Question 1

Recall from Term 1 that the probability density function of the uniform distribution on the interval (a, b) is

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b, \\ 0, & \text{otherwise.} \end{cases}$$

We write $X \sim \mathrm{U}(a, b)$ to indicate that the random variable X follows this distribution.

(a) If $X \sim \mathrm{U}(a, b)$, compute $\mathrm{E}(X)$.

(b) If $X \sim \mathrm{U}(a, b)$, compute $\mathrm{Var}(X)$.

Solution to Question 1

Part (a):

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x = \int_a^b \frac{x}{b-a}\mathrm{d}x = \left[\frac{1}{2} \cdot \frac{x^2}{b-a}\right]_a^b = \frac{a+b}{2}$$

Part (b):

$$\mathrm{E}\left(X^2\right) = \int_{-\infty}^{\infty} x^2 f_X(x)\,\mathrm{d}x = \int_a^b \frac{x^2}{b-a}\mathrm{d}x = \left[\frac{1}{3} \cdot \frac{x^3}{b-a}\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$\Rightarrow \mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - (\mathrm{E}(X))^2 = \frac{b^2+ab+a^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Question 2

Suppose X is uniformly distributed on the interval $[0, 4]$, i.e. $X \sim \text{Unif}(0, 4)$.

(a) Compute $\mathrm{P}(|X-2| \geq 1)$.

(b) Use Chebyshev's inequality to bound the probability that $|X-2| \geq 1$.

(c) Is the bound in (b) informative?

(d) For which values $\epsilon > 0$ can Chebyshev's inequality be used to obtain a nontrivial bound for $\mathrm{P}(|X-2| \geq \epsilon)$?

Solution to Question 2

Part (a):

The solution can be briefly written as

$$\mathrm{P}(|X-2| \geq 1) = \mathrm{P}(\{X \geq 3\} \text{ or } \{X \leq 1\}) = \mathrm{P}(X \geq 3) + \mathrm{P}(X \leq 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

When written out in detail, we first note that:

$$\begin{aligned}
&|X-2| \geq 1 \\
&\Rightarrow X - 2 \geq 1 \text{ or } -(X-2) \geq 1 \\
&\Rightarrow X \geq 3 \text{ or } (X-2) \leq -1 \\
&\Rightarrow X \geq 3 \text{ or } X \leq 1
\end{aligned}$$

Now:

- Let A be the event $X \geq 3$.
- Let B be the event $X \leq 1$.
- Let C be the event $|X-2| \geq 1$.

Since an observed value of X cannot be simultaneously bigger than 3 and less than 1, $A \cap B = \varnothing$. Also, from the above, $C = A \cup B$. Therefore,

$$\begin{aligned}
\mathrm{P}(C) &= \mathrm{P}(A \cup B) \\
&= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B) \\
&= \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(\varnothing) \\
\Rightarrow \mathrm{P}(C) &= \mathrm{P}(A) + \mathrm{P}(B)
\end{aligned}$$

since $\mathrm{P}(\varnothing) = 0$. So,

$$\mathrm{P}(|X-2| \geq 1) = \mathrm{P}(X \geq 3) + \mathrm{P}(X \leq 1).$$

The probability density function of $\text{Unif}(a, b)$ (from PBL Sheet 8) is

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a < x < b, \\ 0, & \text{otherwise.} \end{cases}$$

Therefore, the p.d.f. for $\text{Unif}(0, 4)$ is

$$f_X(x) = \begin{cases} \frac{1}{4}, & \text{if } 0 < x < 4, \\ 0, & \text{otherwise.} \end{cases}$$

The cumulative distribution function for $X \sim \text{Unif}(0,4)$ is then (for $x \in [0,4]$):

$$\begin{aligned} F_X(x) = \mathrm{P}(X \leq x) &= \int_{-\infty}^{x} f_X(x)\,\mathrm{d}x \\ &= \int_{-\infty}^{0} (0)\,\mathrm{d}x + \int_{0}^{x} \frac{1}{4}\,\mathrm{d}x \\ &= \frac{x}{4} \end{aligned}$$

