

GCE



FURTHER MATHEMATICS
AS UNIT 2: FURTHER STATISTICS A

Exponential Distribution



Further Mathematics AS Unit 2: Further Statistics A

Exponential Distribution	2
Continuous Probability Distributions.....	2
The Exponential Distribution.....	3
Specification Content.....	3
Worked Example 1.....	4
Worked Example 2.....	5
Worked Example 3.....	6
Summary of Key Points for the Exponential Distribution	6
Questions	7
Question 1 Worked Solution.....	9
Question 2 Worked Solution.....	9
Question 3 Worked Solution.....	10
Question 4 Worked Solution.....	10
Question 5 Worked Solution.....	11
Question 6 Worked Solution.....	12

Further Mathematics AS Unit 2: Further Statistics A

Exponential Distribution

The exponential distribution is a continuous probability distribution. It can be used to describe the time successive events from a Poisson distribution. For example, the number of cars passing a point on a motorway could be modelled by a Poisson distribution (assuming the modelling assumptions are met). In this case, the intervals of time between successive cars follow an exponential distribution.

Before introducing the exponential distribution, let us revisit the key concepts for continuous probability distributions.

Continuous Probability Distributions

In this unit, you have already been introduced to continuous probability distributions. Recall that a continuous random variable X can be defined using a probability density function, denoted $f(x)$, such that areas under the curve represent probabilities. We assume that $f(x)$ is defined between in some interval $a \leq x \leq b$, and zero elsewhere. Let c and d be such that $a \leq c \leq d \leq b$, then:

$$\int_a^b f(x) dx = 1, \quad P(c \leq X \leq d) = \int_c^d f(x) dx, \quad P(X \leq d) = \int_a^d f(x) dx, \quad P(X \geq c) = \int_c^b f(x) dx.$$

Note that $P(X = x) = 0$ for any value of x . The probability density function can also be used to calculate expected values as well as the variance and standard deviation of the random variable X as follows:

$$E(X) = \int_a^b xf(x) dx, \quad E(X^2) = \int_a^b x^2f(x) dx, \quad E(g(X)) = \int_a^b g(x)f(x) dx.$$

From this, we can calculate the variance and standard deviation using:

$$\text{Var}(X) = E(X^2) - (E(X))^2, \quad \text{SD}(X) = \sqrt{\text{Var}(X)}.$$

Calculation of probabilities is often easier using the cumulative distribution function $F(x)$. This is calculated using:

$$F(x) = \int_a^x f(t) dt.$$

Note that if $f(x)$ is defined for all real values of x , the lower limit of a is replaced by $-\infty$.

Assuming $f(x)$ is defined on the interval $a \leq x \leq b$, and is therefore zero elsewhere, then:

$$F(x) = \begin{cases} 0 & \text{if } x < a \\ \int_a^x f(t) dt & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

Observe that since F is obtained from f by integration, we can also obtain f from F by differentiation. Hence:

$$f(x) = F'(x).$$

Let c and d be such that $a \leq c \leq d \leq b$, then:

$$P(X \leq c) = F(c), \quad P(c \leq X \leq d) = F(d) - F(c), \quad P(X \geq d) = 1 - F(d).$$

In this section, we consider a specific example of a continuous probability distribution function, namely the exponential distribution.

Further Mathematics AS Unit 2: Further Statistics A

The Exponential Distribution

The exponential distribution with parameter λ can be used to model the time between two successive events from a Poisson distribution with mean λ .

Specification Content

- Statistical distributions: exponential distribution
- Find and use the mean and variance of an exponential distribution
 - knowledge and use of: If $Y \sim \text{Exp}(\lambda)$, then $E(Y) = \frac{1}{\lambda}$ and $\text{Var}(Y) = \frac{1}{\lambda^2}$.
- Use the exponential distribution as a model for intervals between events
 - learners will be expected to know that $\frac{d}{dx}(e^{kx}) = ke^{kx}$.

