Continuous random variables: the Gamma distribution and the normal distribution

Chrysafis Vogiatzis

Lecture 8

Learning objectives

After these lectures, we will be able to:

- Give examples of Gamma, Erlang, and normally distributed random variables.
- Recall when to and how to use:
  - Gamma and Erlang distributed random variables.
  - normally distributed random variables.
- Recognize when to use the exponential, the Poisson, and the Erlang distribution.
- Use the standard normal distribution table to calculate probabilities of normally distributed random variables.

Motivation: Congratulations, you are our 100,000th customer!

Last time, we discussed about the probability of the next customer arriving in the next hour, next day, next year. What about the probability of the 10th customer arriving at a certain time? Or, consider a printer that starts to fail and needs maintenance after the 1000th job: what is the probability these failures start happening a month from now?

Motivation: Food poisoning and how to avoid it

A chef is using a new thermometer to tell whether certain foods have been adequately cooked. For example, chicken has to be at an internal temperature of 165 Fahrenheit or more to be adequately cooked; otherwise, we run the risk of salmonella. The restaurant wants to take *no chances*! The chef, then, takes a look at the temperature reading at the thermometer and sees 166. What is the probability that the chicken is adequately cooked, if we assume that the thermometer is right within a margin of error?

The Gamma and the Erlang distribution

Assume again that you are given a rate λ with which some events happen. So far, we have addressed two related questions:

1. What is the probability that the next event happens during some time interval? This is addressed through defining and using an exponentially distributed random variable (continuous).
2. What is the probability that we see a number of events during some time interval? This is addressed through defining and using a Poisson distributed random variable (discrete).

It is time to address a third question: what is the probability that the k-th event happens during some time interval? Like the first question, this is addressed using a continuous random variable; like the second question, we need a number of events to happen first.

Definition 1 (The Gamma distribution) *A continuous random variable X defined over the interval of $[0, \infty)$ is Erlang distributed if it can be written it has probability density function given by*

$$f(x) = \begin{cases} \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{(k-1)!}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0, \end{cases}$$

where $\lambda > 0$ and $k > 0$ are given parameters and $\Gamma(k)$ is the Gamma function. [1] *We sometimes write that $X \sim Gamma(k, \lambda)$ if it follows the Gamma distribution with rate λ and shape parameter k.*

[1] In this class, we will only deal with integer values of k, and hence $\Gamma(k) = (k-1)!$.

When k is a positive integer number, the Gamma distribution is referred to as the **Erlang distribution**. In essence, the definition follows.

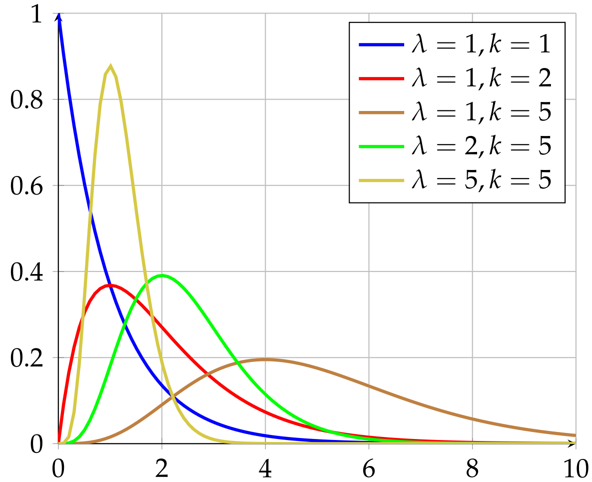
Definition 2 (The Erlang distribution) *A continuous random variable X defined over the interval of $[0, \infty)$ is Gamma distributed if it can be written as **the summation of exponentially distributed random variables** $X = \sum_{i=1}^{k} X_i$, where X_i is an exponentially distributed random variable. When X is Erlang distributed it has probability density function given by*

$$f(x) = \begin{cases} \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{\Gamma(k)}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0, \end{cases}$$

where real $\lambda > 0$ and integer $k > 0$ are given parameters.

Due to the nature of the Γ function, it is typically easier to integrate the pdf after we plug in the value for k that we are interested in. For example, consider the problem from our motivation.

Figure 1: The Erlang distribution probability density function visualized for different values of λ and k.



Congratulations, you are our 10th customer

A store, which is open for 8 hours every day, gives a gift card to the (exactly) 10th customer of every day. The store has observed that customers show up at a rate of 1 every 20 minutes (exponentially distributed). What is the probability the 10th customer of the day shows up in the second half of the day?

