An example from Whitlock and Schluter's lab:

"In 1898, Hermon Bumpus collected data on one of the first examples of natural selection directly observed in nature. Immediately following a bad winter storm, 136 English house sparrows were collected and brought indoors. Of these, 72 subsequently recovered, but 64 died. Bumpus made several measurements on all of the birds, and he was able to demonstrate strong natural selection on some of the traits as a result of this storm."

Measured body mass in sparrows collected after storm.

H_0: Mass did not effect the chance of survival.

H_0: Mass and survival probability are unrelated, so the mean mass in the population of dead and surviving sparrows will be the same.

$$\mu_d = \mu_s$$

H_A: Mass and survival probability are not independent:

$$\mu_d \neq \mu_s$$

Died:

$n_d = 64$ $\bar{Y}_d = 25.86g$ $s_d = 1.63g$

Survived:

$n_s = 72$ $\bar{Y}_s = 25.22g$ $s_s = 1.26g$

Test statistic:

$$\bar{Y}_d - \bar{Y}_s$$

should be close to 0 when H_0 is true.

We can assume that the distribution of mass is approximately normal.

How much big should $\bar{Y}_d - \bar{Y}_s$ be if the null is true?

We should scale $\bar{Y}_d - \bar{Y}_s$ by a standard error of the difference.

A derivation:

$$X_1 \sim \text{normal}\,(\mu = \mu_1, \sigma = \sigma_1)$$

$$X_2 \sim \text{normal}\,(\mu = \mu_2, \sigma = \sigma_2)$$

$$X_1 + X_2 \sim \text{normal}\left(\mu = \mu_1 + \mu_2, \sigma = \sqrt{\sigma_1^2 + \sigma_2^2}\right)$$

$$\bar{Y}_1 \sim \text{normal}\left(\mu = \mu_1, \sigma = \frac{\sigma_1}{\sqrt{n_1}}\right)$$

$$\bar{Y}_2 \sim \text{normal}\left(\mu = \mu_2, \sigma = \frac{\sigma_2}{\sqrt{n_2}}\right)$$

$$\bar{Y}_1 - \bar{Y}_2 \sim \text{normal}\left(\mu = \mu_1 - \mu_2, \sigma = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}\right)$$

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

Assuming that the $\sigma_1 \approx \sigma_2$, and recognizing that we have to estimate the variances:

$$SE_{\bar{Y}_1-\bar{Y}_2} = \sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

s_p^2 is the pooled sample variance:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

Your book writes this as:

$$s_p^2 = \frac{df_1 s_1^2 + df_2 s_2^2}{df_1 + df_2}$$

$$n_d = 64 \qquad \bar{Y}_d = 25.86g \qquad s_d = 1.63g$$

$$n_s = 72 \qquad \bar{Y}_s = 25.22g \qquad s_s = 1.26g$$

$$s_p^2 = \frac{(n_d - 1)s_d^2 + (n_s - 1)s_s^2}{n_d + n_s - 2}$$

$$s_p^2 = \frac{(63)(1.63)^2 + (71)(1.26)^2}{64 + 72 - 2} = 2.09$$

$$\begin{aligned} SE_{\bar{Y}_d - \bar{Y}_s} &= \sqrt{s_p^2 \left(\frac{1}{n_d} + \frac{1}{n_s}\right)} \\ &= \sqrt{2.09 \left(\frac{1}{64} + \frac{1}{72}\right)} = 0.248 \end{aligned}$$

$$t = \frac{\bar{Y}_d - \bar{Y}_s}{SE_{\bar{Y}_d - \bar{Y}_s}}$$

$$t = \frac{25.86 - 25.22}{0.248} = \frac{0.64}{0.248} = 2.58$$

We look up the critical value:

$$t_{0.05(2),df}$$

where $df = n_1 + n_2 - 2$

$$t_{0.05(2),134} \approx 1.98$$

$$t_{0.01(2),134} \approx 2.61$$

95% confidence interval for the difference between two sample means.

We are 95% confident that $\mu_1 - \mu_2$ is between:

$$(\bar{Y}_1 - \bar{Y}_2) - \left(SE_{\bar{Y}_1 - \bar{Y}_2}\right) t_{0.05(2),df}$$