Probability Density Function

The probability density function for the exponential distribution is:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Here, λ is a parameter for the distribution, which is constant. If a random variable X has this distribution, we write $X \sim \text{Exp}(\lambda)$.

Mean and Variance

The mean of the exponential distribution is $\frac{1}{\lambda}$. The standard deviation is also $\frac{1}{\lambda}$. Therefore, the variance is $\frac{1}{\lambda^2}$. Let $X \sim \text{Exp}(\lambda)$, then:

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}, \quad \text{SD}(X) = \frac{1}{\lambda}.$$

These results can be derived using integration, but are not required for this unit as knowledge of integration from Mathematics A2 Unit 3, as well as knowledge of improper integrals from Further Mathematics A2 Unit 4 is needed.

Derivative and integral of e^{kx}

For this module, it is important to know the following derivative:

$$\frac{d}{dx}(e^{kx}) = ke^{kx}.$$

Conversely, we have that:

$$\int ke^{kx} dx = e^{kx} + C \quad \text{and} \quad \int e^{kx} dx = \frac{e^{kx}}{k} + C.$$

Cumulative Distribution Function

We derive the cumulative distribution function as follows:

$$\begin{aligned} F(x) &= \int_0^x \lambda e^{-\lambda t} dt \\ &= \left[\frac{\lambda e^{-\lambda t}}{-\lambda} \right]_0^x \\ &= \left[-e^{-\lambda t} \right]_0^x \\ &= -e^{-\lambda x} + e^0 \\ &= 1 - e^{-\lambda x}. \end{aligned}$$

Therefore, the cumulative distribution function is given by:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Further Mathematics AS Unit 2: Further Statistics A

Calculating Probabilities for Continuous Random Variables

The exponential distribution is an example of a continuous random variable. The rules for calculating probabilities using continuous random variables are different to those for discrete random variables. Let X be a continuous random variable and let a, b, c and d be constants.

Then:

- $P(X = a) = 0$
- $P(X \geq b) = P(X > b) = 1 - P(X \leq b)$
- $P(c \leq X \leq d) = P(c < X \leq d) = P(c \leq X < d) = P(c < X < d) = P(X \leq d) - P(X \leq c)$.

Worked Example 1

The interval, X seconds, between cars passing a point on a motorway follows an exponential distribution with probability density function

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

- State the mean and variance of X .
- State the cumulative distribution function.
- Calculate the probability that (give all answers to 3 significant figures):
 - The interval until the next car passes is between 1 and 2 seconds.
 - The interval until the next car passes is longer than 3 seconds.
 - The interval until the next car passes is less than 1.5 seconds.
- State a distribution that could be used to model the number of cars passing the point each second, giving the values of any parameters.

Solution:

- This is an exponential distribution with parameter $\lambda = 2$. Therefore, the mean is $E(X) = \frac{1}{\lambda} = \frac{1}{2}$ and $\text{Var}(X) = \frac{1}{\lambda^2} = \frac{1}{4}$.

- The cumulative distribution function is $F(x) = \begin{cases} 1 - e^{-2x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$

- Each probability can be calculated using the cumulative distribution function.
 - We calculate $P(1 \leq X \leq 2)$ as follows:

$$\begin{aligned} P(1 \leq X \leq 2) &= P(X \leq 2) - P(X \leq 1) \\ &= F(2) - F(1) \\ &= (1 - e^{-4}) - (1 - e^{-2}) \\ &= e^{-2} - e^{-4} \\ &= 0.117. \end{aligned}$$

- We calculate $P(X > 3)$ as follows:

$$\begin{aligned} P(X > 3) &= 1 - P(X \leq 3) \\ &= 1 - F(3) \\ &= 1 - (1 - e^{-6}) \\ &= e^{-6} \\ &= 0.00248. \end{aligned}$$

- We calculate $P(X < 1.5)$ as follows:

$$\begin{aligned} P(X < 1.5) &= F(1.5) \\ &= 1 - e^{-3} \\ &= 0.950. \end{aligned}$$

- The Poisson distribution with mean 2 could be used to model the number of cars passing the point each second.

