Design of Experiments
Part 2 – Basic Statistical Concepts (1)

- Simple **comparative** experiments
  - The hypothesis testing framework
  - The two-sample *t*-test
  - Checking assumptions, validity

Portland Cement Formulation (page 24)

Table 2-1 Tension Bond Strength Data for the Portland Cement Formulation Experiment

j	Modified Mortar y_{1j}	Unmodified Mortar y_{2j}
1	16.85	17.50
2	16.40	17.63
3	17.21	18.25
4	16.35	18.00
5	16.52	17.86
6	17.04	17.75
7	16.96	18.22
8	17.15	17.90
9	16.59	17.96
10	16.57	18.15

Graphical View of the Data

Dot Diagram, Fig. 2-1, pp. 24

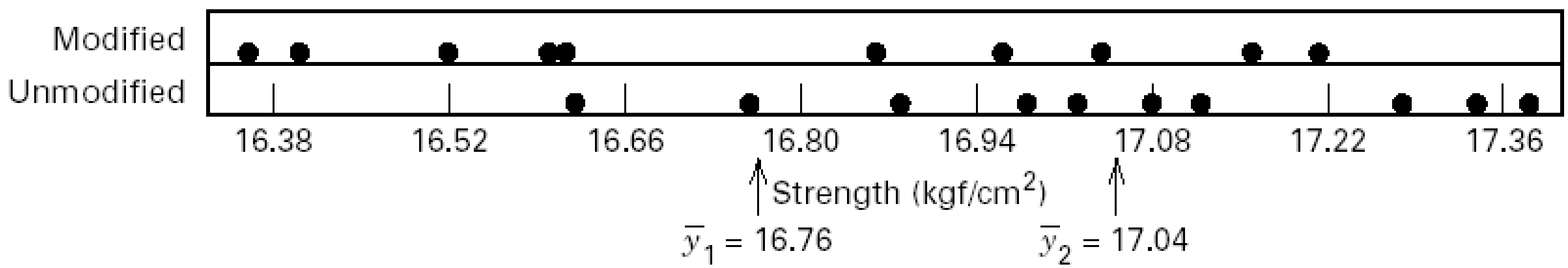

Figure 2-1 Dot diagram for the tension bond strength data in Table 2-1.

