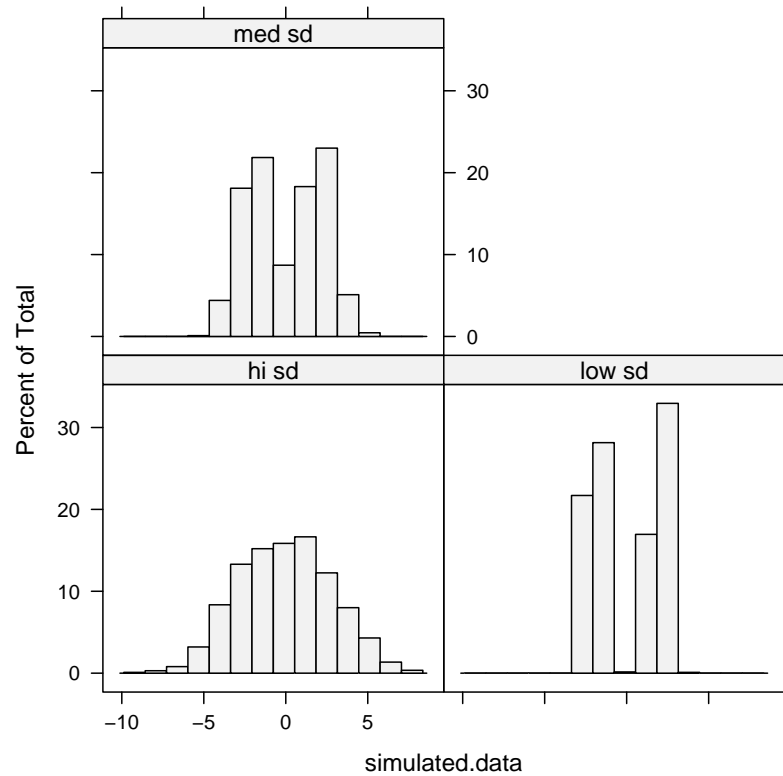A common scientific experiment involves looking at the effect of a factor with a single factor level and a control. In this case, H_0 is $\mu_{\text{control}} = \mu_{\text{treat}}$.

Signal versus noise

Quantifying signal and noise

Signal Differences in means in two mean tests

Noise Standard error of the mean (or related to SEM)

