BIM 105
Probability and Statistics for Biomedical Engineers

David M. Rocke

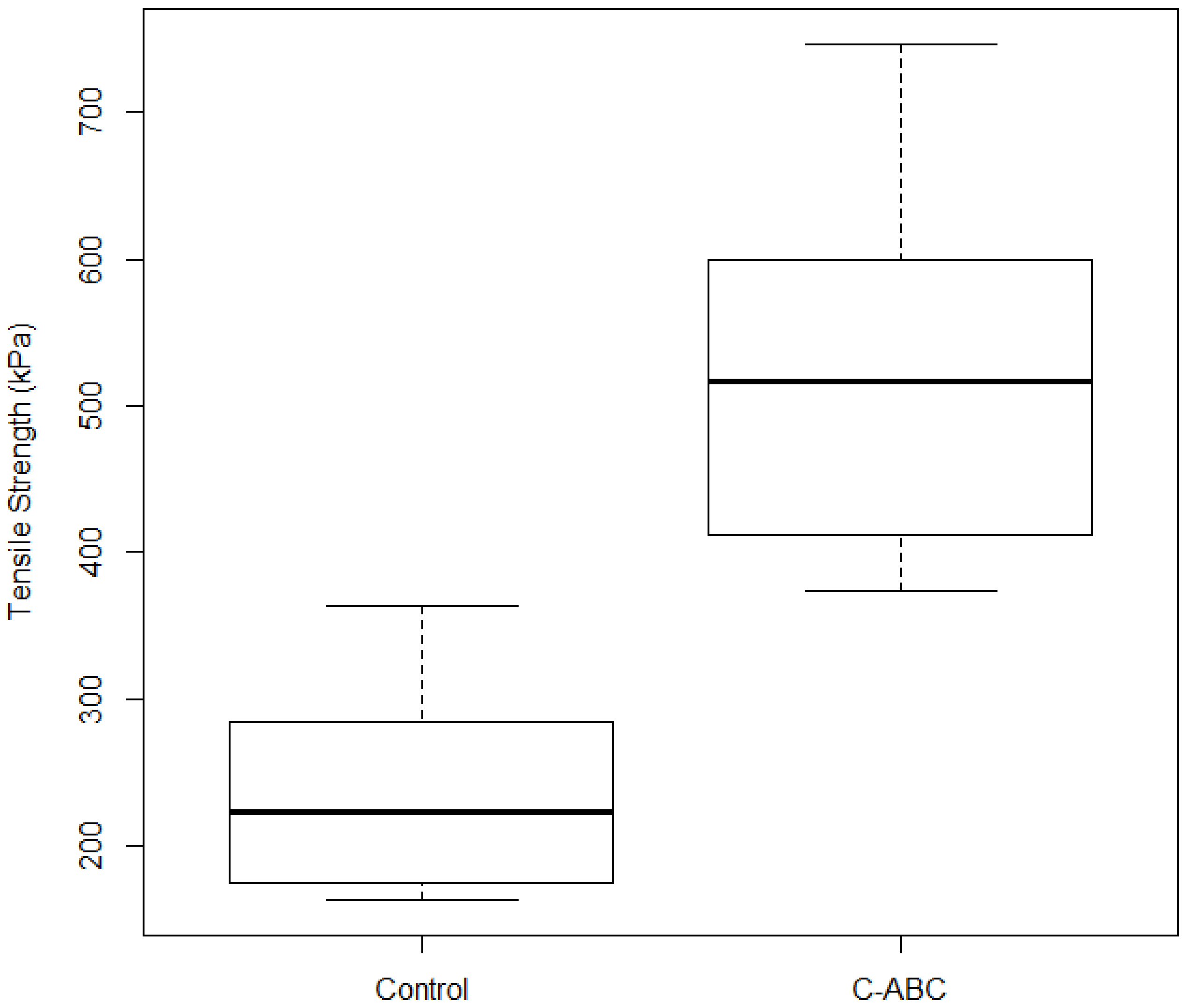
Department of Biomedical Engineering

Tests of Differences

- Most of the time, what we want to know is about differences.
  - Does a new treatment prolong the life of patients with a certain kind of cancer?
  - Does a growth medium with more calcium make stronger engineered cartilage?
  - Does a new manufacturing method for a medical device result in fewer defective devices?
- We want to test for these questions, expecting any deviation to be in the indicated direction, but knowing that it may work the other way.
- Thus, here too we always use two-sided tests and intervals.
- The null hypothesis almost always is the hypothesis that the difference is zero, aka null.