Box Plots, Fig. 2-3, pp. 26

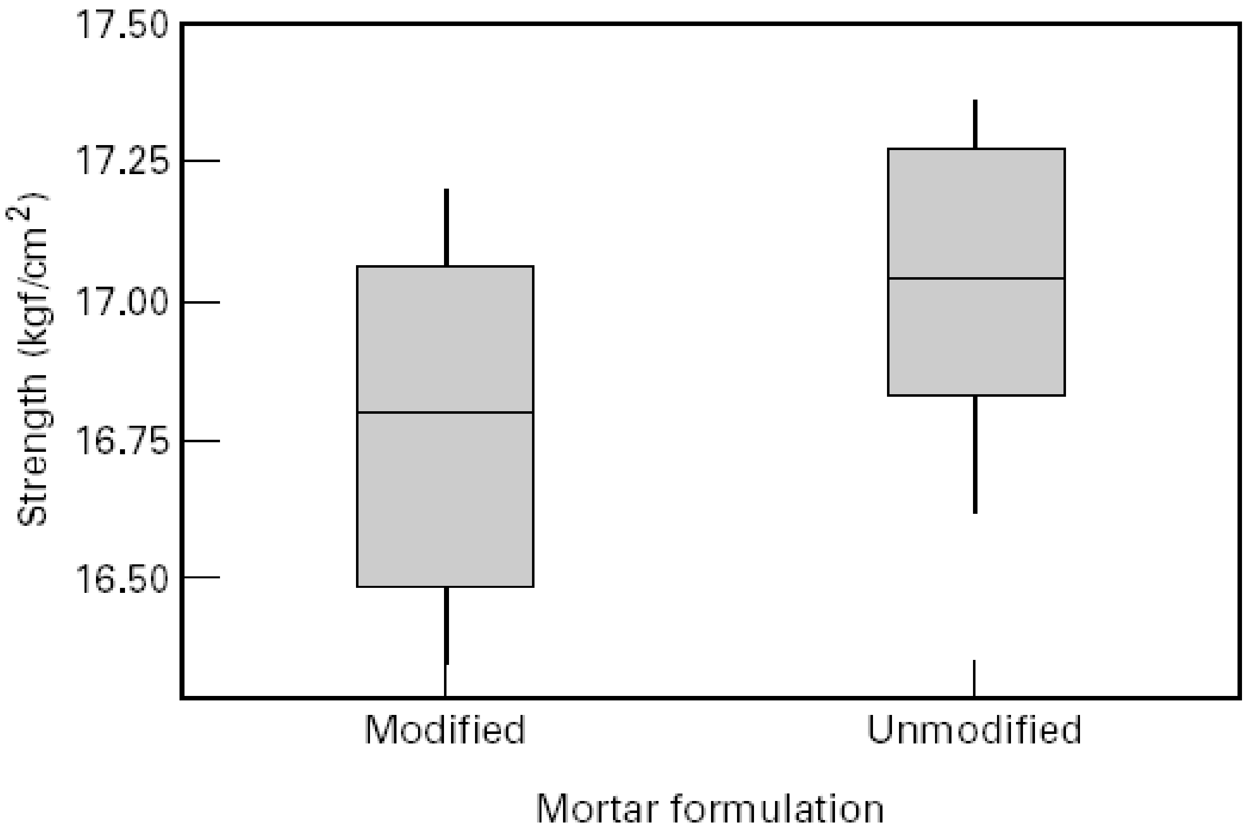

Figure 2-3 Box plots for the portland cement tension bond strength experiment.

The Hypothesis Testing Framework

- **Statistical hypothesis testing** is a useful framework for many experimental situations
- Origins of the methodology date from the early 1900s
- We will use a procedure known as the **two-sample *t*-test**

