Definition of quantiles

Definition of quantiles

Quantiles are cutoff points that divide a dataset into intervals with set probabilities. The qth quantile is the value at which q% of the observations are equal to or less than that value.

Using the `quantile` function

Given a dataset `data` and desired quantile `q`, you can find the `q`th quantile of `data` with:

```
quantile(data,q)
```

Percentiles

Percentiles are the quantiles that divide a dataset into 100 intervals each with 1% probability. You can determine all percentiles of a dataset `data` like this:

```
p <- seq(0.01, 0.99, 0.01)
quantile(data, p)
```

Quartiles

Quartiles divide a dataset into 4 parts each with 25% probability. They are equal to the 25th, 50th and 75th percentiles. The 25th percentile is also known as the *1st quartile*, the 50th percentile is also known as the *median*, and the 75th percentile is also known as the *3rd quartile*.

The `summary()` function returns the minimum, quartiles and maximum of a vector.

Examples

Load the heights dataset from the **dslabs** package:

```
library(dslabs)
data(heights)
```

Use `summary()` on the `heights$height` variable to find the quartiles:

```
summary(heights$height)
```

Find the percentiles of heights$height:

```
p <- seq(0.01, 0.99, 0.01)
percentiles <- quantile(heights$height, p)
```

Confirm that the 25th and 75th percentiles match the 1st and 3rd quartiles. Note that `quantile()` returns a named vector. You can access the 25th and 75th percentiles like this (adapt the code for other percentile values):

```
percentiles[names(percentiles) == "25%"]
percentiles[names(percentiles) == "75%"]
```

Learn About Verified Certificates