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:

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

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:

$$E(X) = \int_a^b x f(x)\,\mathrm{dx}, \qquad E(X^2) = \int_a^b x^2 f(x)\,\mathrm{dx}, \qquad E(g(X)) = \int_a^b g(x) f(x)\,\mathrm{dx}.$$

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

$$\mathrm{Var}(X) = E(X^2) - \big(E(X)\big)^2, \qquad \mathrm{SD}(X) = \sqrt{Var(X)}.$$

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

$$F(x) = \int_a^x f(t)\,\mathrm{dt}.$$

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:

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

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.

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 $Var(Y) = \frac{1}{\lambda^2}$.
- Use the exponential distribution as a model for intervals between events
  - learners will be expected to know that $\frac{d}{dx}\left(e^{kx}\right) = ke^{kx}$.

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:

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

Cumulative Distribution Function

We derive the cumulative distribution function as follows:

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

Therefore, the cumulative distribution function is given by:

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

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

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

(i) State the mean and variance of X.

(ii) State the cumulative distribution function.

(iii) Calculate the probability that (give all answers to 3 significant figures):
  a. The interval until the next car passes is between 1 and 2 seconds.
  b. The interval until the next car passes is longer than 3 seconds.
  c. The interval until the next car passes is less than 1.5 seconds.

(iv) 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:

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

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

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

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

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

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

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

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

(iv) 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.

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

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

Try yourself:

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.

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

Example:

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

a) Find the mean and variance of the lifetime of a light bulb.
b) Find the probability that the lifetime of a bulb is:
   (i) less than 100 hours;
   (ii) 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 $\lambda = 0.025$.

a) Write down the mean and the variance of T.
b) An elderly pedestrian takes 30 seconds to cross the road using this zebra crossing. Calculate the probability that:
   (i) no vehicle arrives whilst the pedestrian is crossing;
   (ii) no vehicle arrives whilst the pedestrian makes two independent crossings.
c) 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:

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

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.

a) It is given that $P(X < 1) = 2P(X > 2)$. Find the mean of the distribution.

b) 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 $P(X > 2) = 1 - P(X \leq 2) = 1 - \left(1 - e^{-2\lambda}\right) = e^{-2\lambda}$. 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):

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