The Hypothesis Testing Framework

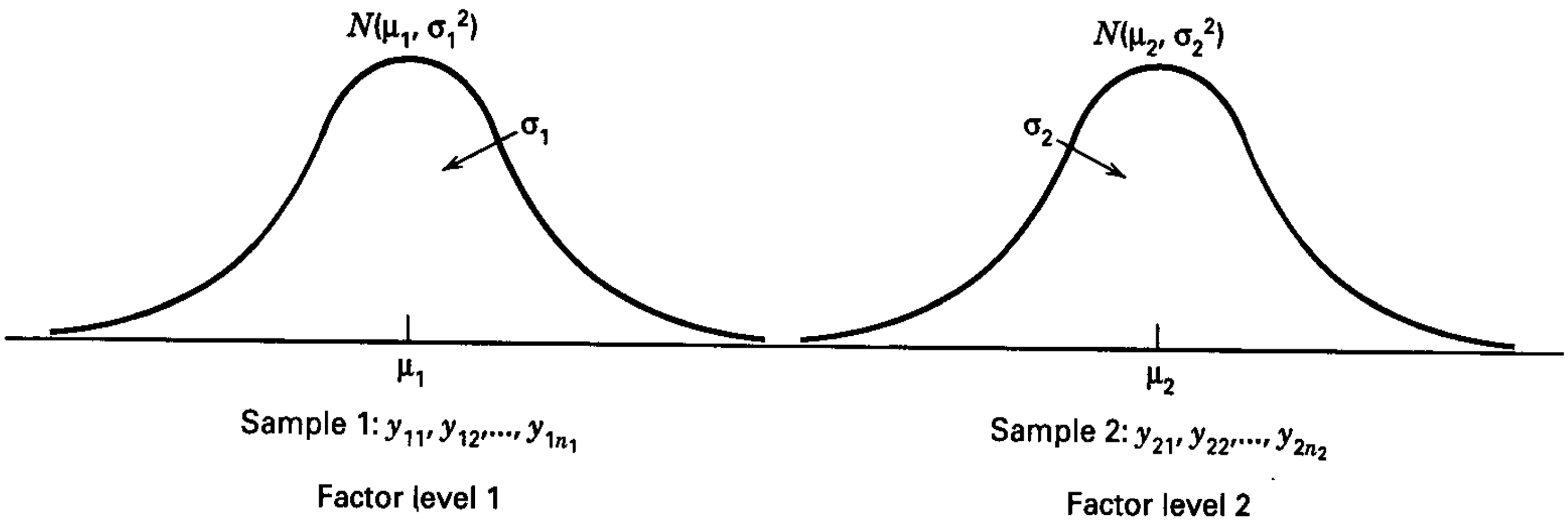

Figure 2-9 The sampling situation for the two-sample *t*-test.

- Sampling from a **normal** distribution
- Statistical hypotheses:

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

Estimation of Parameters

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ estimates the population mean } \mu$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (y_i - \bar{y})^2 \text{ estimates the variance } \sigma^2$$

Summary Statistics (pg. 36)

Formulation 1

"New recipe"

$$\bar{y}_1 = 16.76$$
$$S_1^2 = 0.100$$
$$S_1 = 0.316$$
$$n_1 = 10$$

Formulation 2

"Original recipe"

$$\bar{y}_1 = 17.04$$
$$S_1^2 = 0.061$$
$$S_1 = 0.248$$
$$n_1 = 10$$

How the Two-Sample *t*-Test Works:

Use the sample means to draw inferences about the population means

$$\bar{y}_1 - \bar{y}_2 = 16.76 - 17.04 = -0.28$$

$$\frac{\text{Difference in sample means}}{\text{Standard deviation of the difference in sample means}}$$

$$\sigma_{\bar{y}}^2 = \frac{\sigma^2}{n}$$

This suggests a statistic:

$$Z_0 = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

How the Two-Sample *t*-Test Works:

Use S_1^2 and S_2^2 to estimate σ_1^2 and σ_2^2

The previous ratio becomes $\dfrac{\bar{y}_1 - \bar{y}_2}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}}$

However, we have the case where $\sigma_1^2 = \sigma_2^2 = \sigma^2$

Pool the individual sample variances:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

How the Two-Sample *t*-Test Works:

The test statistic is

$$t_0 = \frac{\bar{y}_1 - \bar{y}_2}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

- Values of t_0 that are near zero are consistent with the null hypothesis
- Values of t_0 that are very different from zero are consistent with the alternative hypothesis
- t_0 is a "distance" measure-how far apart the averages are expressed in standard deviation units
- Notice the interpretation of t_0 as a **signal-to-noise** ratio

The Two-Sample (Pooled) *t*-Test

$$S_p^2 = \frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2} = \frac{9(0.100)+9(0.061)}{10+10-2} = 0.081$$

$$S_p = 0.284$$

$$t_0 = \frac{\bar{y}_1-\bar{y}_2}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{16.76-17.04}{0.284\sqrt{\frac{1}{10}+\frac{1}{10}}} = -2.20$$

The two sample means are a little over two standard deviations apart
Is this a "large" difference?

The Two-Sample (Pooled) *t*-Test

- So far, we haven't really done any "statistics"
- We need an **objective** basis for deciding how large the test statistic t_0 really is
- In 1908, W. S. Gosset derived the **reference distribution** for t_0 … called the *t* distribution
- Tables of the *t* distribution - text, page 606