Effect of C-ABC on the Strength of Engineered Cartilage

- We have measurements in kPa of the tensile strength of pieces of engineered cartilage either with or without the addition of an enzyme called Chondroitinase ABC (C-ABC).
- The 7 control samples measured 364, 224, 183, 165, 163, 275, 293.
- The 6 treated samples measured 462, 747, 571, 599, 373, 413.
- How good is the evidence that the treated samples differ from the controls?


Tensile Strength (kPa)
700
600
500
400
300
200
Control
C-ABC

Tests for a Difference of Means

We have a series of measurements $X_1, X_2, \ldots, X_{n_X}$ and $Y_1, Y_2, \ldots, Y_{n_Y}$

$X_i \sim \mathrm{N}(\mu_X, \sigma_X^2)$

Each value of X has the same mean and the same variance (Identically distributed).

Different values of X are statistically independent. (Independence)

Each X is normally distributed. (Normality)

$Y_i \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$

Each value of Y has the same mean and the same variance (Identically distributed).

Different values of Y are statistically independent. (Independence)

Each Y is normally distributed. (Normality)

Values of X and Y are statistically independent. (Independence)

Tests for a Difference of Means

We have a series of measurements $X_1, X_2, \ldots, X_{n_X}$ and $Y_1, Y_2, \ldots, Y_{n_Y}$

$$X_i \sim \mathrm{N}(\mu_X, \sigma_X^2)$$

$$E(\bar{X}) = \mu_X$$

$$V(\bar{X}) = \sigma_X^2 / n_X$$

$$Y_i \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$$

$$E(\bar{Y}) = \mu_Y$$

$$V(\bar{Y}) = \sigma_Y^2 / n_Y$$

$$E(\bar{X} - \bar{Y}) = \mu_X - \mu_Y$$

$$V(\bar{X} - \bar{Y}) = \sigma_X^2 / n_X + \sigma_Y^2 / n_Y$$

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2 / n_X + \sigma_Y^2 / n_Y}} = Z$$

$$\frac{\bar{X} - \bar{Y}}{\sqrt{s_X^2 / n_X + s_Y^2 / n_Y}} = t_\nu$$

(We will show the degrees of freedom later. Often near $n_X - 1 + n_Y - 1$)

Tests for a Difference of Means

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\sigma_X^2 / n_X + \sigma_Y^2 / n_Y}} = Z$$

$$\frac{\bar{X} - \bar{Y}}{\sqrt{s_X^2 / n_X + s_Y^2 / n_Y}} = t_\nu$$

A test that the difference in means is zero can be based on this statistic.

If n is large, we can simply use the normal distribution. If not, we use

$$\nu = \frac{(s_X^2 / n_X + s_Y^2 / n_Y)^2}{(s_X^2 / n_X)^2 / (n_X - 1) + (s_Y^2 / n_Y)^2 / (n_Y - 1)}$$

If the samples are of equal size n, then this reduces to

$$\nu = \frac{(n-1)(s_X^2 + s_Y^2)^2}{(s_X^2)^2 + (s_Y^2)^2}$$

and if the sample variances are both equal to s^2, this is

$$\nu = 2(n-1)$$

Effect of C-ABC on the Strength of Engineered Cartilage

The $n_X = 6$ treated samples X measured 462, 747, 571, 599, 373, 413.

$\bar{X} = 527.44,\ s_X = 138.93,\ s_X^2 = 19300$

The $n_Y = 7$ control samples Y measured 364, 224, 183, 165, 163, 275, 293.

