Continuous random variables: the uniform and the exponential distribution

Chrysafis Vogiatzis

Lecture 7

Learning objectives

After these lectures, we will be able to:

- Calculate probabilities of continuous random variables using their probability distribution and cumulative distribution functions.
- Give examples of uniform and exponentially distributed random variables.
- Recall when to and how to use:
  - uniformly distributed random variables.
  - exponentially distributed random variables.
- Define the memorylessness property and apply it exponentially distributed random variables.
- Use Poisson random variables and exponential random variables and provide examples of their relationship.

Motivation: continuous vs. discrete random variables

Guess which number I am thinking between 0 and 10 is a tricky proposition. If asked to do so in integer numbers (that is, 0 or 1 or 2...) then it is difficult to guess correctly, but not nearly impossible: we'd get a probability of 1 over 11 or a little more than 9%. On the other hand, when asked to do so with *any* number...

Motivation: Big in Japan

In an earlier example, we discussed the probability of seeing a certain number of earthquakes in the Kanto region of Tokyo in Japan. What if though we are interested in the timing of the next earthquake? How would we go about modeling this using continuous random variables?

Continuous random variables

Let X be a **continuous** random variable. Recall here that a continuous random variable is allowed to take any **real value** within some interval, say in $[a, b]$. Hence, there is an *infinite* number of possible outcomes associated with each continuous random variable!

Definition 1 *We define the* **probability distribution function** *(pdf)* $f(x)$ [1] *of a continuous random variable X as the "relative likelihood" that X will be equal to a specific value x. This definition is a little open-ended, so we will address it more carefully shortly.*

[1] Contrast with the definition of a probability mass function (pmf) $p(x)$ of a discrete random variable here...

We again need to be careful with one item here:

- The actual probability that a continuous random variable X is exactly equal to some value x is 0! [2]

[2] Surprised?

This last note probably changes the way we need to discuss continuous probabilities. What if, instead of asking for the probability that continuous random variable X is exactly equal to some value x, we focus on the probability that continuous random variable X *belongs to some interval* of values?

Continuous random variables

Instead of the probability that:

- the average temperature tomorrow is exactly 78.3 Fahrenheit;
- the next bus passes in exactly 3 minutes and 25 seconds;
- the error of an ammeter (used to measure the current in a circuit) is exactly 0.1 A;

we may ask for the probability that:

- the average temperature tomorrow is between 78 and 79 Fahrenheit;
- the next bus passes between 3 and 4 minutes from now;
- the error of an ammeter is within 0.1 A.

This gives rise to the need for defining and using the **cumulative distribution function**. First, though, let us provide a different definition for the probability density function of a continuous random variable X.

Definition 2 *A random variable is* **continuous** *if it can take uncountably* *many values such that there exists some function $f(x)$ called a* **probability density function** *defined over real values $(-\infty, +\infty)$ such that:*

- $f(x) \geq 0$;
- $\int_{-\infty}^{+\infty} f(x)dx = 1$;
- $P(X \in B) = \int_B f(x)dx$.

The last property essentially states that to find the probability that a random variable X belongs to some interval B, then we need to take the integral of the probability density function of X, $f(x)$, over the interval B.

Continuous random variables

What is the probability that random variable X with pdf $f(x)$ is between 0 and 10?

$$P(0 \leq X \leq 10) = \int_0^{10} f(x)dx.$$

Due to the continuous nature of random variable X, and due to the fact that $P(X = x) = 0$, for any value x, we also get that:

$$P(0 \leq X \leq 10) = P(0 < X < 10) = P(0 < X \leq 10) = P(0 \leq X < 10).$$

Assume that continuous random variable X is distributed with probability density function $f(x)$ in $[0, \infty)$. What is:

a) the probability that X is between 2 and 5?

b) the probability that X is below 5 or above 10?

c) the probability that X is exactly equal to 5?

Definition 3 *We define the* **cumulative distribution function** *of a continuous random variable as the probability that it takes* up to *a value a, i.e.,*

$$F(a) = P(-\infty < X \leq a) = \int_{-\infty}^{a} f(x)dx.$$

By definition, $F'(x) = f(x)$: the derivative of the cdf gives us the pdf. Moreover, what we observed for discrete random variables is also true here and $P(a \leq X \leq b) = F(b) - F(a)$.

