Continuous Probability Distributions

21.1

Introduction

In Block 5.1 we met the idea of a discrete probability distribution. There, the values of a random variable could be written in the form of a list. We learned how to characterise these distributions by calculating the expectation and the variance. In this block we consider so-called continuous random variables whose values cannot be written in the form of a list. We shall see that such random variables can be characterised by a function $f(x)$ called the probability density function. We shall also learn how to determine the expectation and the variance of continuous probability distributions.

Prerequisites

Before starting this Block you should . . .

1. understand the concepts of probability
2. be familiar with the concepts of expectation and variance

Learning Outcomes

After completing this Block you should be able to . . .

✓ explain the term continuous probability distribution
✓ distinguish between discrete and continuous random variables
✓ find the mean and variance of a continuous probability distribution

Learning Style

To achieve what is expected of you . . .

Allocate sufficient study time

Briefly revise the prerequisite material

Attempt every guided exercise and most of the other exercises
1. Continuous random variables

Consider a de-magnetised compass needle mounted at its centre so that it can spin freely. Its initial position is shown in Figure 1(i). It is spun clockwise and when it comes to rest the angle $\theta$, from the vertical, is measured. See Figure 1(ii).

![Figure 1](image)

Let $X$ be the random variable “angle $\theta$ measured after each spin”

Now $X$ is clearly a random variable since it can take different values and we cannot be sure in advance which value it will take.

With respect to this experiment there are two types of questions we can ask:

- What is the probability that $X$ lies between two values $a, b$, i.e. what is $P(a < X < b)$
- What is the probability that $X$ assumes a particular value, i.e. $P(X = c)$

The first type of question is easy to answer. For example, most of us would agree that since the needle is equally likely to be in any given angular range as any other then

$$P\left(0 < X < \frac{\pi}{2}\right) = \frac{1}{4} \quad \text{and} \quad P\left(\frac{\pi}{2} < X < 2\pi\right) = \frac{3}{4} \quad \text{or} \quad P(a < X < b) = \frac{b - a}{2\pi}$$

The second type of question: finding $P(X = c)$ is not a sensible enquiry. In order to answer a question of this kind would require a measuring device (e.g. a protractor) with infinite precision: no such device exists nor could ever be constructed. Hence it can never be verified that the needle, after spinning, takes any particular value; all we can be reasonably sure of is that the needle lies between two particular values.

We conclude that in experiments of this kind we never determine the probability that the random variable assume a particular value but only calculate the probability that it lies within a given range of values. This kind of random variable is called a **continuous random variable** and it is characterised, not by probabilities of the type $P(X = x_i)$, but by a function $f(x)$ called the **probability density function** (p.d.f.). In the case of the rotating needle this function takes the simple form:

$$f(x) = \begin{cases} \frac{1}{2\pi}, & 0 \leq x < 2\pi \\ 0, & \text{elsewhere} \end{cases}$$

The probabilities $P(a < X < b)$ are related to the area under the function curve $f(x)$. Note that $\int_{-\infty}^{\infty} f(x)dx = \int_{0}^{2\pi} \frac{1}{2\pi} dx = 1$, i.e. total probability is 1, and that $f(x) \geq 0$ for all $x$. 

Engineering Mathematics: Open Learning Unit Level 1
21.1: Probability
Suppose we wanted to find \( P \left( \frac{\pi}{6} < X < \frac{\pi}{4} \right) \).
Then
\[
P \left( \frac{\pi}{6} < X < \frac{\pi}{4} \right) = \int_{\frac{\pi}{6}}^{\frac{\pi}{4}} \frac{1}{2\pi} \, dx = \left[ \frac{x}{2\pi} \right]_{\frac{\pi}{6}}^{\frac{\pi}{4}} = \frac{1}{2\pi} \left( \frac{\pi}{4} - \frac{\pi}{6} \right) = \frac{1}{2\pi} \times \frac{\pi}{12} = \frac{1}{24}
\]
This is reasonable since the range \( \frac{\pi}{6} \) to \( \frac{\pi}{4} \) is one twenty-fourth of the range 0 to \( 2\pi \).

We now proceed formally to the definition of a continuous random variable.

**Key Point**

\( X \) is said to be a **continuous random variable** if there exists a function \( f(x) \) associated with \( X \) called the **probability density function** with the properties

- \( f(x) \geq 0 \) for all \( x \)
- \( \int_{-\infty}^{\infty} f(x) \, dx = 1 \)
- \( P(a < X < b) = \int_{a}^{b} f(x) \, dx \)

The first two bullet points in the keypoint are the analogues of the results \( P(X = x_i) \geq 0 \) and \( \sum_i P(X = x_i) = 1 \) for discrete random variables.