Note: the exponential distribution says that the mean length of an interval is 0.5 seconds. This could correspond to an average of 2 cars passing the point every second, as specified in the corresponding Poisson distribution.

Further Mathematics AS Unit 2: Further Statistics A

Worked Example 2

The number of potholes on a random 1 km section of rural road has a Poisson distribution with mean 1.6. Let X km be the distance between successive potholes on the road.

(i) Show that, for $x \geq 0$,

$$P(X > x) = e^{-1.6x}.$$

(ii) Derive the cumulative distribution function (CDF) and probability density function (PDF) for X .

(iii) Calculate the mean and median distance between successive potholes. State which of the mean or median is largest, and what this suggests about the distribution of the distance between successive potholes.

Solution:

(i) The number of potholes on a randomly chosen 1 km stretch of rural road can be modelled by $Po(1.6)$. We know that X is the distance between two successive potholes. If $X > x$, this means there are no faults in the first x km. The number of the faults in the first x km can be modelled using $Y \sim Po(1.6x)$.

Hence:

$$P(X > x) = P(Y = 0) = \frac{e^{-1.6x} \times (1.6x)^0}{0!} = e^{-1.6x}.$$

(ii) The CDF can be obtained using $F(x) = P(X \leq x) = 1 - P(X > x) = 1 - e^{-1.6x}$.

Therefore:

$$F(x) = \begin{cases} 1 - e^{-1.6x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The PDF can then be derived by differentiating, i.e.

$$f(x) = F'(x) = 0 - (-1.6e^{-1.6x}) = 1.6e^{-1.6x}.$$

Hence:

$$f(x) = \begin{cases} 1.6e^{-1.6x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(iii) The mean distance is $\frac{1}{\lambda} = \frac{1}{1.6} = 0.625$ km = 625 m.

The median, m , satisfies $F(m) = 0.5$. Hence:

$$\begin{aligned} 1 - e^{-1.6m} &= 0.5 \\ e^{-1.6m} &= 0.5 \\ -1.6m &= \ln 0.5 \\ m &= \frac{\ln 0.5}{-1.6} \\ m &= 0.433. \end{aligned}$$

Therefore, the median distance is 433 m.

The median distance of 433 m is smaller than the mean distance of 625 m. This suggests that the distribution of the distance between successive potholes is skewed to the right (positively skewed).

Further Mathematics AS Unit 2: Further Statistics A

Worked Example 3

A monitor issues a warning signal when an action is needed as part of a production process. The interval, X hours, between successive signals follows an exponential distribution with parameter 0.08.

- (i) Find the probability that the interval between the next two signals is:
 - a. Between 10 and 20 hours;
 - b. Less than two hours;
 - c. Longer than 50 hours.
- (ii) State the mean and standard deviation of the intervals between successive signals.
- (iii) Following a warning signal, what is the longest time the production process could be left unsupervised whilst ensuring the probability of missing the next signal is less than 0.01?

Solution:

- (i) We have an exponential distribution with parameter $\lambda = 0.08$. The cumulative distribution function is given by $F(x) = \begin{cases} 1 - e^{-0.08x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$
 - a. $P(10 \leq X \leq 20) = F(20) - F(10) = (1 - e^{-1.6}) - (1 - e^{-0.8}) = 0.247$.
 - b. $P(X < 2) = F(2) = 1 - e^{-0.16} = 0.148$.
 - c. $P(X > 50) = 1 - P(X \leq 50) = 1 - e^{-4} = 0.0183$.
- (ii) $E(X) = \frac{1}{\lambda} = \frac{1}{0.08} = 12.5$ hours and $SD(X) = \frac{1}{\lambda} = 12.5$ hours.
- (iii) We wish to find the time, t , for which:

$$\begin{aligned} P(X < t) &= 0.01 \\ 1 - e^{-0.08t} &= 0.01 \\ e^{-0.08t} &= 0.99 \\ -0.08t &= \ln 0.99 \\ t &= \frac{\ln 0.99}{-0.08} \\ t &= 0.1256 \text{ hours} \\ t &= 7.54 \text{ minutes.} \end{aligned}$$

Therefore, the production process should be left for no longer than 7.54 minutes to ensure the probability of missing a signal is less than 0.01.