Then,

$$\mathrm{P}(X \leq 1) = \frac{1}{4}.$$

Since X is a continuous random variable, $\mathrm{P}(X = 3) = 0$, and therefore

$$\begin{aligned} \mathrm{P}(X \geq 3) &= 1 - \mathrm{P}(X < 3) \\ &= 1 - (\mathrm{P}(X < 3) + 0) = 1 - (\mathrm{P}(X < 3) + \mathrm{P}(X = 3)) \\ &= 1 - \mathrm{P}(X \leq 3) \\ &= 1 - \frac{3}{4} \\ &= \frac{1}{4} \end{aligned}$$

Therefore,

$$\mathrm{P}(|X-2| \geq 1) = \mathrm{P}(X \geq 3) + \mathrm{P}(X \leq 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Part (b):

We recall Chebyshev's inequality for a random variable X with mean μ and variance σ^2. For any constant $c > 0$,

$$\mathrm{P}(|X-\mu| \geq c) \leq \frac{\sigma^2}{c^2}.$$

Here we have $X \sim \text{Unif}(0,4)$. For a general $Y \sim \text{Unif}(a,b)$, we computed in PBL Sheet 8, Question 1, that

$$\begin{aligned} \mathrm{E}(Y) &= \frac{a+b}{2} \\ \mathrm{Var}(Y) &= \frac{(b-a)^2}{12} \end{aligned}$$

Therefore, we have

$$\begin{aligned} \mu &= \mathrm{E}(X) = 2 \\ \sigma^2 &= \mathrm{Var}(X) = \frac{4^2}{12} = \frac{16}{12} = \frac{4}{3} \end{aligned}$$

Taking $c = 1$, Chebyshev's inequality then gives us the bound

$$\mathrm{P}(|X-2| \geq 1) \leq \frac{4}{3}.$$

Part (c):

This is not informative, because $\mathrm{P}(|X-2| \geq 1) \in [0,1]$, and so we already had the bound

$$\mathrm{P}(|X-2| \geq 1) \leq 1 < \frac{4}{3}.$$

Part (d):

For any $\epsilon > 0$, Chebyshev's inequality gives

$$\mathrm{P}(|X-2| \geq \epsilon) \leq \frac{(4/3)}{\epsilon^2} = \frac{4}{3\epsilon^2},$$

and this bound is only nontrivial when

$$\begin{aligned} &\frac{4}{3\epsilon^2} < 1 \\ \Rightarrow &\epsilon^2 > \frac{4}{3} \\ \Rightarrow &\epsilon > \frac{2}{\sqrt{3}}. \end{aligned}$$

Question 3

Suppose that a population is taking part in a vote and an unknown proportion p of the voters supports a particular option, labelled A. Suppose it is possible to interview a sample of n randomly selected voters and record $\widehat{p}$, the proportion of that sample that supports option A. What value of n should be chosen so that with high confidence (confidence at least 95%) $\widehat{p}$ is within 0.01 of p?

Solution to Question 3

One will notice the similarity to Exercise 1.3.4 in the notes, and we start the same way. Let us label our sample of n voters from 1 to n, and let X_i be the random variable with value $x_i = 1$ if voter i supports option A, and $x_i = 0$ otherwise. By this construction, each $X_i \sim \mathrm{Bern}(p)$, where p is the unknown parameter we wish to estimate, and $\widehat{p} = \overline{x}$. Since each X_i has mean $\mathrm{E}\left(X_i\right) = p$ and variance $\mathrm{Var}\left(X_i\right) = p(1-p)$, using Proposition 1.2.6, $\mathrm{E}\left(\overline{X}\right) = p$ and $\mathrm{Var}\left(\overline{X}\right) = p(1-p)/n$. Therefore, for any $\epsilon > 0$, Chebyshev's Inequality in Theorem 1.3.4 gives us:

$$\mathrm{P}\left(\left|\overline{X} - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2}.$$

Furthermore, using Corollary 1.1.17, one can remove the unknown p on the right-hand side to obtain

$$\mathrm{P}\left(\left|\overline{X} - p\right| \geq \epsilon\right) \leq \frac{1}{4n\epsilon^2}.$$