**Try each part of this exercise**

Which of the following are not probability density functions?

(i) \( f(x) = \frac{1}{x} \)

(ii) \( f(x) = \frac{1}{x} \) for \( 1 \leq x \leq 2 \)

(iii) \( f(x) = \begin{cases} 
  x^2 - 4x + \frac{10}{3}, & 0 \leq x \leq 3 \\
  0, & \text{elsewhere}
\end{cases} \)

Part (i) Check that the first two statements in the keypoint are satisfied

Part (ii) Repeat as for (i).

Part (iii) Proceed along similar lines to test \( f(x) \).
Try each part of this exercise

Find the probability that $X$ takes a value between $-1$ and $1$ when the p.d.f. is given by the following figure.

![Probability Distribution Function](image)

Part (a) First find $k$

Part (b) What is the formula for $f(x)$?

Part (c) Write down an integral to represent $P(-1 < X < 1)$. Remember to use symmetry, and evaluate the integral.

2. Mean and Variance of a continuous distribution

The cumulative distribution function (c.d.f) $F(x)$ of a continuous random variable is given by

$$F(a) = \int_{-\infty}^{a} f(x)dx$$

for any value $a$. The value $F(a)$ represents the area under the p.d.f. graph to the left of $x = a$ and gives the probability $P(X < a)$.

It follows that $P(a \leq X \leq b) = F(b) - F(a)$.

Now do this exercise

For the p.d.f. in the previous guided exercise obtain the c.d.f. and verify the result obtained for $P(-1 \leq X \leq 1)$.

From the graph $F(-2) = 0$, $F(0) = \frac{1}{2}$, $F(2) = 1$
The mean and variance of a continuous random variable

As we have seen a discrete random variable \( X \) has a mean value (or expectation) \( \mu \) which is a measure of the central position of the distribution. We can find the centre of a continuous random variable also. Here, the centre of the distribution is such that there is just as much area under \( f(x) \) to the left of the centre as there is under \( f(x) \) to the right. Also, the variance of a continuous random variable is a measure of the spread of \( f(x) \) about the centre.

We give the formal definitions of the expectation and variance of a continuous random variable in the following keypoint:

**Key Point**

Let \( X \) be a continuous random variable with associated p.d.f. \( f(x) \). Then its expectation and variance denoted by \( E(X) \) (or \( \mu \)) and \( V(X) \) (or \( \sigma^2 \)) respectively are given by:

\[
E(X) = \int_{-\infty}^{\infty} x f(x) \, dx \quad \text{and} \quad V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) \, dx
\]

As with discrete random variables the variance \( V(X) \) can be written in an alternative form, more amenable to calculation:

\[
V(X) = E(X^2) - [E(X)]^2
\]

where \( E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) \, dx \).

**Try each part of this exercise**

For the variable \( X \) with p.d.f.

\[
f(x) = \begin{cases} 
\frac{1}{2} x, & 0 \leq x \leq 2 \\
0, & \text{elsewhere}
\end{cases}
\]

find \( E(X) \) and then \( V(X) \).

Part (a) First we find \( E(X) \)

Answer

Part (b) Now we find \( E(X^2) \).

Answer

Part (c) Now find \( V(X) \)

Answer
More exercises for you to try

1. The mileage (in 1000’s of miles) for which a certain type of tyre will last is a random variable with p.d.f.

\[ f(x) = \begin{cases} 
\frac{1}{20}e^{-x/20} & \text{for all } x > 0 \\
0 & \text{for all } x \leq 0 
\end{cases} \]

Find the probability that the tyre will last:

(a) at most 10,000 miles,
(b) between 16,000 and 24,000 miles
(c) at least 30,000 miles.

2. If the p.d.f. of a random variable \(X\) is

\[ f(x) = \begin{cases} 
kx & \text{for } 0 \leq x \leq 5 \\
0 & \text{elsewhere} 
\end{cases} \]

(a) calculate \(k\)
(b) calculate \(P(1 \leq X \leq 3), P(2 \leq X \leq 4), P(X \leq 3)\).