Summary of Key Points for the Exponential Distribution

If $X \sim \text{Exp}(\lambda)$, then:

1. The probability density function $f(x)$ is:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

2. The mean, variance and standard deviation are:

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}, \quad SD(X) = \frac{1}{\lambda}.$$

3. The cumulative distribution function $F(x)$ is:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Useful Results:

$$\frac{d}{dx}(e^{kx}) = ke^{kx}, \quad \int ke^{kx} dx = e^{kx} + C, \quad \int e^{kx} dx = \frac{e^{kx}}{k} + C.$$

Further Mathematics AS Unit 2: Further Statistics A

Questions

- The lifetime of a light bulb is X hours, where X can be modelled by an exponential distribution with parameter $\lambda = 0.0125$.
 - Find the mean and variance of the lifetime of a light bulb.
 - Find the probability that the lifetime of a bulb is:
 - less than 100 hours;
 - between 50 hours and 150 hours.
- The time, T seconds, between the arrival of successive vehicles at a zebra crossing on a road can be modelled by an exponential distribution with parameter $\lambda = 0.025$.
 - Write down the mean and the variance of T .
 - An elderly pedestrian takes 30 seconds to cross the road using this zebra crossing. Calculate the probability that:
 - no vehicle arrives whilst the pedestrian is crossing.
 - no vehicle arrives whilst the pedestrian makes two independent crossings.
 - A person starts crossing the road immediately after a vehicle has passed. How long should this person take to cross the road to ensure the probability of a vehicle arriving before they have crossed is less than 0.2?
- The interval, X metres, between consecutive minor faults on a roll of cloth has a distribution whose probability density function is given by:
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$
where λ is a positive constant.
 - It is given that $P(X < 1) = 2P(X > 2)$. Find the mean of the distribution.
 - Find the probability that the interval between two minor faults is greater than 3 metres.
- The probability that the lifetime, H , of a certain type of electrical component is more than h hours is given by $P(H > h) = e^{-\frac{h}{1000}}$, $h > 0$.
 - Calculate the probability that a randomly selected component has a lifetime of:
 - more than 1500 hours;
 - between 1000 and 2000 hours;
 - precisely 1200 hours.
 - Calculate the probability that three components, chosen at random, all have lifetimes of more than 1500 hours.
 - Derive the probability density function for H , and hence state the mean and variance of the component lifetime.
- The number of arrivals per hour at a toll booth has a Poisson distribution with mean 90. Let T minutes denote the time between successive arrivals at the toll booth.
 - Show that $P(T > t) = e^{-1.5t}$.
 - Use your answer in a) to find the CDF and PDF for T . Hence state the distribution of T .
 - Calculate the mean and standard deviation of T .
 - Determine the median time between successive arrivals at the toll booth.
 - Use your answers from c) and d) to comment on the distribution of time between successive arrivals at the toll booth.
- The number of flaws in a randomly chosen 100 metre length of ribbon is modelled by a Poisson distribution with mean 1.6. The random variable X metres is the distance between two successive flaws.

Further Mathematics AS Unit 2: Further Statistics A

- a) Show that the cumulative distribution function of X is given by:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-0.016x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

- b) Deduce that X has a negative exponential distribution, stating its mean.
c) Find the median and interquartile range for the distance between successive flaws.
d) Find the probability that there is a distance of at least 50 metres between two successive flaws.

Further Mathematics AS Unit 2: Further Statistics A

Question 1 Worked Solution

The lifetime of a light bulb is X hours, where X can be modelled by an exponential distribution with parameter $\lambda = 0.0125$.