Timing chemical reactions

The time until a chemical reaction is over is measured in milliseconds (ms). The probability that the reaction is over by time x is given by the following cdf:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-0.01x}, & x \geq 0. \end{cases}$$

Answer the following questions.

1. What is the pdf?
2. What proportion of chemical reactions are performed in less than or equal to 200 ms?
3. What proportion of chemical reactions take more than or equal to 100 ms and less than or equal to 200 ms?

1. For the pdf, by definition we have that

$$f(x) = F'(x) = \begin{cases} 0 \\ \left(1 - e^{-0.01x}\right)' \end{cases} = \begin{cases} 0, & x < 0 \\ 0.01 \cdot e^{-0.01x}, & x \geq 0 \end{cases}$$

2. We need to calculate $F(200)$:

$$P(x \leq 200) = F(200) = 1 - e^{-2} = 0.8647.$$

3. We now need $P(100 \leq x \leq 200)$:

$$P(100 \leq x \leq 200) = \int_{100}^{200} f(x)dx = \int_{100}^{200} 0.01 \cdot e^{-0.01x} dx =$$
$$= e^{-1} - e^{-2} = 0.2325.$$

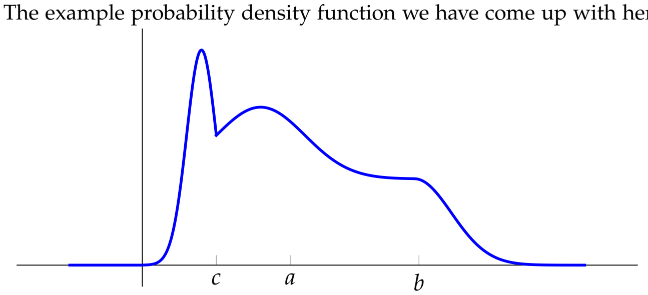
To answer this last part, we could have also used the fact that $P(a \leq X \leq b) = F(b) - F(a)$:

$$P(100 \leq x \leq 200) = F(200) - F(100) = 0.2325.$$

We may also represent the cdf visually. If we plot $f(x)$ (the pdf), then the cdf is the area under the curve. For example, consider the $f(x)$ plotted in Figure 1. It could be a valid pdf as it satisfies $f(x) \geq 0$, and it also can be shown to satisfy $\int_{-\infty}^{+\infty} f(x)dx = 1$, even though it would be impossible to do so without knowing the exact function.

That said, we may observe that it is equal to 0 for small enough and large enough numbers, which indicates that the integral from minus to plus infinity is equal to a finite number (i.e., the area under the curve is a finite number).

Figure 1: The example probability density function we have come up with here.



Then, in Figures 2 and 3, we show what the cdf appears to be visually.

Figure 2: How to visually represent the cumulative probability distribution $F(x)$ (here, we specifically present $F(c)$).

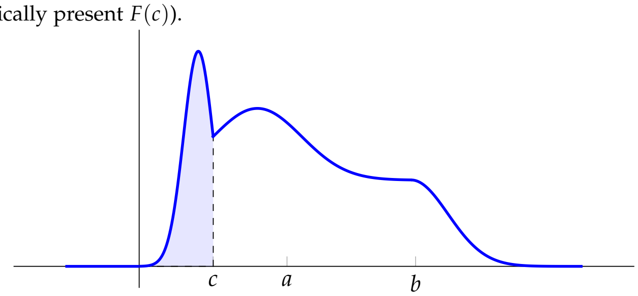
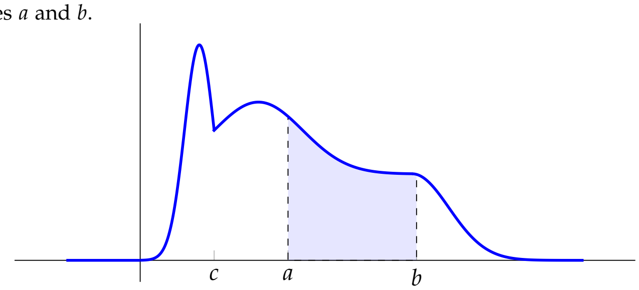


Figure 3: How to visually represent the probability that a random variable is between two values a and b.