$\bar{Y} = 238.15,\ s_Y = 75.79,\ s_Y^2 = 5744$

The t-statistic to test for a difference is

$$\frac{527.44 - 238.15}{\sqrt{19300/6 + 5744/7}} = \frac{289.29}{63.54} = 4.553$$

$$\nu = \frac{(3217 + 820.57)^2}{(3217)^2/5 + (820.57)^2/6} = 7.47$$

$\nu = 7 \qquad p = 0.0026$

$\nu = 7.47 \qquad p = 0.0022$

you need to round ν down if you use a table not a computer

Assumption: Equal Variance

- If we could assume that the population variance of *X* and the population variance of *Y* are the same, then there is another version of the t-test for a difference in means.
- This is hardly ever the best option. If $\sigma_X = \sigma_Y$ then it is slightly better.
- If $\sigma_X \neq \sigma_Y$ then it can be a lot worse, so there is little to gain and much to lose.
- In small to moderate size samples, you can't tell if the variances differ. In large samples, we always use the form that does not assume equal variances anyway.
- Bottom line: always use the version of the two-sample t-test that does not assume equality of variances.

MATLAB

```
>> control'
  364.1265  223.7607  183.0584  165.3209  162.8021  274.7644  293.1934

>> treat'
  461.5631  747.0499  570.8312  599.1513  373.4106  412.6264

>> [h,p,ci,stats] = ttest2(treat,control,'Vartype','unequal')
h =
     1
p =
    0.0022
ci =
  140.9449
  437.6393
stats =
    tstat: 4.5530
       df: 7.4712
       sd: [2x1 double]
>> getfield(stats,'sd')

  138.9250
   75.7892
```

Difference of Proportions

Suppose we have two methods of inducing stem cell properties
in fibroblasts, the standard method, and a new method. In an experiment,
5 out of 200 cells were converted using the standard method and
12 out of 250 were converted using the new method.
We want to test the hypothesis that the new method is the same as the old one.

$$E(\hat{p}_{\text{standard}}) = p_{\text{standard}} \qquad E(\hat{p}_{new}) = p_{new}$$

$$V(\hat{p}_{standard}) = p_{standard}(1 - p_{standard}) / n_{standard}$$

$$V(\hat{p}_{new}) = p_{new}(1 - p_{new}) / n_{new}$$

A test of the hypothesis of equality of proportions is given by

$$z = \frac{(\hat{p}_{new} - \hat{p}_{standard}) - 0}{\sqrt{p_{pooled}(1 - p_{pooled})(1/n_{new} + 1/n_{standard})}}$$

where

$$p_{pooled} = \frac{x_{new} + x_{standard}}{n_{new} + n_{standard}} = \frac{17}{450} = 0.03778$$

$$\sqrt{(0.03778)(0.96222)(1/200 + 1/250)} = 0.01809$$

$$0.023 / 0.01809 = 1.2716$$

$$p = 2(0.1018) = 0.2035$$

This is consistent with the proportions being equal.

Under the null, both samples come from the same binomial distribution.

This is different from the confidence interval.

For a confidence interval, we use a standard error of the difference of

$$\sqrt{\frac{\hat{p}_{new}(1-\hat{p}_{new})}{n_{new}}+\frac{\hat{p}_{standard}(1-\hat{p}_{standard})}{n_{standard}}}$$

which does not assume that the proportions are equal.

For a hypothesis test of no difference, we assume the proportions are equal and use the standard error

$$\sqrt{p_{pooled}(1-p_{pooled})(1/n_{new}+1/n_{standard})}$$

where

$$p_{pooled}=\frac{x_{new}+x_{standard}}{n_{new}+n_{standard}}$$

Paired Data

- For a two sample test, we have two independent samples.
- Sometimes, we are interested in a difference of means when the X and Y values each are paired up, meaning associated with one and only one of the other.
  - A drug treatment is supposed to reduce arthritis pain. For each patient, we have a pain measure X before and a pain measure Y after treatment.
  - The standard and mini Wright meters measure peak air flow. Each subject has one measurement with each meter, so they are paired.

We have n pairs of data x_i, y_i and we want to know if the means of X and Y differ.

$d_i = x_i - y_i$

$\mu_D = \mu_X - \mu_Y$

$H_0 : \mu_D = 0$

We now have a one-sample problem with a null hypothesis that the mean is 0.

$$\frac{\bar{d} - 0}{s_d / \sqrt{n}} = t_{n-1}$$

For the wright data, there are 17 pairs of values, mini and standard, which leads to 17 differences (mini – standard)

$\bar{d} = 2.117$

$s_d = 38.77$

$s_d / \sqrt{17} = 38.77 / 4.123 = 9.402$

$t_{16} = 2.117 / 9.402 = 0.225$

$p = 0.8246$

No significant difference