- Find the mean and variance of the lifetime of a light bulb.
- Find the probability that the lifetime of a bulb is:
 - less than 100 hours;
 - between 50 hours and 150 hours.

a) The mean is $\frac{1}{\lambda} = \frac{1}{0.0125} = 80$ and the variance is $\frac{1}{\lambda^2} = \frac{1}{0.0125^2} = 6400$.

b) Firstly, we note that the CDF is given by:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-0.0125x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) $P(X < 100) = e^{-1.25} = 0.713$.

(ii)
$$\begin{aligned} P(50 \leq X \leq 150) &= P(X \leq 150) - P(X \leq 50) \\ &= (1 - e^{-1.875}) - (1 - e^{-0.625}) \\ &= e^{-0.625} - e^{-1.875} \\ &= 0.382. \end{aligned}$$

Question 2 Worked Solution

The time, T seconds, between the arrival of successive vehicles at a zebra crossing on a road can be modelled by an exponential distribution with parameter $\lambda = 0.025$.

- Write down the mean and the variance of T .
- An elderly pedestrian takes 30 seconds to cross the road using this zebra crossing. Calculate the probability that:
 - no vehicle arrives whilst the pedestrian is crossing;
 - no vehicle arrives whilst the pedestrian makes two independent crossings.
- A person starts crossing the road immediately after a vehicle has passed. How long should this person take to cross the road to ensure the probability of a vehicle arriving before they have crossed is less than 0.2?

a) The mean is $\frac{1}{\lambda} = \frac{1}{0.025} = 40$, and the variance is $\frac{1}{\lambda^2} = \frac{1}{0.025^2} = 1600$.

b) Firstly, we note that the CDF is given by:

$$F(t) = P(T \leq t) = \begin{cases} 1 - e^{-0.025t}, & t \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(i) The probability that no vehicle arrives is:

$$P(T > 30) = 1 - P(T \leq 30) = 1 - (1 - e^{-0.75}) = e^{-0.75} = 0.4724.$$

(ii) From the previous part, we deduce the probability of no vehicles arriving during two independent crossings is $(e^{-0.75})^2 = e^{-1.5} = 0.2231$.

c) We wish to find the value of t such that $P(T \leq t) = 0.2$. Therefore:

$$\begin{aligned} 1 - e^{-0.025t} &= 0.2 \\ e^{-0.025t} &= 0.8 \\ -0.025t &= \ln 0.8 \\ t &= \frac{\ln 0.8}{-0.025} \\ t &= 8.93. \end{aligned}$$

Further Mathematics AS Unit 2: Further Statistics A

Therefore, a person needs to cross in 8.93 seconds to ensure the probability of a vehicle arriving before they have crossed is less than 0.2.

Question 3 Worked Solution

The interval, X metres, between consecutive minor faults on a roll of cloth has a distribution whose probability density function is given by:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where λ is a positive constant.

- It is given that $P(X < 1) = 2P(X > 2)$. Find the mean of the distribution.
- Find the probability that the interval between two minor faults is greater than 3 metres.

a) The cumulative distribution function is:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We have $P(X < 1) = 1 - e^{-\lambda}$, and $P(X > 2) = 1 - P(X \leq 2) = 1 - (1 - e^{-2\lambda}) = e^{-2\lambda}$. Therefore, $1 - e^{-\lambda} = 2e^{-2\lambda}$. Multiplying by $e^{2\lambda}$ gives $e^{2\lambda} - e^\lambda = 2$, i.e. $e^{2\lambda} - e^\lambda - 2 = 0$.

Let $y = e^\lambda$. Then $y^2 - y - 2 = 0$ and so $(y - 2)(y + 1) = 0$.