Now, we want to find the value of n so that (when $\epsilon = 0.01$)

$$\mathrm{P}\left(\left|\overline{X} - p\right| \geq 0.01\right) \leq 1 - 0.95 = 0.05.$$

Instead of trying to directly bound $\mathrm{P}\left(\left|\overline{X} - p\right| \geq 0.01\right)$ to be less than 0.05, we instead can bound $\frac{1}{4n\epsilon^2}$; i.e. $\mathrm{P}\left(\left|\overline{X} - p\right| \geq 0.01\right) \leq \frac{1}{4n\epsilon^2} \leq 0.05$. We solve:

$$\begin{aligned}
&\frac{1}{4n\epsilon^2} \leq \frac{5}{100} \\
\Rightarrow &4n\epsilon^2 \geq \frac{100}{5} \\
\Rightarrow &4n(0.01)^2 \geq 20 \\
\Rightarrow &n \geq \left(\frac{20}{4}\right)(100)^2 = 50000
\end{aligned}$$

Therefore, taking a sample of at least 50,000 voters will give us an estimate of p to within $\epsilon = 0.01$ with confidence 95%. The statements of Proposition 1.2.6, Theorem 1.3.4 and Corollary 1.1.17 from the notes are below.

Proposition 1.2.6. Suppose that the sample $X_1, X_2, \ldots, X_n$ are independently sampled from a distribution F_X that has mean μ and finite variance σ^2. Then

1. $\mathrm{E}\left(\overline{X}\right) = \mu$,
2. $\mathrm{Var}\left(\overline{X}\right) = \dfrac{\sigma^2}{n}$,
3. $\mathrm{E}\left(S^2\right) = \sigma^2$.

Theorem 1.3.4 (Chebyshev's Inequality)**.** If X is a random variable with mean μ and variance σ^2, then for all $c > 0$,

$$\mathrm{P}\left(|X - \mu| \geq c\right) \leq \frac{\sigma^2}{c^2}.$$

Corollary 1.1.17. Suppose $X \sim \mathrm{Bern}(p)$, for some $p \in [0, 1]$. Then $\mathrm{Var}\left(X\right) = p(1-p) \leq \frac{1}{4}$.

Question 4

Consider the probability space $(\Omega, \mathcal{F}, \mathrm{P})$. Recall from Term 1 the definition of an indicator variable for an event $A \in \mathcal{F}$, denoted $\mathbb{I}_A$ (or $\mathbb{I}(A)$) and defined for $\omega \in \Omega$ by

$$\mathbb{I}_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A. \end{cases}$$

(a) Is $\mathbb{I}_A$ a discrete random variable or a continuous random variable?

(b) If $\mathbb{I}_A$ is discrete, write down its probability mass function, or if it is continuous, write down its probability density function.

(c) Compute $\mathrm{E}(\mathbb{I}_A)$.

Solution to Question 4

Part (a): Since $\mathrm{Im}(\mathbb{I}_A) = \{\mathbb{I}_A(\omega) : \omega \in \Omega\} = \{0, 1\}$, and is therefore countable, the random variable $\mathbb{I}_A$ is a discrete random variable (Definition 7.2.1 in Prof. Veraart's notes from Term 1).

Part (b): Since $\mathbb{I}_A$ is discrete, it has a probability mass function (p.m.f.). One might directly be able to write down the p.m.f. as:

$$p_{\mathbb{I}_A}(x) = \begin{cases} \mathrm{P}(A), & \text{if } x = 1, \\ 1 - \mathrm{P}(A), & \text{if } x = 0, \\ 0, & \text{if } x \notin \{0, 1\}. \end{cases}$$

However, **the box on the next page contains a careful derivation**.

Part (c): There are two (similar) approaches to computing the expectation of $\mathbb{I}_A$.

