Exponential Distribution Cheat Sheet (A Level Only)

The exponential distribution is a continuous distribution used to model probabilities of waiting time until the first event or intervals between events. It is defined by the average number of events within a unit interval and is usually written as $X \sim \exp(\lambda)$. Examples include:

- Time between phone calls received
- Distance between potholes in a road
- Time between collisions of particles in a gas

The conditions required for a situation to be modelled with the exponential distribution are:

- Events occur independently
- The average rate at which events occur is constant
- Events occur one at a time

In fact, these are the same conditions required by the Poisson distribution. The exponential and Poisson distributions are related in the following way: The Poisson distribution gives the probability of a number of events occurring within a set interval whereas the exponential distribution gives the probability of an interval between events. This relationship is explored mathematically:

Suppose $X \sim \text{Po}(\mu)$ represents the number of busses that arrive in 1 hour so $P(X = x) = \frac{e^{-\mu}\mu^x}{x!}$. Let Y be the number of busses that arrive within t hours such that $Y \sim \text{Po}(\mu t)$, then $P(Y = y) = \frac{e^{-\mu t}(\mu t)^y}{y!}$ and $P(Y = 0) = \mu e^{-\mu t}$, but $Y = 0$ for t units of time is the same as $T > t$, where T is the time between intervals. Therefore, $T \sim \exp(\mu)$.

Finding Probabilities with the Exponential Distribution

If $X \sim \exp(\lambda)$ then the probability density function is given by:

$$f(x) = \lambda e^{-\lambda x}$$

This can be integrated to find the cumulative probability distribution (which gives the probability of an interval between 0 and x):

$$F(x) = \int_0^x f(x)\,dx = \int_0^x \lambda e^{-\lambda x}\,dx = -e^{-\lambda x}\Big|_0^x = 1 - e^{-\lambda x}$$

Example 1: Phone calls are received at a consistent, average rate of 3 per hour. Assume that the time intervals between calls are given by an exponential distribution. a.) Calculate the probability of waiting less than 15 minutes for the first call. b.) Calculate the probability of waiting more than 30 minutes for the first call. c.) Given that a call has not been received for 1 hour, what is the probability of waiting at least another 30 minutes for the next call?

a.) The cumulative probability function gives the probability of the interval between calls being between 0 and t (T/t is used instead of X/x as we are dealing with times). This makes it the easier way to do these types of problems. Here, $x = \frac{1}{4}$ because the set up implies that x is in units of hours.	$T \sim \exp(3) \Rightarrow F(t) = 1 - e^{-3t}$ $P\left(T < \frac{1}{4}\right) = F\left(\frac{1}{4}\right) = 1 - e^{-\frac{3}{4}}$ ≈ 0.528
b.) More than 30 minutes is the same as not less than 30 minutes.	$P(T > 0.5) = 1 - P(T < 0.5)$ $= 1 - \left(1 - e^{-\frac{3}{2}}\right) = e^{-\frac{3}{2}}$ ≈ 0.223
c.) Here, the conditional probability formula must be used.	$P(T > 1.5 \mid T > 1) = \frac{P(T > 1.5) \cap P(T > 1)}{P(T > 1)}$
But $T > 1.5$ requires $T > 1.5$ so $P(T > 1.5) \cap P(T > 1) = P(T > 1.5)$	$P(T > 1.5 \mid T > 1) = \frac{P(T > 1.5)}{P(T > 1)}$ $= \frac{e^{-\frac{9}{2}}}{e^{-3}} = e^{-\frac{3}{2}} \approx 0.223$

Parts b and c of the above example have the same answer; because events occur independently, intervals between events must also be independent. This means that information about what has already occurred is irrelevant. The solution could have been found much quicker using this fact.

Mean and Variance

The mean of a continuous probability distribution is given by:

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

For an exponential distribution, using the fact that $x > 0$, this becomes:

$$E(X) = \int_0^{\infty} x\lambda e^{-\lambda x}\,dx.$$

The integral can be evaluated by parts:

$$E(X) = -xe^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} -e^{-\lambda x}\,dx$$

$$= (0 - 0) - \frac{1}{\lambda}e^{-\lambda x}\Big|_0^{\infty} = \frac{1}{\lambda}$$

$$\Rightarrow E(X) = \frac{1}{\lambda}$$

This result is expected. If $Y \sim \text{Po}(\lambda)$, then there are λ events distributed randomly within a unit time. Therefore, it is expected that on average the events will be separated by $\frac{1}{\lambda}$ units of time, i.e. $E(X) = \frac{1}{\lambda}$.

Variance is given by:

$$Var(X) = E(X^2) - E(X)^2,$$

And $E(X^2)$ is given by:

$$E(X^2) = \int_0^{\infty} x^2 \lambda e^{-\lambda x}\,dx.$$

Integrating by parts gives:

$$E(X^2) = -x^2 e^{-\lambda x}\Big|_0^{\infty} - \int_0^{\infty} -2xe^{-\lambda x}\,dx,$$

This integral has already been evaluated above in the calculation of $E(X)$.

$$E(X^2) = (0 - 0) + \frac{2}{\lambda}E(X) = \frac{2}{\lambda^2}$$

Therefore:

$$Var(X) = \frac{1}{\lambda^2} \Rightarrow \sigma = \frac{1}{\lambda}$$

Just as with the Poisson distribution, the standard deviation and the mean are equal.

Example 2: Let N be the number of beta particles emitted by a radioactive material in a minute. a.) Given $N \sim \text{Po}(2)$, write down the distribution of T, the time intervals between consecutive emissions. b.) Calculate the probability that the first emission is between 1 and 2 minutes after observation. c.) Write down the mean and variance of T.

a.) Since N follows a Poisson distribution, T must follow an exponential distribution.	$T \sim \exp(2)$
b.) Either the probability density function or the cumulative probability function could be used. Here, the probability density function is used (it is useful to be able to do these calculations in both ways).	$P(2 > T > 1) = \int_1^2 2e^{-2t}\,dt$ $= -e^{-2t}\Big\|_1^2$ $= -(e^{-4} - e^{-2})$ ≈ 0.117
c.) The results from above can be used directly.	$E(T) = \frac{1}{2}$ $Var(T) = \frac{1}{4}$

Example 3: The continuous random variable X is modelled by an exponential distribution with mean μ. Find the interquartile range of X.

The distribution of X can be written down using $E(X) = \mu$.	$X \sim \exp\left(\frac{1}{\mu}\right)$
The n^{th} percentile can be found by solving $\int_0^a f(x)\,dx = \frac{n}{100}$ for a. The lower quartile is found first.	$\int_0^a \frac{1}{\mu} e^{-\frac{1}{\mu}x}\,dx = 0.25$ $-e^{-\frac{1}{\mu}x}\Big\|_0^a = 0.25$ $1 - e^{-\frac{1}{\mu}a} = 0.25$ $\frac{a}{\mu} = -\ln 0.75$ $a = -\mu \ln 0.75$
The upper quartile is found in the same way, by solving $\int_0^b f(x)\,dx = 0.75$.	$\int_0^b \frac{1}{\mu} e^{-\frac{1}{\mu}x}\,dx = 0.75$ $-e^{-\frac{1}{\mu}x}\Big\|_0^b = 0.75$ $1 - e^{-\frac{1}{\mu}b} = 0.75$ $b = -\mu \ln 0.25$
The interquartile range can now be calculated as the difference between b and a.	$b - a = \mu \ln 3$