Let T be the time the $k = 10$-th customer arrives. T can then be written as the summation of the arrival times of the first plus the second plus the third, all the way to the 10-th customer: since it is a summation of exponentially distributed random variables, T is Erlang distributed with parameters λ (the rate) and $k = 10$.
We are interested in $P(T > 4 \text{ hours})$. For convenience, we translate the rate to hours, getting that $\lambda = 3$ per hour:

$$P(T > 4 \text{ hours}) = \int_4^8 f(x)dx = \int_4^8 \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{\Gamma(k)} dx =$$
$$= \int_4^8 \frac{3^{10} \cdot x^9 \cdot e^{-3 \cdot x}}{9!} dx = \frac{59049}{362880} \int_4^8 x^9 \cdot e^{-3 \cdot x} =$$
$$= \frac{59049}{362880} \cdot 1.487 = 0.242.$$

Replacing parts

A machine requires a component to work. The component is replaced every two times the machine is doing a job. The machine works at a rate of 3 jobs per 8 hours. What is the probability the component is not replaced in the first 8 hours?

Once again, we use $\lambda = 3/8$ per hour, and then define T as the Erlang distributed random variable of the time the 2nd job appears ($k = 2$, integer and hence Erlang). We then have:

$$P(T > 8 \text{ hours}) = \int_8^\infty f(x)dx = \int_8^\infty \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{\Gamma(k)} dx = \\ = \int_8^\infty \frac{\left(\frac{3}{8}\right)^2 \cdot x \cdot e^{-\frac{3}{8}x}}{\Gamma(2)} dx = \frac{9}{128} \int_8^\infty x \cdot e^{-\frac{3}{8}x}. \quad (1)$$

Recall that we can integrate by parts to get:

$$\int_8^\infty x \cdot e^{-\frac{3}{8}x} = -\frac{8}{3} \int_8^\infty x \cdot \left(e^{-\frac{3}{8}x}\right)' = -\frac{8}{3} \cdot \left(x \cdot e^{-\frac{3}{8}x}\Big|_8^\infty - \int_8^\infty e^{-\frac{3}{8}x} \right) = \\ = -\frac{8}{3} \cdot \left(8e^{-3} + \frac{8}{3} \cdot e^{-3}\right) = \frac{256}{9} \cdot e^{-3} = 1.4162. \quad (2)$$

Plugging the result from (2) into (1), we get 0.0996.

The lifetime of a printer toner is exponentially distributed: it needs to be replaced once every 9 months. What probability distribution would you use for each of the following cases?

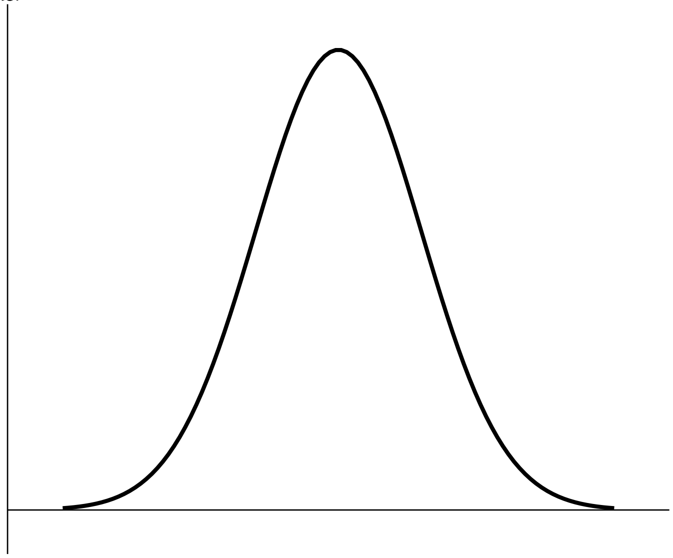
a) The number of toners you need to buy in the next 3 years.

b) The time until the toner is replaced.

c) The time until you run out of toners, if you have bought a package with 3 toners.

d) The time until the toner is replaced, given that the toner currently in use has not been replaced for 6 months already.

Figure 2: An example of how the normal distribution probability density function looks like.



The normal distribution

We have come to a big one. This is arguably the most well-studied, used, and applied distribution among the ones we have studied so far. It is defined through two parameters referred to as the mean (μ) and the variance (σ). We then say that a normally distributed random variable X is $\mathcal{N}(\mu, \sigma^2)$. [2]

[2] Note that we replace the standard deviation σ, with its square σ^2.