Since we discussed the validity of a pdf, let's see an example of how we could use that.

Valid pdf?

Assume a continuous random variable taking values between 0 and 10 with a pdf of $f(x) = c \cdot x$. What is c?

First of all, c has to be nonnegative ($c \geq 0$), otherwise $f(x)$ may become negative, which is not allowed. We then employ the fact that $\int_{-\infty}^{+\infty} f(x)dx = 1$:

$$\int_{-\infty}^{+\infty} f(x)dx = 1 \implies \int_{0}^{10} cxdx = 1 \implies c \cdot \frac{x^2}{2}\Big|_0^{10} = 1 \implies c = 0.02.$$

Are the following valid pdfs?

- $f(x) = 0.01, 0 \leq x \leq 100$?
- $f(x) = \lambda \cdot e^{-\lambda \cdot x}, x \geq 0$?
- $f(x) = \lambda \cdot e^{-\mu \cdot x}, x \geq 0$ if we are told that $\lambda \neq \mu$?

The uniform distribution

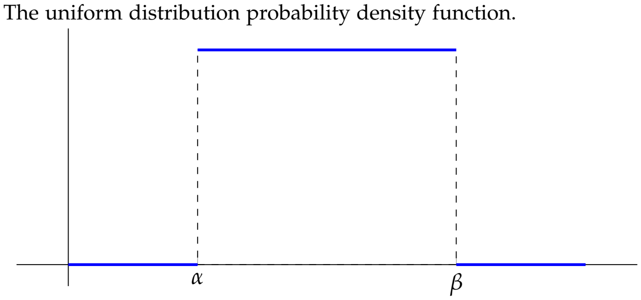
We begin this time from the simplest continuous distribution, the uniform distribution. In essence, it mimics its discrete counterpart, where everything is equally likely. However, since we are discussing continuous random variables, this implies that all values of $f(x)$ are equal, having equal relative likelihood. Its pdf and cdf are shown next.

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha}, & \text{if } \alpha \leq x \leq \beta \\ 0, & \text{otherwise.} \end{cases}$$

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 0, & \text{if } x < \alpha \\ \frac{x-\alpha}{\beta-\alpha}, & \text{if } \alpha \leq x \leq \beta \\ 1, & \text{if } x > \beta \end{cases}$$

Visually, the uniform distribution is presented in Figure 4.

Figure 4: The uniform distribution probability density function.



Totally random buses

Assume you are told that the next bus will arrive at any point in the next 5 to 15 minutes. Hence, in this case, the time until the next bus shows up is uniformly distributed. Then, what is the probability the bus arrives before:

a) 2′? b) 7′? c) 10′? d) 18′?

Note that here we have that $\alpha = 5$, $\beta = 15$. Thus:

a) 2′: $x = 2 < \alpha \implies F(2) = 0$.

b) 7′: $x = 7 \implies F(7) = \frac{7-5}{15-5} = 0.2$.

c) 10′: $x = 10 \implies F(10) = \frac{10-5}{15-5} = 0.5$.

d) 18′: $x = 18 > \beta \implies F(18) = 1$.

We visually present the probability for the bus to arrive in the next 10 minutes as an area under the curve in Figure 5.

Figure 5: The area under the curve for the probability of the bus arriving in 10′ from the example. The area is marked in green. We can tell that it is half the total area under the curve of the pdf, and hence corresponds to a probability of 50%.

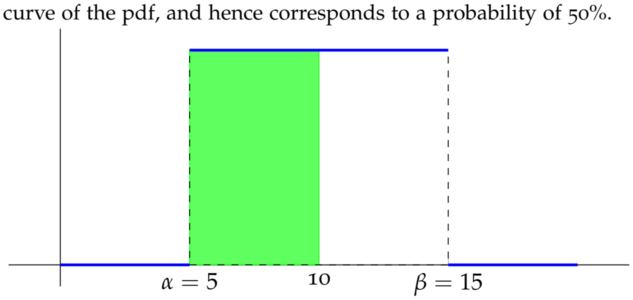
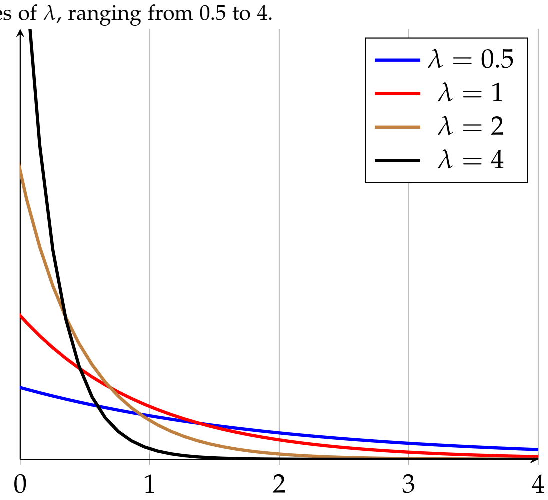