The ratio of these two is called a **test-statistic**

Students t-test

Student's t test is designed to scale the signal by the combined noise. It is the ratio of the difference in means and the standard error of the difference in those means.
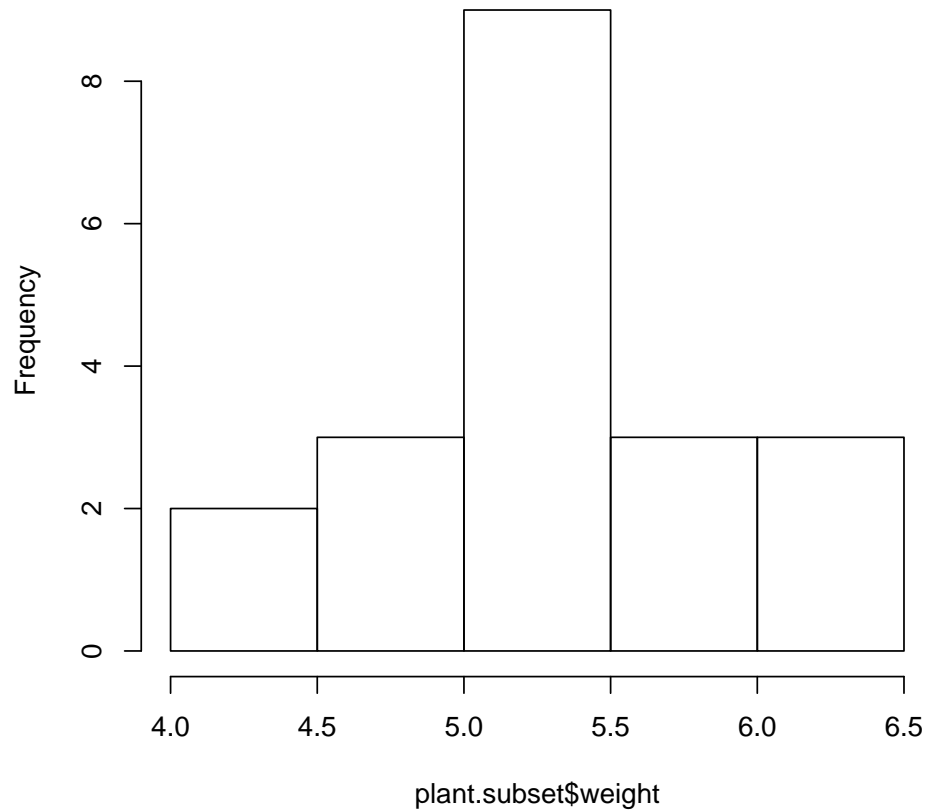
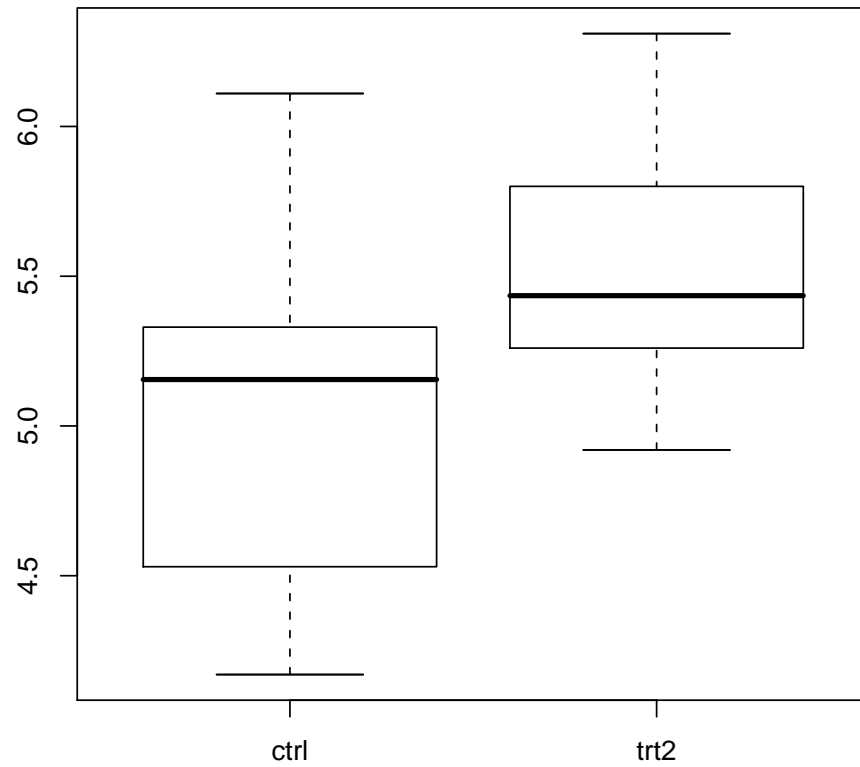$$t = \frac{\bar{Y} - \bar{X}}{\sqrt{\frac{s_Y^2}{n_Y} + \frac{s_X^2}{n_X}}}$$

t-**test assumptions**

- observations chosen at random from the two treatment levels for each of the two samples.

- sampled populations are distributed normally

- data are continuous

- the variances in the two treatment levels are equal

t-**test implemented**

Histogram of plant.subset$weight

plant.subset$weight

data

```
> X <- plant.subset$weight[plant.subset$group == "ctrl"]
> Y <- plant.subset$weight[plant.subset$group == "trt2"]
> cbind(X, Y)

        X    Y
[1,] 4.17 6.31
[2,] 5.58 5.12
[3,] 5.18 5.54
[4,] 6.11 5.50
[5,] 4.50 5.37
```

```
 [6,] 4.61 5.29
 [7,] 5.17 4.92
 [8,] 4.53 6.15
 [9,] 5.33 5.80
[10,] 5.14 5.26
```

Test statistic and α

```
> that <- (mean(X) - mean(Y))/sqrt(var(X)/length(X) + var(Y)/length(Y))
> df <- (length(Y) + length(X)) - 2
> that

[1] -2.134020

> df

[1] 18

> 2 * (1 - pt(abs(that), df))

[1] 0.04685138
```

Builtin R function

```
> t.test(X, Y, var.equal = T)

        Two Sample t-test

data:  X and Y
t = -2.134, df = 18, p-value = 0.04685
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.980338117 -0.007661883
sample estimates:
mean of x mean of y
    5.032     5.526
```

Builtin R function on dataframe

```
> names(plant.subset)

[1] "weight" "group"
```

```
> t.test(plant.subset$weight ~ plant.subset$group, var.equal = T)

        Two Sample t-test

data:  plant.subset$weight by plant.subset$group
t = -2.134, df = 18, p-value = 0.04685
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.980338117 -0.007661883
sample estimates:
mean in group ctrl mean in group trt2
           5.032              5.526
```