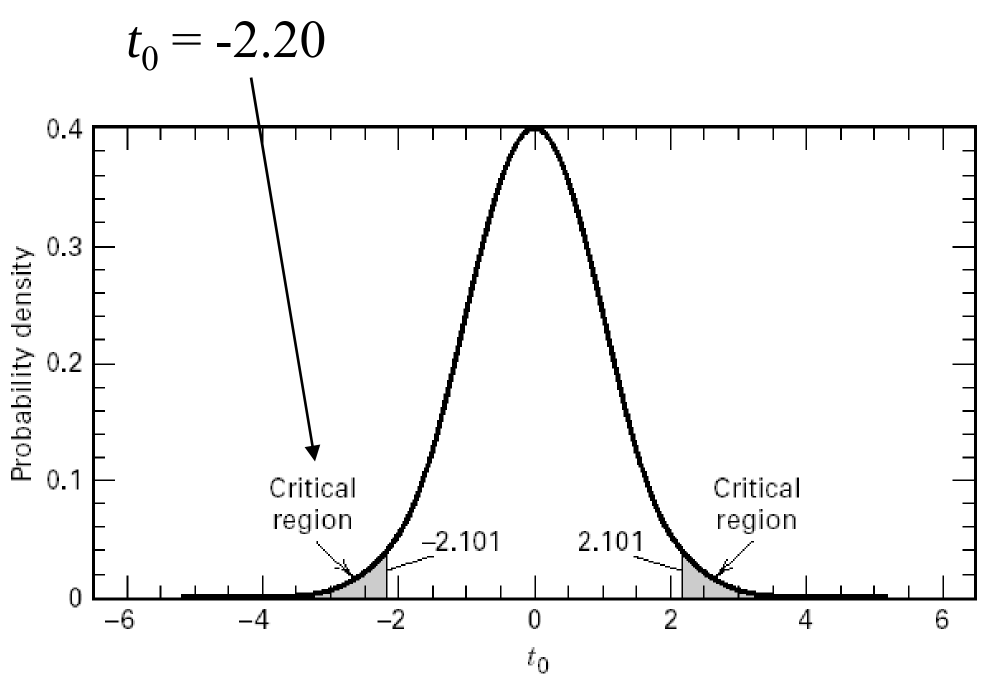

Figure 2-10 The t distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$.

The Two-Sample (Pooled) *t*-Test

- A value of t_0 between –2.101 and 2.101 is consistent with equality of means
- It is possible for the means to be equal and t_0 to exceed either 2.101 or –2.101, but it would be a "**rare event**" … leads to the conclusion that the means are different
- Could also use the ***P*-value** approach

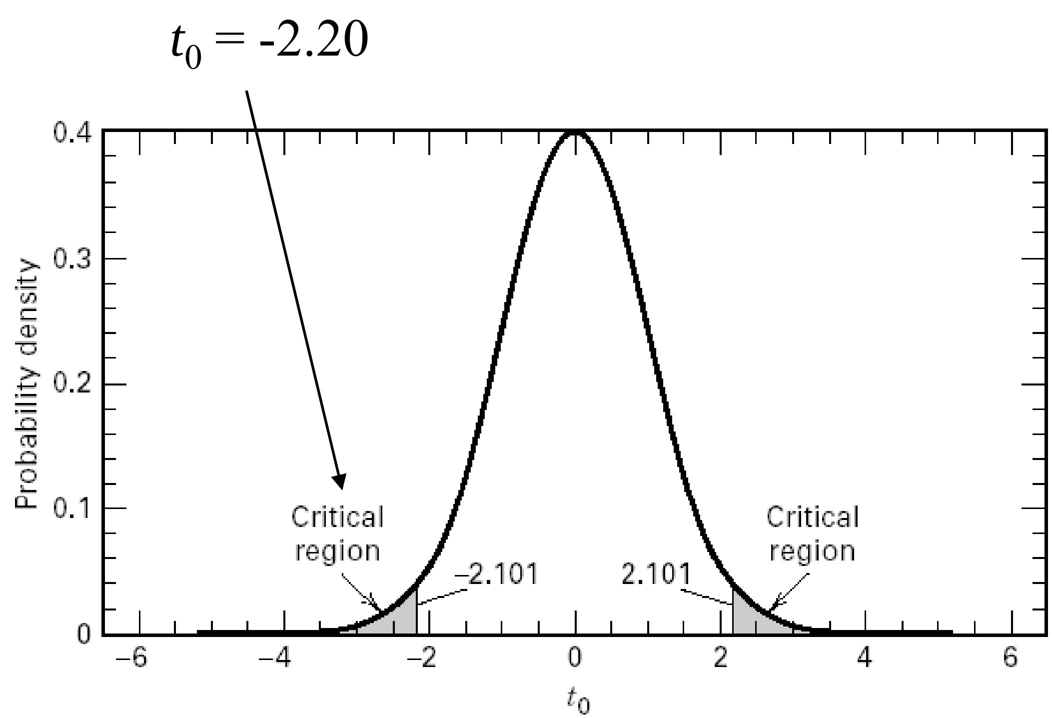

Figure 2-10 The t distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$.

