqrt{2\pi(\sigma_x^2 + \sigma_n^2)^2}} \frac{1}{e^{2(\sigma_x^2 + \sigma_n^2)} - x^2 \sigma_x^2 - (y-x)^2 \sigma_n^2}.$$

We can simplify the exponent as follows.

$$\frac{y^2}{2(\sigma_x^2 + \sigma_n^2)} - \frac{x^2}{2\sigma_x^2} - \frac{(y-x)^2}{2\sigma_n^2} = \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{y^2 \sigma_x^2 \sigma_n^2 - x^2 \sigma_n^2 (\sigma_x^2 + \sigma_n^2)}{\sigma_x^2 + \sigma_n^2} - \frac{y - x)^2 \sigma_x^2}{\sigma_x^2 + \sigma_n^2} \right)$$

$$= \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{x^2 \sigma_n^2 - x^2 \sigma_n^2 (\sigma_x^2 + \sigma_n^2)}{\sigma_x^2 + \sigma_n^2} - \frac{x^2 \sigma_n^2}{\sigma_x^2 + \sigma_n^2} \right)$$

$$= \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{x^2 \sigma_n^2}{(\sigma_x^2 + \sigma_n^2)^2} \right)$$

$$= \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{y^2 \sigma_x^2 \sigma_n^2 - x^2 \sigma_n^2 (\sigma_x^2 + \sigma_n^2)}{(\sigma_x^2 + \sigma_n^2)^2} - \frac{x^2 \sigma_n^2}{(\sigma_x^2 + \sigma_n^2)^2} \right)$$

$$= \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{-x^2 \sigma_n^2}{(\sigma_x^2 + \sigma_n^2)^2} \right)$$

$$= \frac{\sigma_x^2 + \sigma_n^2}{2\sigma_x^2 \sigma_n^2} \left( \frac{x - y \frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2}}{(\sigma_x^2 + \sigma_n^2)^2} \right)^2.$$

Note that for $n = 1$, this equation reduces to $E[T] = 1/q$, which is the mean of a geometrically-distributed random variable, as expected.
Thus, we obtain

\[
f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi\frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2}}} e^{-\frac{(x-y)^2}{2\frac{\sigma_x^2}{\sigma_x^2 + \sigma_n^2}}}.\]

Looking at this formula, we see that the conditional density is normal with mean \(\frac{\sigma_x^2 y}{\sigma_x^2 + \sigma_n^2}\) and variance \(\frac{\sigma_x^2 \sigma_n^2}{\sigma_x^2 + \sigma_n^2}\).

6. Let \(R_i\) be the number rolled on the \(i^{th}\) die. Since each number is equally likely to rolled, the PMF of each \(R_i\) is uniformly distributed from 1 to 6. The PMF of \(X_1\) is obtained by convolving the PMFs of \(R_1\) and \(R_2\). Similarly, the PMF of \(X_2\) is obtained by convolving the PMFs of \(R_3\) and \(R_4\). \(X_1\) and \(X_2\) take on values from 2 to 12 and are independent and identically distributed random variables. The PMF of either one is given by

\[
p_{X_1}^{(i)}, p_{X_2}^{(i)}
\]

Note that the sum \(X_1 + X_2\) takes on values from 4 to 24. The discrete convolution formula tells us that for \(n\) from 4 to 24:

\[
P(X_1 + X_2 = n) = \sum_{i=1}^{n} P(X_1 = i)P(X_2 = n - i)
\]

so

\[
P(X_1 + X_2 = 8) = \sum_{i=1}^{8} P(X_1 = i)P(X_2 = 8 - i)
\]

and thus we find the desired probability is \(\frac{35}{36^2} = .027\).

