Problem $X . Y$ stands for the $Y$ th Problem of the $X$ th Chapter, e.g. 2.19 is the 19 th Problem of Chapter 2
Problem 2.19-(b). The solution of this problem is described in the textbook. The questions I have received about this problem convinced me that preparing a tree representation for the bisection strategy would be helpful.

In the following figures "Is $P \leq X$ ?" stands for "Is the prize in a box numbered less than or equal to $X$ ?"


Figure 1.
Note that with the bisection strategy described in Figure 1, when the prize is in any of the boxes 1, 2, 6, 7 it takes four questions to learn the box number and when the prize is in any of the boxes $3,4,5,8,9,10$ it takes three questions to learn the box number. Since the prize is equally likely to be in any of the ten boxes expected number of questions is given by $\mathbf{E}[\mathbf{Q}]=4 \frac{4}{10}+3 \frac{6}{10}=3.4$.

Remark 1. The bisection strategy described in Figure 1 is not the only strategy consistent with the general strategy described in the problem. However, any strategy consistent with the description given in the problem will have four boxes reached with four questions and six boxes reached with three questions.


Figure 2.
Remark 2. Bisection strategies that are not consistent with the one described in the problem can achieve $\mathbf{E}[\mathbf{Q}]=3.4$, as well. For example consider the strategy that first asks if $P \leq 6$ and if the answer is yes asks if $P \leq 4$ and follows the bisection strategy described in the problem for the rest. Then boxes $1,2,3,4$ will be reached with for questions and boxes $5,6,7,8$, 9,10 will be reached with three questions.

Remark 3. The strategies that reveal the box number with the minimum expected number of questions can be found more generally via the Huffman coding. See https://en.wikipedia.org/wiki/Huffman_coding, if you want to learn more about Huffman coding. But the Huffman coding is NOT a subject that you are responsible for in EE230.

Solution of Problem 2.23. The solution manual presents an elegant solution to this problem. The following solution differs from the one in the manual in the way it derives the PMF of the random variable $X$ representing the number of coin tosses in both parts and in the way it calculates $\mathbf{E}[\mathrm{X}]$ in part (b).
(a) Any outcome of the experiment is either a string composed of $\ell$ consecutive $H T$ 's followed by one of the stings from the set $\{H H, H T T\}$ or a string composed of $\ell$ consecutive $T H$ 's followed by one of the stings from the set $\{T T, T H H\}$ for some non-negative integer $\ell$. Thus the sample space $\Omega$ is given by

$$
\Omega=\{H H, T T, H T T, T H H, H T H H, T H T T, H T H T T, T H T H H, \ldots\} .
$$

Since we have independent tosses of a fair coin probability of length $k$ string of $H$ 's and $T$ 's is equal to $2^{-k}$.
Let the random variable $X$ be the number of coin tosses,

$$
X(\omega):=\text { the length of the string } \omega \text {. }
$$

Note for any $k \in\{2,3, \ldots\}$ there are two outcomes of length $k$ in $\Omega$ each with probability $2^{-k}$ and for $k<2$ there are no outcomes of length $k$ in $\Omega$. Thus

$$
\mathbf{P}(\mathrm{X}=k)= \begin{cases}2^{1-k} & k \in\{2,3, \ldots\} \\ 0 & \text { otherwise }\end{cases}
$$

Let the random variable Y be $\mathrm{Y}=\mathrm{X}-1$, then

$$
\mathbf{P}(\mathrm{Y}=k)= \begin{cases}\left(\frac{1}{2}\right)^{k} & k \in\{1,2, \ldots\} \\ 0 & \text { otherwise }\end{cases}
$$

Thus Y is a geometric random variable with $p=\frac{1}{2}$. Consequently, $\mathbf{E}[\mathrm{Y}]=\frac{1}{p}=2$ and $\operatorname{var}(\mathrm{Y})=\frac{1-p}{p^{2}}=2$. On the other hand $\mathbf{E}[\mathrm{X}]=\mathbf{E}[\mathrm{Y}]+1$ by the linearity property of the expectation and $\operatorname{var}(\mathrm{X})=\operatorname{var}(\mathrm{Y})$ because adding a constant to a random variable does not change its variance. Hence,

$$
\mathbf{E}[\mathrm{X}]=3 \quad \text { and } \quad \operatorname{var}(\mathrm{X})=2
$$

(b) Any outcome of the experiment is a string that is composed of $\ell_{1}$ consecutive $H$ 's followed by $\ell_{2}$ consecutive $T$ 's followed by the string $T H$ for some non-negative integers $\ell_{1}$ and $\ell_{2}$. Thus the sample space $\Omega$ is given by

$$
\Omega=\{\omega \mid \exists \ell_{1}, \ell_{2} \in\{0,1,2,3\} \text { such that } \omega=\underbrace{H \cdots H}_{\ell_{1}} \underbrace{T \cdots T}_{\ell_{2}} T H\}
$$

Since we have independent tosses of a fair coin probability of length $k$ string of $H$ 's and $T$ 's is equal to $2^{-k}$.
Let the random variable $X$ be the number of coin tosses,

$$
X(\omega):=\text { the length of the string } \omega \text {. }
$$

Note for any $k \in\{2,3, \ldots\}$ there are $k-1$ distinct $\left(\ell_{1}, \ell_{2}\right)$ pairs satisfying the identity $\ell_{1}+\ell_{2}=k-2$ and for $k<2$ there are no outcomes of length $k$ in $\Omega$. Thus

$$
\mathbf{P}(\mathrm{X}=k)= \begin{cases}(k-1) 2^{-k} & k \in\{2,3, \ldots\} \\ 0 & \text { otherwise }\end{cases}
$$