The Two-Sample (Pooled) *t*-Test

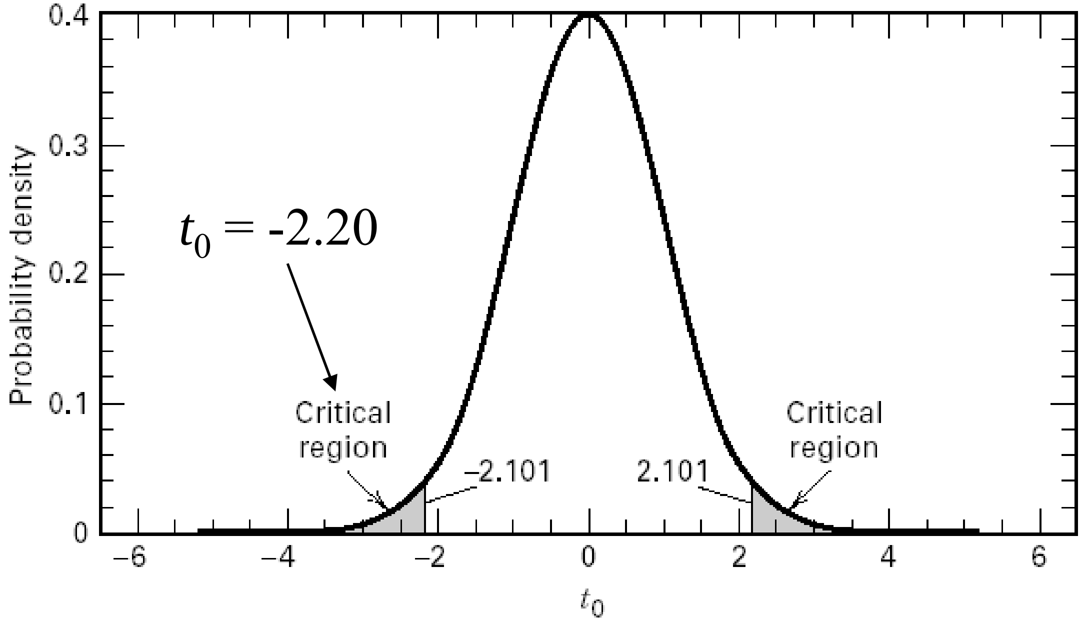

Figure 2-10 The t distribution with 18 degrees of freedom with the critical region $\pm t_{0.025,18} = \pm 2.101$.

- The ***P*-value** is the risk of **wrongly rejecting** the null hypothesis of equal means (it measures rareness of the event)
- The P-value in our problem is $P = 0.042$