7. The PDF for \(X\) and \(Y\) are as follows,
Because $X$ and $Y$ are independent and $W = X + Y$, the pdf of $W$, $f_W(w)$, can be written as the convolution of $f_X(x)$ and $f_Y(y)$:

$$f_W(w) = \int_{-\infty}^{\infty} f_X(x)f_Y(w - x)\,dx$$

There are five ranges for $w$:

1. $w \leq 0$

2. $0 \leq w \leq 1$

3. $1 \leq w \leq 2$
4. $2 \leq w \leq 3$

$$f_W(w) = \begin{cases} 
  f_X(x)f_Y(w-x) \, dx, & 0 \leq w \leq 1 \\
  f_{w-1}^x f_X(x)f_Y(w-x) \, dx, & 1 \leq w \leq 2 \\
  f_{x-1}^w f_X(x)f_Y(w-x) \, dx, & 2 \leq w \leq 3 \\
  0, & \text{otherwise}
\end{cases}$$

Therefore,

$$f_W(w) = \begin{cases} 
  2w - \frac{3}{2}w^2 + \frac{1}{6}w^3, & 0 \leq w \leq 1 \\
  \frac{7}{6} - \frac{7}{2}w, & 1 \leq w \leq 2 \\
  \frac{9}{7} - \frac{9}{2}w + \frac{3}{2}w^2 - \frac{1}{6}w^3, & 2 \leq w \leq 3 \\
  0, & \text{otherwise}
\end{cases}$$

5. $3 \leq w$

$$f_W(w) = \begin{cases} 
  f_X(x)f_Y(w-x) \, dx, & 0 \leq w \leq 1 \\
  f_{w-1}^x f_X(x)f_Y(w-x) \, dx, & 1 \leq w \leq 2 \\
  f_{x-1}^w f_X(x)f_Y(w-x) \, dx, & 2 \leq w \leq 3 \\
  0, & \text{otherwise}
\end{cases}$$

Therefore,

$$f_W(w) = \begin{cases} 
  2w - \frac{3}{2}w^2 + \frac{1}{6}w^3, & 0 \leq w \leq 1 \\
  \frac{7}{6} - \frac{7}{2}w, & 1 \leq w \leq 2 \\
  \frac{9}{7} - \frac{9}{2}w + \frac{3}{2}w^2 - \frac{1}{6}w^3, & 2 \leq w \leq 3 \\
  0, & \text{otherwise}
\end{cases}$$

G1. To compute $f_W(w)$, we will start by computing the joint PDF $f_{Y,Z}(y,z)$. Computing the joint density is quite simple. Define the joint CDF $F_{Y,Z}(y,z) = P(Y \leq y, Z \leq z)$. Now, $F_Z(z) = P(Z \leq z) = z^n$, because the maximum is less than $z$ if and only if every one of the $X_i$ is less than $z$, and all the $X_i$’s are independent. We can also compute $P(y \leq Y, Z \leq Z) = (z - y)^n$ because the minimum is greater than $y$ and the maximum is less than $z$ if and only if every $X_i$ falls between $y$ and $z$. Subtraction gives

$$F_{Y,Z}(y,z) = z^n - (z - y)^n.$$ 

Now, we find the joint PDF by differentiating, which gives $f_{Y,Z}(y,z) = n(n-1)(z-y)^{n-2}, 0 \leq y \leq z \leq 1$. Because $Y$ and $Z$ are not independent, convolving the individual densities for $Y$ and $Z$ will not give us the density for $W$. Instead, we must calculate the CDF $F_W(w)$ by integrating $P_{Y,Z}(y,z)$ over the appropriate region. We consider the cases $w \leq 1$ and $w > 1$ separately.
When \( w \leq 1 \), we need to compute
\[
\int_0^w \int_y^{w-y} f_{Y,Z}(y, z) \, dz \, dy = \frac{w^n}{2}.
\]

When \( w > 1 \), we can compute the CDF from
\[
1 - \int_w^1 \int_{w-z}^z f_{Y,Z}(y, z) \, dy \, dz = 1 - \frac{(2 - w)^n}{2}.
\]

Finally, we take the derivative to get
\[
f_W(w) = \begin{cases} 
  \frac{nw^{n-1}}{2} & ; 0 \leq w \leq 1 \\
  \frac{n(2-w)^{n-1}}{2} & ; 1 \leq w \leq 2 \\
  0 & ; \text{otherwise}
\end{cases}
\]

To prove the concentration result, it is easier to look at \( F_W(w) \). The CDF is exponential in \( n \). Thus, \( P(W \leq 1 - \epsilon) = \frac{(1-\epsilon)^n}{2} \) and \( P(W \geq 1 + \epsilon) = 1 - (1 - \frac{(2-(1+\epsilon))^n}{2}) = \frac{(1-\epsilon)^n}{2} \). It is easily seen that both of these probabilities go to 0 as \( n \to \infty \), which proves the desired concentration result.