Thus

$$
\begin{aligned}
\mathbf{E}[\mathrm{X}] & =\sum_{k=2}^{\infty} k(k-1)\left(\frac{1}{2}\right)^{k} \\
& =\frac{1}{4} \sum_{k=2}^{\infty} k(k-1)\left(\frac{1}{2}\right)^{k-2} \\
& =\left.\frac{1}{4} \sum_{k=2}^{\infty} \frac{\mathrm{d}^{2}}{\mathrm{~d} s^{2}} s^{k}\right|_{s=\frac{1}{2}} \\
& =\left.\frac{1}{4} \sum_{k=0}^{\infty} \frac{\mathrm{d}^{2}}{\mathrm{~d} s^{2}} s^{k}\right|_{s=\frac{1}{2}} \\
& =\left.\frac{1}{4} \frac{\mathrm{~d}^{2}}{\mathrm{~d} s^{2}} \sum_{k=0}^{\infty} s^{k}\right|_{s=\frac{1}{2}} \\
& =\left.\frac{1}{4} \frac{\mathrm{~d}^{2}}{\mathrm{~d} s^{2}} \frac{1}{1-s}\right|_{s=\frac{1}{2}} \\
& =\left.\frac{1}{4} \frac{2}{(1-s)^{3}}\right|_{s=\frac{1}{2}} \\
& =4 .
\end{aligned}
$$

Solution of Problem 3.6. Let the random variable $X$ be Jane's waiting time and the event $A$ be the event that Jane finds a customer ahead of her when she arrives
Then given event the $A$, the waiting time of Jane is equal to $Z-\tau$ where $Z$ is the service time of the customer ahead of Jane and $\tau$ is the difference between Jane's arrival time and the beginning of the service.

$$
\begin{equation*}
\mathbf{P}(\mathbf{X}>x \mid \mathrm{A})=\mathbf{P}(\mathrm{Z}-\tau>x \mid \mathrm{A}) \quad \forall x, \tau \in \mathbb{R}_{\geq 0} \tag{1}
\end{equation*}
$$

Since $Z$ is exponentially distrusted, it has the memorylessness property that can be summarized as follows:

$$
\begin{equation*}
\mathbf{P}\left(Z>t_{1}+t_{2} \mid Z>t_{1}\right)=\mathbf{P}\left(Z>t_{2}\right) \quad \forall t_{1}, t_{2} \in \mathbb{R}_{\geq 0} \tag{2}
\end{equation*}
$$

Since Z is exponentially distrusted with parameter $\lambda$

$$
\mathbf{P}\left(\mathbf{Z}>t_{2}\right)= \begin{cases}e^{-\lambda t_{2}} & \text { if } t_{2} \geq 0  \tag{3}\\ 1 & \text { if } t_{2}<0\end{cases}
$$

On the other hand event A is equivalent to the event $Z>\tau$. Thus (1), (2), and (3) imply

$$
\mathbf{P}(\mathrm{X}>x \mid \mathrm{A})= \begin{cases}e^{-\lambda x} & \text { if } x \geq 0  \tag{4}\\ 1 & \text { if } x<0\end{cases}
$$

We are told in the problem description that there is either one or zero customers ahead of Jane. Thus $\mathrm{A}^{c}$ is the event that Jane finds no customers in the que when she arrives. In this case Jane's waiting time will be 0 . Thus

$$
\begin{equation*}
\mathbf{P}\left(\mathbf{X}=0 \mid \mathrm{A}^{c}\right)=1 \tag{5}
\end{equation*}
$$

Consequently,

$$
\mathbf{P}\left(\mathrm{X}>x \mid \mathrm{A}^{c}\right)= \begin{cases}0 & \text { if } x \geq 0  \tag{6}\\ 1 & \text { if } x<0\end{cases}
$$

On the other hand as a result of total probability theorem we have

$$
\mathbf{P}(\mathbf{X}>x)=\mathbf{P}(\mathrm{A}) \mathbf{P}(\mathbf{X}>x \mid \mathrm{A})+\mathbf{P}\left(\mathrm{A}^{c}\right) \mathbf{P}\left(\mathbf{X}>x \mid \mathrm{A}^{c}\right)
$$

Since Jane having one or zero customers ahead of her are equally likely $\mathbf{P}(A)=\mathbf{P}\left(A^{c}\right)=1 / 2$. Thus

$$
\begin{equation*}
\mathbf{P}(\mathbf{X}>x)=\frac{1}{2}\left[\mathbf{P}(\mathbf{X}>x \mid \mathrm{A})+\mathbf{P}\left(\mathbf{X}>x \mid \mathrm{A}^{c}\right)\right] . \tag{7}
\end{equation*}
$$

Using (4), (6), and (7) we get

$$
\mathbf{P}(\mathrm{X}>x)= \begin{cases}\frac{1}{2} e^{-\lambda x} & \text { if } x \geq 0 \\ 1 & \text { if } x<0\end{cases}
$$

Thus

$$
\begin{align*}
f_{\mathrm{X}}(x) & =\mathbf{P}(\mathbf{X} \leq x)  \tag{8}\\
& =1-\mathbf{P}(\mathbf{X}>x)  \tag{9}\\
& = \begin{cases}\frac{2-e^{-\lambda x}}{2} & \text { if } x \geq 0 \\
0 & \text { if } x<0\end{cases} \tag{10}
\end{align*} .
$$

Remark 4. We tacitly assumed that the difference between Jane's arrival and the commencement of service for the last customer before Jane is a fixed time in (1). We, however, need not make that assumption; the memorylessness of the exponential distribution would imply the same result, even when that difference is some non-negative random variable.