3. In the manufacture of petroleum the distilling temperature \((T^\circ C)\) is crucial in determining the quality of the final product. \(T\) can be considered as a random variable uniformly distributed over 150\(^\circ\)C to 300\(^\circ\)C. Its costs \(LC_1\) to produce 1 gallon of petroleum. If the oil distills at temperatures less than 200\(^\circ\)C the product sells for \(LC_2\) per gallon. If it distills at a temperature greater than 200\(^\circ\)C it sells for \(LC_3\) per gallon. Find the expected net profit per gallon.

4. A target is made of three concentric circles of radii \(1/\sqrt{3}, 1\) and \(\sqrt{3}\) metres. Shots within the inner circle count 4 points, in the next ring 3 points and within the third ring 2 points. (Shots outside the target count zero.) The distance of a shot from the centre of the target is a random variable \(R\) with density function.

\[ f(r) = \frac{2}{\pi(1+r^2)}, \quad r > 0. \]

Calculate the expected value of the score after five shots.
(i) We can write \( f(x) = \begin{cases} x, & 0 \leq x \leq 1 \\ 0, & \text{elsewhere} \end{cases} \)

\( f(x) \geq 0 \) for all \( x \) but

\[
\int_{-\infty}^{\infty} f(x) \, dx = \int_{0}^{1} x \, dx = \frac{1}{2} \neq 1.
\]

Thus this function is not a valid probability density function.
Note that \( f(x) = \begin{cases} 
1 - \frac{1}{2}x, & 0 \leq x \leq 2 \\
0, & \text{elsewhere} 
\end{cases} \) for all \( x \)

\[
\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{2} \left(1 - \frac{1}{2}x\right)\,dx = \left[ x - \frac{x^2}{4} \right]_0^2 = 2 - 1 = 1
\]
(Alternatively, the area of the triangle is \( \frac{1}{2} \times 1 \times 2 = 1 \))

This implies that \( f(x) \) is a valid probability density function.
(iii) 

\[
\int_{-\infty}^{\infty} f(x) \, dx = \int_{0}^{3} \left( x^2 - 4x + \frac{10}{3} \right) \, dx = \left[ \frac{x^3}{3} - 2x^2 + \frac{10}{3} x \right]_{0}^{3} \\
= (9 - 18 + 10) = 1
\]

but \( f(x) < 0 \) for \( 1 \leq x \leq 3 \). Hence (i) and (iii) are not p.d.f's.

Back to the theory
\[ \int_{-\infty}^{\infty} f(x) \, dx = 2 \frac{1}{2} k \times 2 \quad \text{(by symmetry)} \]

\[ = 2k = 1 \]

Hence \( k = \frac{1}{2} \)

Back to the theory.
\[ f(x) = \begin{cases} 
\frac{1}{2} - \frac{1}{4}x, & 0 \leq x \leq 2 \\
\frac{1}{2} + \frac{1}{4}x, & -2 \leq x < 0 \\
0, & \text{elsewhere.}
\end{cases} \]
\[ \int_{-1}^{1} f(x)dx = 2 \int_{0}^{1} \left( \frac{1}{2} - \frac{1}{4}x \right) dx = 2 \left[ \frac{1}{2}x - \frac{1}{8}x^2 \right]_{0}^{1} = 2 \left( \frac{1}{2} - \frac{1}{8} \right) = \frac{3}{4} \]
\[ F(x) = \begin{cases} 
0, & x \leq -2 \\
\frac{1}{2} + \frac{1}{2}x + \frac{1}{8}x^2, & -2 < x < 0 \\
\frac{1}{2} + \frac{1}{2}x - \frac{1}{8}x^2, & 0 < x \leq 2 \\
1, & x \geq 2 
\end{cases} \]

\[ P(-1 \leq x \leq 1) = F(1) - F(-1) = \left( \frac{1}{2} + \frac{1}{2} - \frac{1}{8} \right) - \left( \frac{1}{2} - \frac{1}{4} + \frac{1}{8} \right) = \frac{1}{2} + \frac{1}{2} - \frac{1}{8} - \frac{1}{8} = \frac{3}{4}. \]
\[ E(X) = \int_{0}^{2} \frac{1}{2} x \, dx = \left[ \frac{1}{6} x^3 \right]_0^2 = \frac{8}{6} = \frac{4}{3}. \]