Hence, $y = 2$ or $y = -1$. Therefore, $e^\lambda = 2$ or $e^\lambda = -1$. Since $e^\lambda = -1$ is impossible, we must have $e^\lambda = 2$, and so $\lambda = \ln 2$. Therefore, the mean is $\frac{1}{\lambda} = \frac{1}{\ln 2} = 1.4427$.

b) Using the CDF from a):

$$P(X > 3) = 1 - P(X \leq 3) = 1 - F(3) = 1 - (1 - e^{-3 \ln 2}) = e^{-3 \ln 2} = e^{\ln \frac{1}{8}} = \frac{1}{8}.$$

Question 4 Worked Solution

The probability that the lifetime, H , of a certain type of electrical component is more than h hours is given by $P(H > h) = e^{-\frac{h}{1000}}$, $h > 0$.

- Calculate the probability that a randomly selected component has a lifetime of:
 - more than 1500 hours;
 - between 1000 and 2000 hours;
 - precisely 1200 hours.
- Calculate the probability that three components, chosen at random, all have lifetimes of more than 1500 hours.
- Derive the probability density function for H , and hence state the mean and variance of the component lifetime.

a) Firstly, we note that $F(h) = P(H \leq h) = 1 - e^{-\frac{h}{1000}} = 1 - e^{-0.001h}$.

(i) $P(H > 1500) = e^{-\frac{1500}{1000}} = e^{-1.5} = 0.223$.

(ii) $P(1000 \leq H \leq 2000) = P(H \leq 2000) - P(H \leq 1000) = (1 - e^{-2}) - (1 - e^{-1}) = 0.233$.

(iii) $P(H = 1200) = 0$, since H is a continuous random variable.

b) From a)(i), we know that the probability of a randomly selected component having a lifetime of more than 1500 hours is $e^{-1.5}$. Hence, the probability of three randomly selected components all having lifetimes more than 1500 is $(e^{-1.5})^3 = e^{-4.5} = 0.0111$.

c) We find $f(h)$ from $F(h)$ through differentiation:

Further Mathematics AS Unit 2: Further Statistics A

$$f(h) = 0 - (-0.001)e^{-0.001h} = 0.001e^{-0.001h} \text{ for } h > 0.$$

i.e.

$$f(h) = \begin{cases} 0.001e^{-0.001h}, & h \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

This means that H follows an exponential distribution with parameter $\lambda = 0.001$. Hence, the mean lifetime is $\frac{1}{0.001} = 1000$ and the variance is $\frac{1}{0.001^2} = 1,000,000$.

Question 5 Worked Solution

The number of arrivals per hour at a toll booth has a Poisson distribution with mean 90. Let T minutes denote the time between successive arrivals at the toll booth.

- Show that $P(T > t) = e^{-1.5t}$.
- Use your answer in a) to find the CDF and PDF for T . Hence state the distribution of T .
- Calculate the mean and standard deviation of T .
- Determine the median time between successive arrivals at the toll booth.
- Use your answers from c) and d) to comment on the distribution of time between successive arrivals at the toll booth.

a) The number of arrivals per hour can be modelled by $Po(90)$. Therefore, the number of arrivals per minute can be modelled by $Po\left(\frac{90}{60}\right) = Po(1.5)$. We know that T is the time between two successive arrivals. If $T > t$, this means there are no arrivals for t minutes. The number of the faults in the first t minutes can be modelled using $Y \sim Po(1.5t)$. Hence:

$$P(T > t) = P(Y = 0) = \frac{e^{-1.5t} \times (1.5t)^0}{0!} = e^{-1.5t}.$$

b) The CDF can be obtained using $F(t) = P(T \leq t) = 1 - P(T > t) = 1 - e^{-1.5t}$. Therefore:

$$F(t) = \begin{cases} 1 - e^{-1.5t}, & t \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

The PDF can then be derived by differentiating: $f(t) = F'(t) = 0 - (-1.5e^{-1.5t}) = 1.5e^{-1.5t}$.

$$\text{Hence, } f(t) = \begin{cases} 1.5e^{-1.5t}, & t \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

This is the PDF for an exponential distribution with parameter 1.5. Therefore, $T \sim \text{Exp}(1.5)$.

c) $E(T) = \frac{1}{\lambda} = \frac{1}{1.5} = \frac{2}{3}$ and $SD(T) = \frac{1}{\lambda} = \frac{2}{3}$. Therefore, the mean and standard deviation of T is $\frac{2}{3} \times 60 = 40$ seconds.