Checking Assumptions – The Normal Probability Plot

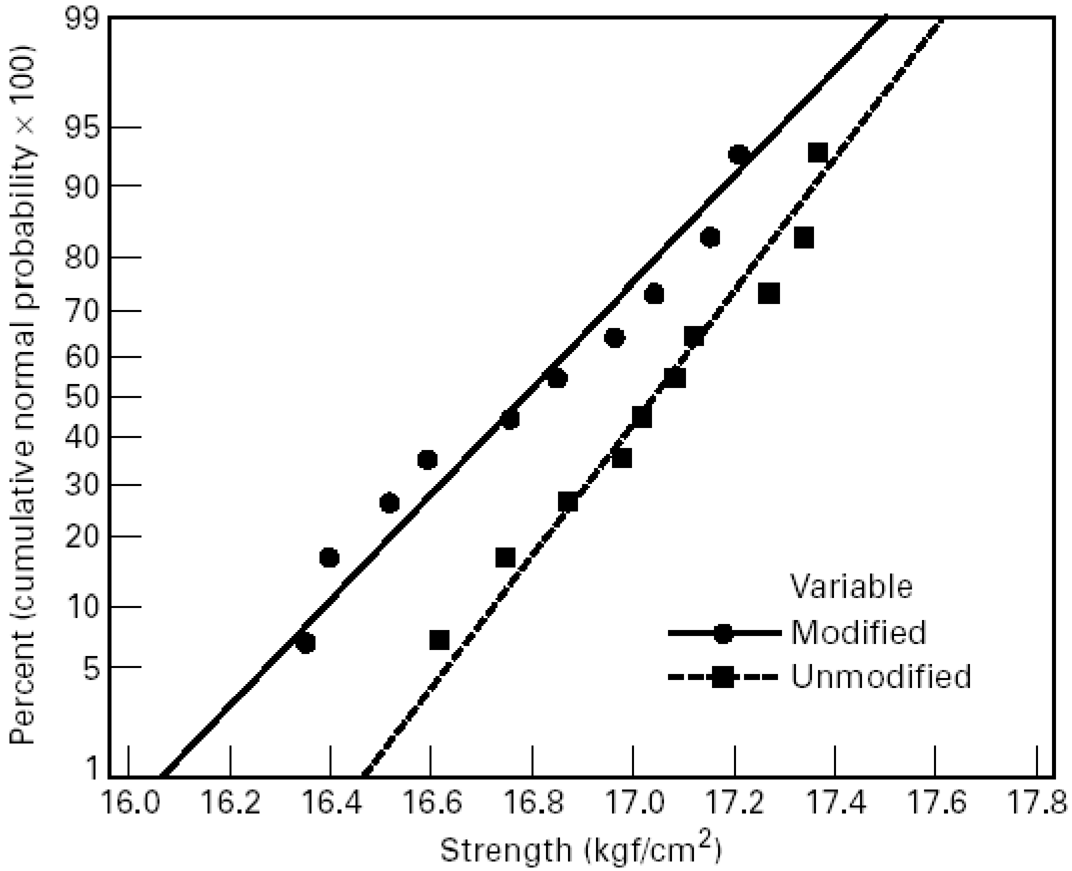

Figure 2-11 Normal probability plots of tension bond strength in the portland cement experiment.

Importance of the *t*-Test

- Provides an **objective** framework for simple comparative experiments
- Could be used to test all relevant hypotheses in a two-level factorial design, because all of these hypotheses involve the mean response at one "side" of the cube versus the mean response at the opposite "side" of the cube

Confidence Intervals (See pg. 43)

- Hypothesis testing gives an objective statement concerning the difference in means, but it doesn't specify "how different" they are
- **General form** of a confidence interval

$$L \leq \theta \leq U \text{ where } P(L \leq \theta \leq U) = 1 - \alpha$$

- The 100(1- α)% **confidence interval** on the difference in two means:

$$\bar{y}_1 - \bar{y}_2 - t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)} \leq \mu_1 - \mu_2 \leq$$
$$\bar{y}_1 - \bar{y}_2 + t_{\alpha/2, n_1+n_2-2} S_p \sqrt{(1/n_1) + (1/n_2)}$$

R Two-Sample *t*-Test Results

(assuming variances are equal)

```
> t.test(Modified, Unmodified, paired=F, var.equal=T)

        Two Sample t-test

data:  Modified and Unmodified
t = -2.1869, df = 18, p-value = 0.0422
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.54507339 -0.01092661
sample estimates:
mean of x mean of y
   16.764    17.042
```

R Two-Sample *t*-Test Results

(assuming variances are not equal)

```
> t.test(Modified, Unmodified, paired=F,
var.equal=F)

        Welch Two Sample t-test

data:  Modified and Unmodified
t = -2.1869, df = 17.025, p-value = 0.043
alternative hypothesis: true difference in
means is not equal to 0
95 percent confidence interval:
 -0.54617414 -0.00982586
sample estimates:
mean of x mean of y
   16.764    17.042
```