Figure 6: The exponential distribution probability density function visualized for different values of λ, ranging from 0.5 to 4.



The exponential distribution

It is time to move to one of the most important and consequential probability distributions. The exponential distribution takes its name from the fact that it is based on the exponential function. This is shown when considering its pdf and cdf [3]:

[3] In this lecture's worksheet, you are asked to derive $F(x)$, so brush up on your integration skills!

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

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

We also present the two functions visually in Figures 10 and 7. Formally, the exponential distribution is defined as in Definition 4.

Definition 4 (The exponential distribution) *A continuous random variable X defined over the interval of $[0, \infty)$ is exponentially distributed if it has probability density function given by*

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

where $\lambda > 0$ is a parameter. We sometimes write that $X \sim Exp(\lambda)$ if it follows the exponential distribution with rate λ. [4]

[4] Where have we seen *rates* before?

One of the many applications that the exponential distribution sees in practice has to do with quantifying the probability of **the time until the next event**. When events happen with some rate λ,

Figure 7: The exponential distribution cumulative density function visualized for different values of λ, ranging from 0.5 to 4.

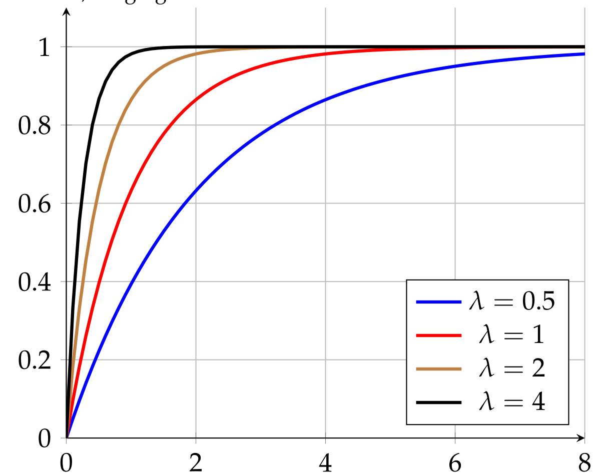


Figure 8: Taken from *The Simpsons*.

Figure 9: Taken from *The Office*.

we can quantify the risk or chance that the next event will happen within some time interval using the exponential distribution. Let us motivate this better with an example.

No accident in x days

How many days until the next accident? In many real-life cases, we assume that the **time to the next event** follows an exponential distribution. Assume that you work in a facility that has typically a rate of $\lambda = 2$ major accidents per year.

What is the probability the next accident happens in the next year? What is the probability that the next accident happens in the next 1 month?

Before we answer the question, we need to address the déjà vue feeling we may be experiencing. **We have seen** this family of questions before! When dealing with the **Poisson random variables**, we were talking about rates and about the probability of having a certain number of events within some time interval. The relationships do not stop here: both distributions make use of the exponential function, and both distributions rely on rates $\lambda > 0$. This brings us to the next subsection: how are Poisson random variables and exponentially distributed random variables related?

The exponential distribution and the Poisson distribution

We have already motivated the fact that these two appear to be "sibling" distributions. Let us go back to a question we addressed in a previous worksheet. [5] We repeat this here for convenience.

[5] Recall Lecture 6 Worksheet and, specifically, Problem 10.

Lecture 6 worksheet: Problem 10 repeat

We saw in class the probability mass function for a Poisson distributed random variable with rate λ. Assume that $\lambda = 3$ per year. What is the probability that there will be no events in the next year? Can you say that this means that the next event will happen more than a year from now? Let T be the time of the next event: what is $P(T > 1 \text{ year})$?