Back to the theory.
\[ E(X^2) = \int_0^2 \frac{1}{2}x.x^2 \, dx = \left[ \frac{1}{8}x^4 \right]_0^2 = 2. \]
\[ V(X) = E(X^2) - \{E(X)\}^2 \]
\[ = 2 - \frac{16}{9} = \frac{2}{9}. \]
1(a) \[ P(a < X < b) = \int_a^b f(x)dx \]

\[ P(X < 10) = \int_{-\infty}^{10} \frac{1}{20} e^{-x/20}dx \]

\[ = \int_{0}^{10} \frac{1}{20} e^{-x/20}dx = \left[-e^{-x/20}\right]_{0}^{10} = 0.3934 \]

1(b)

\[ P(16 < X < 24) = \int_{16}^{24} \frac{1}{20} e^{-x/20}dx \]

\[ = \left[-e^{-x/20}\right]_{16}^{24} = -e^{-1.2} + e^{-0.8} = 0.148 \]

1(c)

\[ P(X > 30) = \int_{30}^{\infty} \frac{1}{20} e^{-x/20}dx \]

\[ = \left[-e^{-x/20}\right]_{30}^{\infty} = e^{-1.5} = 0.223 \]

2(a) \[ \int_{-\infty}^{\infty} f(x)dx = 1 \]

\[ \therefore \int_{0}^{5} kxdx = \left[ \frac{kx^2}{2} \right]_{0}^{5} = \frac{25k}{2} = 1 \]

\[ \therefore k = 0.08 \]

2(b)

\[ P(1 \leq X \leq 3) = \int_{1}^{3} 0.08x \, dx = \left[0.04x^2\right]_{1}^{3} = 0.32 \]

\[ P(2 \leq X \leq 4) = \int_{2}^{4} 0.08x \, dx = (0.04)12 = 0.48 \]

\[ P(X \leq 3) = \int_{0}^{3} 0.08x \, dx = 0.36 \]

3.

\[ P(X < 200) = 50 \cdot \frac{1}{150} = \frac{1}{3} \quad P(X > 200) = \frac{2}{3} \]
Let $F$ be a random variable defining profit. $F$ can take two values $\mathcal{L}(C_2 - C_1)$ or $\mathcal{L}(C_3 - C_1)$

<table>
<thead>
<tr>
<th>$x$</th>
<th>$C_2 - C_1$</th>
<th>$C_3 - C_1$</th>
<th>$P(F = x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/3</td>
<td>2/3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$$E(F) = \frac{C_2 - C_1}{3} + \frac{2}{3}[C_3 - C_1] = \frac{C_2 - 3C_1 + 2C_3}{3}$$

4.

$$P(\text{inner hit}) = P\left(0 < r < \frac{1}{\sqrt{3}}\right)$$

$$= \int_0^{\frac{1}{\sqrt{3}}} \frac{2}{\pi(1 + r^2)} dr = \frac{2}{\pi} [\tan^{-1} r]_0^{\frac{1}{\sqrt{3}}}$$

$$= \frac{2}{\pi} \tan^{-1} \frac{1}{\sqrt{3}} = \frac{2}{\pi} \left(\frac{\pi}{6}\right) \frac{1}{3}$$

5 (continued)

$$P(\text{middle band}) = P\left(\frac{1}{\sqrt{3}} < r < 1\right)$$

$$= \int_{\frac{1}{\sqrt{3}}}^{1} \frac{2}{\pi(1 + r^2)} dr = \frac{2}{\pi} [\tan^{-1} r]_{\frac{1}{\sqrt{3}}}^{1} = \frac{2}{\pi} \tan^{-1} 1 - \frac{1}{3} = \frac{1}{6}$$

$$P(\text{outer band}) = P(1 < r < \sqrt{3}) = \frac{2}{\pi} [\tan^{-1} r]_{1}^{\sqrt{3}} = \frac{2}{\pi} \tan^{-1} \sqrt{3} - \frac{1}{2} = \frac{1}{6}$$

$$P(\text{miss target}) = 1 - \frac{1}{6} - \frac{1}{6} - \frac{1}{3} = \frac{1}{3}$$

Let $S$ be a random variable equal to ‘score’.

<table>
<thead>
<tr>
<th>$s$</th>
<th>0</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(S = s)$</td>
<td>1/3</td>
<td>1/6</td>
<td>1/6</td>
<td>1/3</td>
</tr>
</tbody>
</table>

$$E(S) = \frac{5}{6} + \frac{4}{3} = \frac{13}{6}$$

The expected score after 5 shots is this value times 5 which is: $= 5 \left(\frac{13}{6}\right) = 10.83$.

Back to the theory.