Andrew Gard - equitable.equations@gmail.com

The Gamma Distribution in *R*

The **gamma distribution** models the elapsed time before a fixed number of occurrences of a randomly-occurring event, like calls to a pizza place or defective products in a factory.

The **gamma distribution** models the elapsed time before a fixed number of occurrences of a randomly-occurring event, like calls to a pizza place or defective products in a factory.

The gamma distribution is completely specified by a two parameters, θ, representing the mean waiting time between events, and α, the total number of occurrences needed.

The **gamma distribution** models the elapsed time before a fixed number of occurrences of a randomly-occurring event, like calls to a pizza place or defective products in a factory.

The gamma distribution is completely specified by a two parameters, θ, representing the mean waiting time between events, and α, the total number of occurrences needed. These are known as the **scale** and **shape** parameters, respectively.

The **gamma distribution** models the elapsed time before a fixed number of occurrences of a randomly-occurring event, like calls to a pizza place or defective products in a factory.

The gamma distribution is completely specified by a two parameters, θ, representing the mean waiting time between events, and α, the total number of occurrences needed. These are known as the **scale** and **shape** parameters, respectively. Instead of θ, the mean number of occurrences per interval is sometimes specified instead. This is known as the **rate** parameter, $\lambda = 1/\theta$.

The **gamma distribution** models the elapsed time before a fixed number of occurrences of a randomly-occurring event, like calls to a pizza place or defective products in a factory.

The gamma distribution is completely specified by a two parameters, θ, representing the mean waiting time between events, and α, the total number of occurrences needed. These are known as the **scale** and **shape** parameters, respectively. Instead of θ, the mean number of occurrences per interval is sometimes specified instead. This is known as the **rate** parameter, $\lambda = 1/\theta$.

Base *R* functions use the rate parameter λ by default.

There are 4 basic functions in R for calculating in the gamma distribution.

There are 4 basic functions in *R* for calculating in the gamma distribution.

- `rgamma(n, alpha, lambda)` generates `n` random values from the gamma distribution with shape parameter `alpha` and rate parameter `lambda`.

There are 4 basic functions in *R* for calculating in the gamma distribution.

- `rgamma(n, alpha, lambda)` generates `n` random values from the gamma distribution with shape parameter `alpha` and rate parameter `lambda`. For instance,

```
rgamma(6, 2, .5)
```

```
## [1] 6.891120 2.254109 1.302274 4.101624 7.770818 5.081080
```

There are 4 basic functions in *R* for calculating in the gamma distribution.

- `rgamma(n, alpha, lambda)` generates `n` random values from the gamma distribution with shape parameter `alpha` and rate parameter `lambda`. For instance,

```
rgamma(6, 2, .5)
```

```
## [1] 6.891120 2.254109 1.302274 4.101624 7.770818 5.081080
```

Here and elsewhere, the scale parameter θ may be used instead, but it must be named explicitly.

```
rgamma(6, 2, scale = 2)
```

```
## [1] 6.891120 2.254109 1.302274 4.101624 7.770818 5.081080
```

- `pgamma(x, alpha, lambda)` is the cumulative distribution function. It returns the probability that the waiting time for alpha occurrences is no more than `x`. As usual in *R*, `x` can be a vector.

- `pgamma(x, alpha, lambda)` is the cumulative distribution function. It returns the probability that the waiting time for alpha occurrences is no more than `x`. As usual in *R*, `x` can be a vector. For instance,

```
pgamma(1:6, 2, .5)
```

```
## [1] 0.09020401 0.26424112 0.44217460 0.59399415 0.71270250 0.80085173
```

- `pgamma(x, alpha, lambda)` is the cumulative distribution function. It returns the probability that the waiting time for alpha occurrences is no more than `x`. As usual in *R*, `x` can be a vector. For instance,

```
pgamma(1:6, 2, .5)
```

```
## [1] 0.09020401 0.26424112 0.44217460 0.59399415 0.71270250 0.80085173
```

- `qgamma(p, alpha, lambda)` is the inverse distribution function. It returns the value `x` such that `pgamma(x, alpha, lambda) = p`. In other words, it computes quantiles in the specified gamma distribution. Again, `p` can be a vector.

- `pgamma(x, alpha, lambda)` is the cumulative distribution function. It returns the probability that the waiting time for alpha occurrences is no more than `x`. As usual in *R*, `x` can be a vector. For instance,

```
pgamma(1:6, 2, .5)
```

```
## [1] 0.09020401 0.26424112 0.44217460 0.59399415 0.71270250 0.80085173
```

- `qgamma(p, alpha, lambda)` is the inverse distribution function. It returns the value `x` such that `pgamma(x, alpha, lambda) = p`. In other words, it computes quantiles in the specified gamma distribution. Again, `p` can be a vector. For instance,

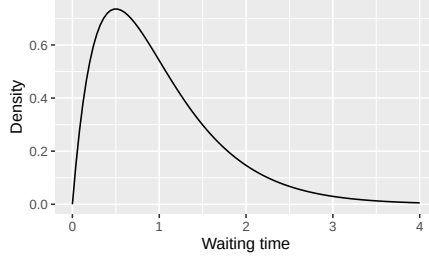
```
qgamma(c(.2, .4, .6, .8), 2, .5)
```

```
## [1] 1.648777 2.752843 4.044626 5.988617
```

- `dgamma(x, alpha, lambda)` is the probability density function (pdf) of the gamma distribution with shape parameter `alpha` and rate parameter `lambda`.

- `dgamma(x, alpha, lambda)` is the probability density function (pdf) of the gamma distribution with shape parameter `alpha` and rate parameter `lambda`. This is useful in creating density plots, for instance.

```
library(tidyverse)
ggplot() +
  xlim(c(0, 4)) +
  geom_function(fun = dgamma,
                args = list(shape = 2, scale = .5)) +
  labs(x = "Waiting time", y = "Density")
```



Example. Calls to a customer-service line come at an average rate of 1 every 3 minutes.

1. What is the probability that more than an hour elapses before 25 calls come in?

2. What is the 95^{th} percentile for time needed for 5 calls to come in?

3. Simulate waiting times for 5 calls 1000 times. Plot the results.