Let X be the number of events during the next year. Then, we have that:

$$P(T > 1 \text{ year}) = P(X = 0) = e^{-\lambda} \cdot \lambda^0/0! = e^{-3} = 0.05.$$

Note that we could also find the probability that the next event does happen during the next year:

$$P(T \leq 1 \text{ year}) = 1 - P(X = 0) = 1 - e^{-\lambda} \cdot \lambda^0/0! = 1 - e^{-3} = 0.95.$$

Let us put the previous result in perspective. The time to the next event is exponentially distributed if the number of events is distributed as a Poisson random variable! The full relationship between the exponential and the Poisson distributions is presented in tabular form in Table 1.

Table 1: The relationship between an exponentially distributed and a Poisson distributed random variable.

Exponential distribution	Poisson distribution
Rate λ	Rate λ
Time to next event	Number of events within some time
Continuous, $[0, \infty)$	Discrete $\{0, 1, \ldots\}$

No accident in x days (cont'd)

How many days until the next accident? In many real-life cases, we assume that the **time to the next event** follows an exponential distribution. Assume that you work in a facility that has typically a rate of $\lambda = 2$ major accidents per year.

What is the probability the next accident happens in the next year? What is the probability that the next accident happens in the next 1 month?

Let X be the time until the next accident. Recall that $\lambda = 2$ per year, or equivalently $\lambda = 2$ per 12 months. We then have:

1. $P(X \leq 1 \text{ year}) = F(1) = 1 - e^{-2\cdot 1} = 0.8647$.
2. $P(X \leq 1 \text{ month}) = F(1) = 1 - e^{-\frac{2}{12}\cdot 1} = 0.1535$.

Note how we used $\lambda = 2/12$ for the second question.

Historically, an emergency room after hours (10pm–6am) sees 48 patient requests every 8 hours. The time until the next patient arrives is exponentially distributed with that rate.

a) What is the probability that the next patient arrives in the next 10 minutes?

b) What is the probability there are 5 patients during the next hour (60 minutes)?

Memorylessness

Definition 5 (Memoryless random variables) *A random variable X is said to be* memoryless *(without memory) if:*

$$P(X > s + t | X > s) = P(X > t).$$

In English, the memorylessness property states that information available to us for what has happened so far does not alter our perception for the future. Let us see that with an example.

Memorylessness and the exponential distribution

A car transmission fails in time that is exponentially distributed with a rate of 1 every 80,000 miles. What is the probability that the transmission does not fail within its first 40,000 miles?

We need $P(T > 40{,}000 \text{ miles})$, when knowing that T is exponentially distributed with $\lambda = 1/80000$. We have:

$$P(T > 40{,}000 \text{ miles}) = 1 - P(T \leq 40{,}000 \text{ miles}) = 1 - F(40000) =$$
$$= e^{-\frac{1}{80000} \cdot 40000} = e^{-0.5} = 0.6065 = 60.65\%.$$

The next part is left as an exercise to the reader. [6]

[6] See also the Worksheet of Lecture 7!

Memorylessness and the exponential distribution

For the car from the previous example, assume we know that its transmission has been working for 80,000 miles already. What is the probability that the transmission does not fail in the next 40,000 miles?

From our answer to the previous question, we get to the point we wanted to make:

Exponentially distributed random variables are memoryless.

[7]

[7] Again, the proof is something we will do in the Worksheet of Lecture 7.

Memorylessness and the uniform distribution

Assume that the time (in minutes) until the next customer shows up is uniformly distributed in $[0, 60]$. What is the probability the next customer shows up after the first minute? What is the probability the next customer shows up between the 59th and 60th minute, given that no customer has shown up until the 58th minute?

