Exercise 1. Solution: By definition, p is a density if and only if $p(x) \geq 0$ for any $x \in \mathbf{R}$ and $\int_{\mathbf{R}} p(x)\mathrm{d}x = 1$. The first condition implies that $a \geq 0$. Obviously, $a \neq 0$ (since otherwise $\int_{\mathbf{R}} p(x)\mathrm{d}x = 0$). To find b, notice that

$$1 = \int_{\mathbf{R}} p(x)\mathrm{d}x = e^b \int_0^\infty ae^{-ax}\mathrm{d}x = e^b.$$

Hence, p is a density function if and only if $a > 0, b = 0$.

To find the expectation and variance, one can use the "integration by parts" trick. Let X denote the random variable with the density p. We have

$$\begin{aligned}\mathbb{E}(X) = \int_{\mathbf{R}} xp(x)\mathrm{d}x &= \int_0^\infty x \cdot \underbrace{\left(ae^{-ax}\right)}_{=(-e^{-ax})'} \mathrm{d}x \\ &= \underbrace{\left(-xe^{-ax}\right)\Big|_{x=0}^{x=\infty}}_{=0} + \int_0^\infty e^{-ax}\mathrm{d}x = \frac{1}{a}.\end{aligned}$$

Above, we used the fact that e^x grows faster than any power x^k as $x \to \infty$ (why?).

We next have to compute the variance of the random variable X, i.e., $\mathrm{var}(X)$. Recall that $\mathrm{var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Thus,

$$\begin{aligned}\mathbb{E}(X^2) = \int_{\mathbf{R}} x^2 p(x)\mathrm{d}x &= \int_0^\infty x^2 \cdot \underbrace{\left(ae^{-ax}\right)}_{=(-e^{-ax})'} \mathrm{d}x \\ &= \underbrace{\left(-x^2 e^{-ax}\right)\Big|_{x=0}^{x=\infty}}_{=0} + \frac{2}{a}\int_0^\infty xae^{-ax}\mathrm{d}x = \frac{2}{a^2},\end{aligned}$$

(the latter term is nothing else than the expectation, which has already been computed, multiplied by $\frac{2}{a}$). As a result, $\mathrm{var}(X) = \frac{1}{a^2}$.

Exercise 2. Solution: Let us first note that the two events of interest are complementary, i.e., their probabilities sum up to 1. Therefore, it suffices to compute the probability of one of these two events. We begin by defining two random events: Let

- W denote the event "the drawn ball is white,"
- U_k denote the event "the k-th urn has been chosen."

Note that event B is equivalent to the event U_1 given W (in words, the remaining ball is black if the first urn was chosen). So, to compute the probability of the event B, we need to compute $\mathbb{P}(U_1|W)$.

Now, observe that $\mathbb{P}(W|U_1) = 1/2$, $\mathbb{P}(W|U_2) = 0$, and $\mathbb{P}(W|U_3) = 1$ (recall that the conditional probability $\mathbb{P}(W|U_k)$ simply states the probability of the event that *the drawn ball is white* given that *the urn U_k is chosen*). Moreover, notice that $\mathbb{P}(U_k) = \frac{1}{3}$ for $k = 1, 2, 3$. As a result, one can use *the law of total probability* to compute $P(W)$ as follows:

$$\mathbb{P}(W) = \sum_{k=1}^{3} \mathbb{P}(W|U_k)\mathbb{P}(U_k) = \frac{1}{2}.$$

Considering the *Bayes' law*, we next have

$$\mathbb{P}(W \cap U_1) = \mathbb{P}(W|U_1)\mathbb{P}(U_1) = \frac{1}{6},$$

and hence

$$\mathbb{P}(U_1|W) = \frac{\mathbb{P}(W \cap U_1)}{\mathbb{P}(W)} = \frac{1}{3}.$$

This means that the probability of event A (which is equal to $\mathbb{P}(U_3|W)$) is $\frac{2}{3}$. Therefore, event A is more probable.

Exercise 3. Solution: We first need to transform the provided description of the problem to a mathematical one. Note that there are two sources of stochasticity: (i) the individual who is tested might be healthy (H) or infected (I) and (ii) the test result might be positive (P) or negative (N). Correspondingly, we define two r.v.'s X and Y with spaces $\mathbb{X} = \{H, I\}$ and $\mathbb{Y} = \{P, N\}$, respectively. It is easy to see that the exercise is asking for the conditional probability $\mathbb{P}(X = I \mid Y = P)$. Now, let's see what information we can extract from the statement of the exercise. First of all, since only 0.1% of the population is infected, we have

$$\mathbb{P}(X = I) = 1 - \mathbb{P}(X = H) = 0.001.$$

Second, a false positive rate of 5% means

$$\mathbb{P}(Y = P \mid X = H) = 1 - \mathbb{P}(Y = N \mid X = H) = 0.05.$$

Similarly, a false negative rate of 5% means

$$\mathbb{P}(Y = N \mid X = I) = 1 - \mathbb{P}(Y = P \mid X = I) = 0.05.$$

Observe that the information that we have is conditioned upon X, while the information that we are looking for is conditioned upon Y. This is a clear signal for using the Bayes' rule:

$$\mathbb{P}(X = I \mid Y = P) = \frac{\mathbb{P}(Y = P \mid X = I) \cdot \mathbb{P}(X = I)}{\mathbb{P}(Y = P)} = \frac{0.95 \times 0.001}{\mathbb{P}(Y = P)}.$$