The first approach is to directly use the definition of expectation for a discrete random variable:

$$\begin{aligned} \mathrm{E}(X) &= \sum_{x \in \mathrm{Im}(\mathbb{I}_A)} x p_{\mathbb{I}_A}(x) \\ &= 1 \cdot p_{\mathbb{I}_A}(1) + 0 \cdot p_{\mathbb{I}_A}(0) \\ &= 1 \cdot \mathrm{P}(A) + 0 \cdot (1 - \mathrm{P}(A)) \\ \Rightarrow \mathrm{E}(X) &= \mathrm{P}(A). \end{aligned}$$

The second approach takes a shortcut. One can consider $\mathbb{I}_A$ to be random variable following a Bernoulli distribution with parameter λ, i.e. $\mathbb{I}_A \sim \mathrm{Bern}(\lambda)$, where $\lambda = p_{\mathbb{I}_A}(1) = \mathrm{P}(A)$. Therefore, $\mathrm{E}(\mathbb{I}_A) = \lambda = \mathrm{P}(A)$.

(One would usually use "p" as the parameter for a Bernoulli distribution, but "λ" was used here to avoid confusion with the p.m.f. $p_{\mathbb{I}_A}$.)

The first approach is from [2, Chap. 24, p. 203-204] and the second approach is from [1, Theorem 4.4.2, p. 164].

Part (b): (in detail)

Recall the definition of a probability mass function from Prof. Veraart's notes from Term 1:

Definition 7.2.4 (Probability mass function)**.** The **probability mass function** (p.m.f.) of the discrete random variable X is defined as the function $p_X : \mathbb{R} \to [0,1]$ given by

$$p_X(x) = \mathrm{P}\left(\{\omega \in \Omega \,:\, X(\omega) = x\}\right).$$

Now, since $\mathrm{Im}\left(\mathbb{I}_A\right) = \{\mathbb{I}_A(\omega) \,:\, \omega \in \Omega\} = \{0,1\}$, there are only three cases to consider and so one can directly use Definition 7.2.4 and the definition of $\mathbb{I}_A$ to obtain:

$$p_{\mathbb{I}_A}(x) = \mathrm{P}\left(\{\omega \in \Omega \,:\, \mathbb{I}_A(\omega) = x\}\right)$$

$$= \begin{cases} \mathrm{P}\left(\{\omega \in \Omega \,:\, \mathbb{I}_A(\omega) = 1\}\right), & \text{if } x = 1, \\ \mathrm{P}\left(\{\omega \in \Omega \,:\, \mathbb{I}_A(\omega) = 0\}\right), & \text{if } x = 0, \\ \mathrm{P}\left(\{\omega \in \Omega \,:\, \mathbb{I}_A(\omega) = x\}\right), & \text{if } x \notin \{0,1\} \end{cases}$$

$$= \begin{cases} \mathrm{P}\left(\{\omega \in \Omega \,:\, \omega \in A\}\right), & \text{if } x = 1, \\ \mathrm{P}\left(\{\omega \in \Omega \,:\, \omega \notin A\}\right), & \text{if } x = 0, \\ \mathrm{P}\left(\varnothing\right), & \text{if } x \notin \{0,1\} \end{cases} \tag{1}$$

$$= \begin{cases} \mathrm{P}\left(A\right), & \text{if } x = 1, \\ \mathrm{P}\left(A^c\right), & \text{if } x = 0, \\ 0, & \text{if } x \notin \{0,1\} \end{cases} \tag{2}$$

$$\Rightarrow p_{\mathbb{I}_A}(x) = \begin{cases} \mathrm{P}\left(A\right), & \text{if } x = 1, \\ 1 - \mathrm{P}\left(A\right), & \text{if } x = 0, \\ 0, & \text{otherwise} \end{cases}$$

Equation (1) leads to Equation (2) by recalling that $A \in \mathcal{F} \Rightarrow A \subseteq \Omega$.

Equation (1) uses $\varnothing$ to denote the empty set; the set is empty because there are no elements of Ω with $\mathbb{I}_A(\omega) = x$ and $x \notin \mathrm{Im}\left(\mathbb{I}_A\right) = \{0,1\}$.

References

[1] J. K. Blitzstein and J. Hwang. *Introduction to probability.* Chapman and Hall/CRC, 2nd edition, 2014.

[2] M. Taboga. *Lectures on probability theory and mathematical statistics.* CreateSpace Independent Publishing Platform, 3rd edition, 2017.