Let T be the time to the next customer arrival. In the first part, we are looking for $P(T > 1)$:

$$P(T > 1) = \frac{59}{60}.$$

Now, note that this answer is significantly different than the answer we would get by calculating $P(T > 59 | T > 58)$, which is:

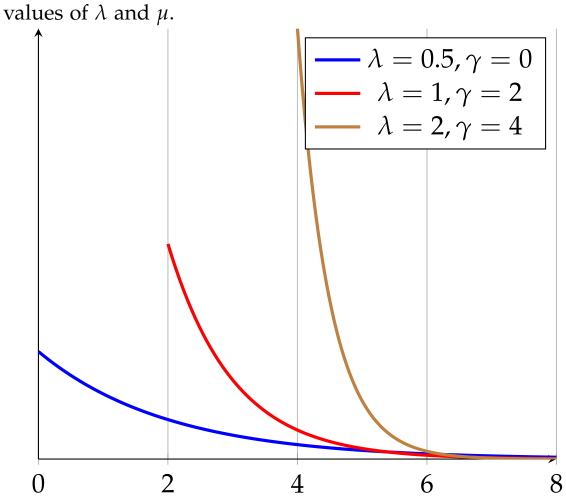
$$P(T > 59|T > 58) = \frac{P(T > 59)}{P(T > 58)} = \frac{1/60}{2/60} = \frac{1}{2}.$$

From the example, we may deduce that the uniform distribution is not memoryless. Memorylessness is a rare property (actually, the only **continuous distribution** to possess the property is the exponential).

General exponential distribution

The exponential distribution we have discussed so far only requires a single parameter: the rate $\lambda > 0$. In some cases, we may be interested in a "shifted version" of the exponential distribution (as in Figure **??**).

Figure 10: The general exponential distribution probability density function visualized for different values of λ and μ.



We call the "shift" the location parameter and we represent it with $\gamma > 0$ [8]. Hence, the general exponential distribution is a two-parameter distribution requiring the presence of a rate $\lambda > 0$ and a location parameter $\gamma > 0$, rendering the pdf and the cdf of the general exponential distribution as follows:

[8] Some textbooks employ μ instead of γ. In this class, I will try to reserve μ for something different.

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

$$F(x) = \begin{cases} 1 - e^{-\lambda(x-\gamma)}, & \text{if } x \geq \gamma \\ 0, & \text{if } x < \gamma. \end{cases}$$

Doctor FaceTime

A doctor sees patients in time that is exponentially distributed with rate 1 patient every 40 minutes. However, every patient will spend at least 10 minutes logged in the appointment while they are answering survey questions. In essence, this means that no patient will leave before these 10 minutes are up. What is the probability the next patient is seen for:

a) more than 30 minutes?

b) more than 1 hour?

c) more than 2 hours?

Let T be the time the next patient will require: T is exponentially distributed with rate $\lambda = 1/40$ minutes and $\gamma = 10$ minutes. Then, we have:

a) $P(T > 30 \text{ minutes}) = 1 - P(T \leq 30 \text{ minutes}) = 1 - F(30) = e^{-\frac{1}{40}\cdot(30-10)} = e^{-0.5} = 0.607$.

b) $P(T > 1 \text{ hour}) = 1 - F(60) = e^{-\frac{1}{40}\cdot(60-10)} = e^{-5/4} = 0.287$.

c) $P(T > 2 \text{ hours}) = 1 - F(120) = e^{-\frac{1}{40}\cdot(120-10)} = e^{-11/4} = 0.064$.

Summary

We have seen a lot of material in this lecture. To help place everything together, we provide in Table 2 a summary of all results from Lecture 7. You could again refer to these (and the keyword that follow) for all information about the uniform and the exponential probability distributions.

Table 2: A summary of all results from Lecture 7.

Name	**Parameters**	**Values**	**pmf**
Uniform	–	$[a, b]$	$f(x) = \begin{cases} \frac{1}{\beta-\alpha}, & \text{if } \alpha \leq x \leq \beta \\ 0, & \text{otherwise.} \end{cases}$
Exponential	$\lambda > 0$	$[0, +\infty)$	$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0. \end{cases}$
General exponential	$\lambda, \gamma > 0$	$[\gamma, +\infty)$	$f(x) = \begin{cases} \lambda \cdot e^{-\lambda(x-\gamma)}, & \text{if } x \geq \gamma \\ 0, & \text{if } x < \gamma. \end{cases}$

Some keywords that might help you narrow down your search. For convenience we also include the Poisson distribution, seeing as it is related to the exponential distribution.

Uniform: "equally probable"; "$f(x) = c$, where c is a constant".

Exponential: "time to next event"; "rate of events"; "memoryless distribution/memorylessness property".

General exponential: "time to next event"; "rate of events"; "location parameter"; "no event before a certain point".

Poisson: "number of events in an interval"; "rate of events".