Now, we only need to compute the *marginal* probability of Y using the conditional probabilities on X. This is a clear signal for using the law of total probability:

$$\begin{aligned}\mathbb{P}(Y = P) &= \mathbb{P}(Y = P \mid X = H) \cdot \mathbb{P}(X = H) + \mathbb{P}(Y = P \mid X = I) \cdot \mathbb{P}(X = I) \\ &= 0.05 \times 0.999 + 0.95 \times 0.001\end{aligned}$$

Therefore,

$$\mathbb{P}(X = I \mid Y = P) = \frac{0.95 \times 0.001}{0.05 \times 0.999 + 0.95 \times 0.001} \approx 0.02.$$

So, there is only 2% chance that a random individual who tested positive is actually infected! Try to reflect on this result and connect it to the percentage of the population that is infected and the reliability of the test.

Exercise 4. Solution: For the first lemma, we have

$$\mathbb{E}(X + Y) = \sum_{(x,y) \in \mathbb{X} \times \mathbb{Y}} (x + y) \cdot p(x, y)$$

$$
\begin{aligned}
&= \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} x \cdot p(x,y) + \sum_{(x,y)\in\mathbb{X}\times\mathbb{Y}} y \cdot p(x,y) \\
&= \sum_{x\in\mathbb{X}} \left(x \cdot \sum_{y\in\mathbb{Y}} p(x,y) \right) + \sum_{y\in\mathbb{Y}} \left(y \cdot \sum_{x\in\mathbb{X}} p(x,y) \right) \\
&= \sum_{x\in\mathbb{X}} \left(x \cdot \sum_{y\in\mathbb{Y}} p(x,y) \right) + \sum_{y\in\mathbb{Y}} \left(y \cdot \sum_{x\in\mathbb{X}} p(x,y) \right) \\
&= \sum_{x\in\mathbb{X}} x \cdot p_X(x) + \sum_{y\in\mathbb{Y}} y \cdot p_Y(y) \\
&= \mathbb{E}(X) + \mathbb{E}(Y).
\end{aligned}
$$

For the second lemma, we have

$$
\begin{aligned}
\mathbb{E}\big(\mathbb{E}(X|Y)\big) &= \sum_{y\in\mathbb{Y}} \mathbb{E}(X|Y=y) \cdot p_Y(y) \\
&= \sum_{y\in\mathbb{Y}} \sum_{x\in\mathbb{X}} x \cdot p_{X|Y}(x|y) \cdot p_Y(y) \\
&= \sum_{y\in\mathbb{Y}} \sum_{x\in\mathbb{X}} x \cdot p(x,y) \\
&= \sum_{x\in\mathbb{X}} \left(x \cdot \sum_{y\in\mathbb{Y}} p(x,y) \right) \\
&= \sum_{x\in\mathbb{X}} x \cdot p_X(x) \\
&= \mathbb{E}(X).
\end{aligned}
$$

Exercise 5. Solution: The equality immediately follows from the first-order optimality condition:

$$
\begin{aligned}
& 0 = \frac{\partial}{\partial x}(x^\top Q x + q^\top x)\bigg|_{x=x^\star} \\
\iff & 0 = 2Qx^\star + q \\
\iff & x^\star = -\frac{1}{2}Q^{-1}q,
\end{aligned}
$$

where for the last equality we used the fact that Q is positive definite and hence invertible.

Exercise 6. Solution: For $k=1$ (*the base case*), we clearly have $1 = \frac{1}{6} \cdot 1 \cdot 2 \cdot 3$. Next, for some $k \in \mathbf{N}$, assume (*induction hypothesis*)

$$
1^2 + 2^2 + \ldots + k^2 = \frac{1}{6}k(k+1)(2k+1).
$$

Now, observe that

$$
\begin{aligned}
1^2 + 2^2 + \ldots + k^2 + (k+1)^2 &= \frac{1}{6}k(k+1)(2k+1) + (k+1)^2 \\
&= \frac{1}{6}(k+1)(2k^2 + 7k + 6) \\
&= \frac{1}{6}(k+1)\big((k+2)(2k+3)\big)
\end{aligned}
$$

$$= \frac{1}{6}(k+1)\big((k+1)+1\big)\big(2(k+1)+1\big)$$

where for the first equality, we used the induction hypothesis. This completes the proof.