Definition 3 (Normal distribution) *A random variable X is said to follow a normal distribution with mean μ and standard deviation σ, if it has a probability density function of:*

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}.$$

We then commonly write that $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$.

We present an example for how the normal distribution pdf looks like in Figure 2. We observe that it is **symmetric** and **bell-shaped**. From the definition of the normal distribution, we may also get the cumulative distribution function as:

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

This is clearly an integral that we would rather not have to deal with!

Parameters

The two parameters that describe a normal distribution affect its *location* (μ) and its *spread* (σ). Visually, we show this relationship in

Figure 3: Some examples of how the normal distribution is affected by its mean and standard deviation.

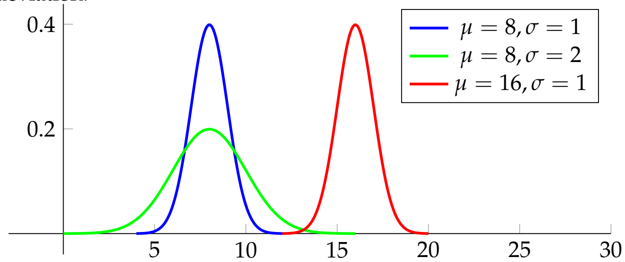


Figure 3. Note how the mean affects the location of the normal distribution, whereas the standard deviation affects how far it spreads.

The standard normal distribution

When $\mu = 0$ and $\sigma = 1$, we call the resulting normal distribution, the **standard normal distribution** and denote it as $\mathcal{N}(0,1)$. Due to the applicability of the normal distribution in many real-life instances, the standard normal distribution has been extensively studied and we have in our possession tables containing the values of the cumulative density function. An example of such a table is provided to you in the next page.

For convenience, we refer to the pdf and the cdf of the standard normal distribution $\mathcal{N}(0,1)$ as $\phi(z)$ and $\Phi(z)$, respectively. Note the use of z rather than the typically used x!

NORMAL CUMULATIVE DISTRIBUTION FUNCTION ($\Phi(z)$)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

A normally distributed random variable $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$ can be converted to the standard normal distribution $\mathcal{N}\left(0, 1\right)$ through one small, simple transformation, called the z-transform:

$$\text{If } X \text{ is } \mathcal{N}(\mu, \sigma^2), \text{ then } Z = \frac{X-\mu}{\sigma} \text{ is } \mathcal{N}(0, 1).$$

This implies that for any normally distributed random variable X, $P(X = x)$ can be written as $P(Z = \frac{x-\mu}{\sigma})$, where Z is distributed following the standard normal distribution!

Doing transformations

Let X be $\mathcal{N}(400, 400)$ (i.e., $\sigma^2 = 400 \implies \sigma = 20$). What is:

a) $P(X \leq 400)$? b) $P(X \leq 451)$? c) $P(X \leq 375)$?

a) $x = 400 \implies z = \frac{400-\mu}{\sigma} = 0.$

b) $x = 451 \implies z = \frac{51}{20} = 2.55.$

c) $x = 375 \implies z = \frac{-25}{20} = -1.25.$

With z at hand, calculating a probability becomes merely a look-up operation! Indeed, all you need to do is find the z value in the cdf table. The rows reveal the two most important digits and the columns the third most important digit. For example, finding $z = 1.37$, we'd go to the 1.3 row and the 0.07 column.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	**0.9147**	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319

Using the z-table

Let X be $\mathcal{N}(400, 400)$ (i.e., $\sigma^2 = 400 \implies \sigma = 20$). What is:

a) $P(X \leq 400)$? b) $P(X \leq 451)$? c) $P(X \leq 375)$?

We already have found that:

a) $z = 0$. b) $z = 2.55$. c) $z = -1.25$.

Now is the time to find these values in the table. For the first one:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	**0.5000**	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359

For the second one:

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	**0.9946**	0.9948	0.9949	0.9951	0.9952

For the third one, we run into a problem. The table provided does not give any negative values for z!