d) The median, m , satisfies $F(m) = 0.5$. Hence:

$$\begin{aligned} 1 - e^{-1.5m} &= 0.5 \\ e^{-1.5m} &= 0.5 \\ -1.5m &= \ln 0.5 \\ m &= \frac{\ln 0.5}{-1.5} \\ m &= 0.462. \end{aligned}$$

Therefore, the median time is $0.462 \times 60 = 27.726$ seconds.

e) Since the median time is less than the mean time, this suggests that the distribution of the time between successive arrivals at the toll booth is skewed to the right (positively skewed).

Further Mathematics AS Unit 2: Further Statistics A

Question 6 Worked Solution

The number of flaws in a randomly chosen 100 metre length of ribbon is modelled by a Poisson distribution with mean 1.6. The random variable X metres is the distance between two successive flaws.

a) Show that the cumulative distribution function of X is given by:

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-0.016x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

b) Deduce that X has a negative exponential distribution, stating its mean.

c) Find the median and interquartile range for the distance between successive flaws.

d) Find the probability that there is a distance of at least 50 metres between two successive flaws.

a) The number of flaws in 100 metre of ribbon can be modelled by $Po(1.6)$, hence the number of flaws in 1 metre of ribbon can be modelled by $Po(0.016)$. We know that X is the distance between two successive serious faults. If $X > x$, this means there are no faults in the first x metres. The number of the faults in the first x metres can be modelled using $Y \sim Po(0.016x)$.

Hence:

$$P(X > x) = P(Y = 0) = \frac{e^{-0.016x} \times (0.016x)^0}{0!} = e^{-0.016x}.$$

The CDF can be obtained using $F(x) = P(X \leq x) = 1 - P(X > x) = 1 - e^{-0.016x}$.

Therefore:

$$F(x) = \begin{cases} 1 - e^{-0.016x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

b) The PDF can then be derived by differentiating, i.e.

$$f(x) = F'(x) = 0 - (-0.016e^{-0.016x}) = 0.016e^{-0.016x}.$$

Hence:

$$f(x) = \begin{cases} 0.016e^{-0.016x}, & x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

This is the PDF for an exponential distribution with parameter 0.016, therefore $X \sim \text{Exp}(0.016)$. The mean is $\frac{1}{\lambda} = \frac{1}{0.016} = 62.5$.

c) The median, m , satisfies $F(m) = 0.5$. Hence:

$$\begin{aligned} 1 - e^{-0.016m} &= 0.5 \\ e^{-0.016m} &= 0.5 \\ -0.016m &= \ln 0.5 \\ m &= \frac{\ln 0.5}{-0.016} \\ m &= 43.32. \end{aligned}$$

Therefore, the median is 43.32 metres.

The lower quartile, l , satisfies $F(l) = 0.25$. Hence:

$$\begin{aligned} 1 - e^{-0.016l} &= 0.25 \\ e^{-0.016l} &= 0.75 \\ -0.016l &= \ln 0.75 \\ l &= \frac{\ln 0.75}{-0.016} \\ l &= 17.98. \end{aligned}$$

Further Mathematics AS Unit 2: Further Statistics A

The upper quartile, u , satisfies $F(u) = 0.25$. Hence:

$$\begin{aligned}1 - e^{-0.016u} &= 0.75 \\e^{-0.016u} &= 0.25 \\-0.016u &= \ln 0.25 \\u &= \frac{\ln 0.25}{-0.016} \\u &= 86.64.\end{aligned}$$

Therefore, the interquartile range is $u - l = 86.64 - 17.98 = 68.66$ metres.

d) We calculate:

$$\begin{aligned}P(X > 50) &= 1 - P(X \leq 50) \\&= 1 - F(50) \\&= 1 - (1 - e^{-0.016 \times 50}) \\&= e^{-0.8} \\&= 0.449.\end{aligned}$$