Recall of two facts:

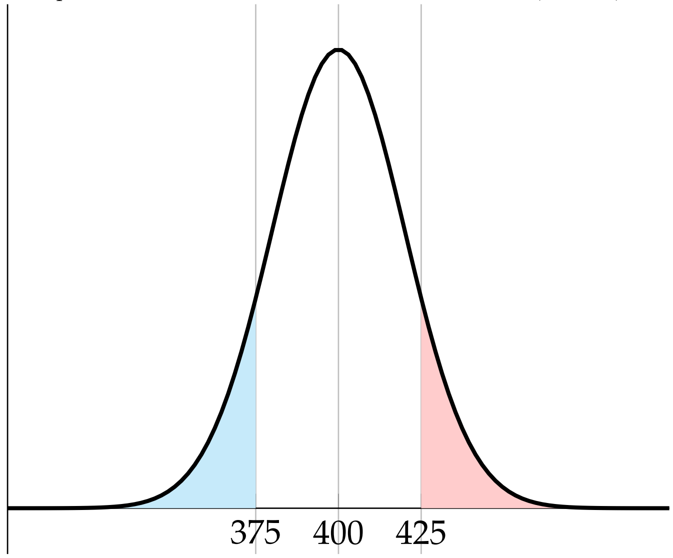
1. The normal distribution is symmetric.
2. For any continuous random variable, the probability can be found by looking at the area under the curve of the pdf.

Let us combine these two facts in an image. Consider the random variable $X \sim \mathcal{N}(400, 400)$ in Figure 4. As a reminder, we are interested in $P(X \leq 375)$.

Due to symmetry, the two shaded areas (in blue and red) have to be equal, as they are symmetric from the mean (400). Hence, we have that $P(X \leq 375) = P(X \geq 425)$. That said, we do know that $P(X \geq 425) = 1 - P(X \leq 425)$. Finally, recall that $P(X \leq 425)$ can be found in the z-table, as it corresponds to a positive value! Hence, to recap, when dealing with negative values of z, we can follow the next steps:

1. Instead of $z < 0$, search for $-z$.
2. Find the value in the z-table, $\Phi(-z)$.
3. Then, $\Phi(z) = 1 - \Phi(-z)$.

Figure 4: The pdf of the distribution of random variable $X \sim \mathcal{N}(400, 400)$.



Negative values of z

Let X be $\mathcal{N}(400, 400)$ (i.e., $\sigma^2 = 400 \implies \sigma = 20$). What is:

a) $P(X \leq 400)$? b) $P(X \leq 451)$? c) $P(X \leq 375)$?

We have solved the first two:

a) $P(X \leq 400) = 0.$ b) $P(X \leq 451) = 0.9946.$ c) $P(X \leq 375)$?

Finally, for $P(X \leq 375)$, with corresponding $z = -1.25$, we find $-z$ on the table.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	**0.8944**	0.8962	0.8980	0.8997	0.9015

Then, we report that

$$P(X \leq 375) = \Phi(-1.25) = 1 - \Phi(1.25) = 0.1056.$$

Finally, like in all distributions, if we are interested in the probability of a quantity being within a range of values (say, $P(a \leq X \leq b)$), then we may calculate $F(b) - F(a)$, or using the corresponding z-values (z_a, z_b), we may calculate that probability as $P(a \leq X \leq b) = F(b) - F(a) = \Phi(z_b) - \Phi(z_a)$.

A newsvendor is deciding how many newspapers to order for the following day. The demand for newspapers follows a normal distribution with a mean of 100 and a standard deviation of 10.

- What is the probability of selling all the newspapers they order if they place an order for:
  a) 120 newspapers?
  b) 80 newspapers?
- How many newspapers should the newsvendor order if:
  a) the newsvendor is risk-averse and would like at least a 90% chance of selling all of them?
  b) the newsvendor is risk-seeking and would like at least a 90% chance of satisfying all demand?

Summary

In Table 1, we provide all of the results from Lectures 7 and 8. We are bundling them together (even though we already provided a summary of results for Lecture 7 alone) so that we have everything for continuous distributions in one place.

Table 1: A summary of all results from Lectures 7 and 8.

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	$\lambda, \gamma > 0$	$[\gamma, +\infty)$	$f(x) = \begin{cases} \lambda \cdot e^{-\lambda(x-\gamma)}, & \text{if } x \geq \gamma \\ 0, & \text{if } x < \gamma. \end{cases}$
Gamma	$\lambda > 0, k > 0$	$[0, +\infty)$	$f(x) = \begin{cases} \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{\Gamma(k)}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0, \end{cases}$
Erlang	$\lambda > 0$, integer $k > 0$	$[0, +\infty)$	$f(x) = \begin{cases} \frac{\lambda^k \cdot x^{k-1} \cdot e^{-\lambda x}}{(k-1)!}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0, \end{cases}$
Normal	μ, σ^2	$(-\infty, +\infty)$	$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$

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".

Erlang: "time to k-th event"; "rate of events".

Normal: "normally distributed"; "average/summation of multiple identical random variables"; "central limit theorem". [3]

[3] These keywords are provided here, but are explained in Lecture 10.