and

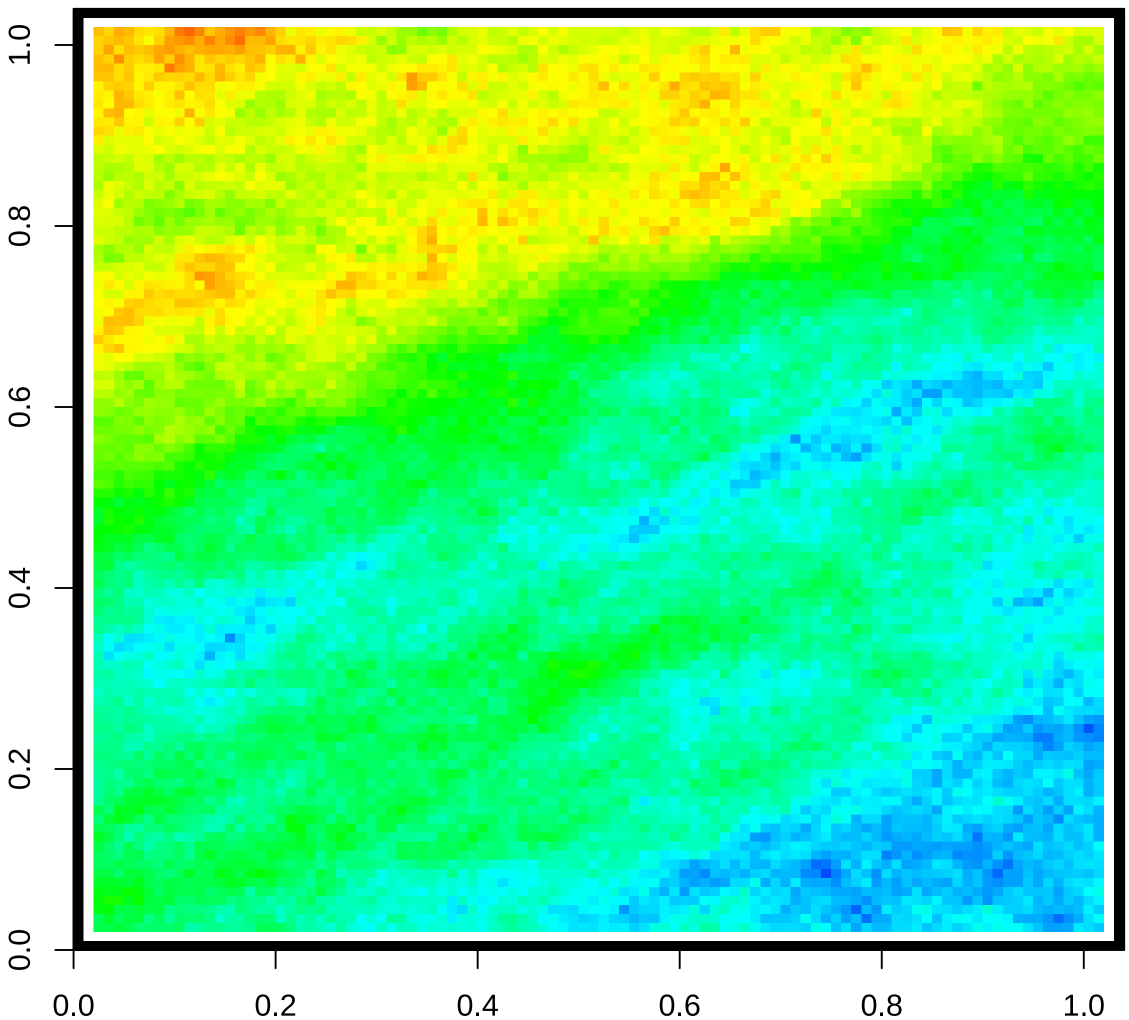
$$(\bar{Y}_1 - \bar{Y}_2) + \left(SE_{\bar{Y}_1 + \bar{Y}_2}\right) t_{0.05(2),df}$$

In the sparrow example, we are 95% confident that:

$$0.64 - 0.248 * 1.98 < \mu_d - \mu_s < 0.64 - 0.248 * 1.98$$

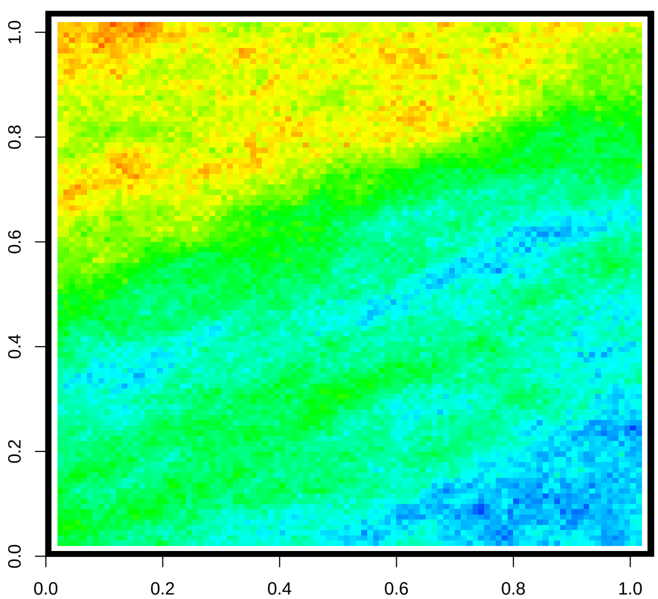
$$0.143 < \mu_d - \mu_s < 1.126$$

red ≈ 10 expected
blue ≈ 250 expected

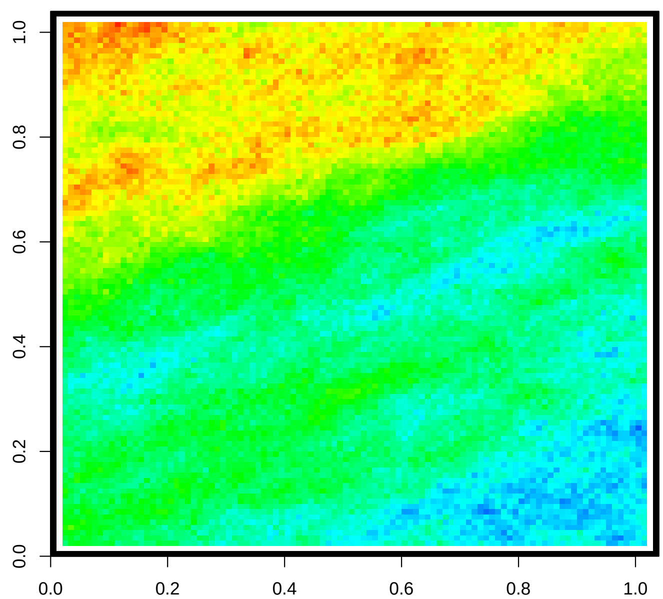


If we could treat the entire landscape, we would see a small effect:

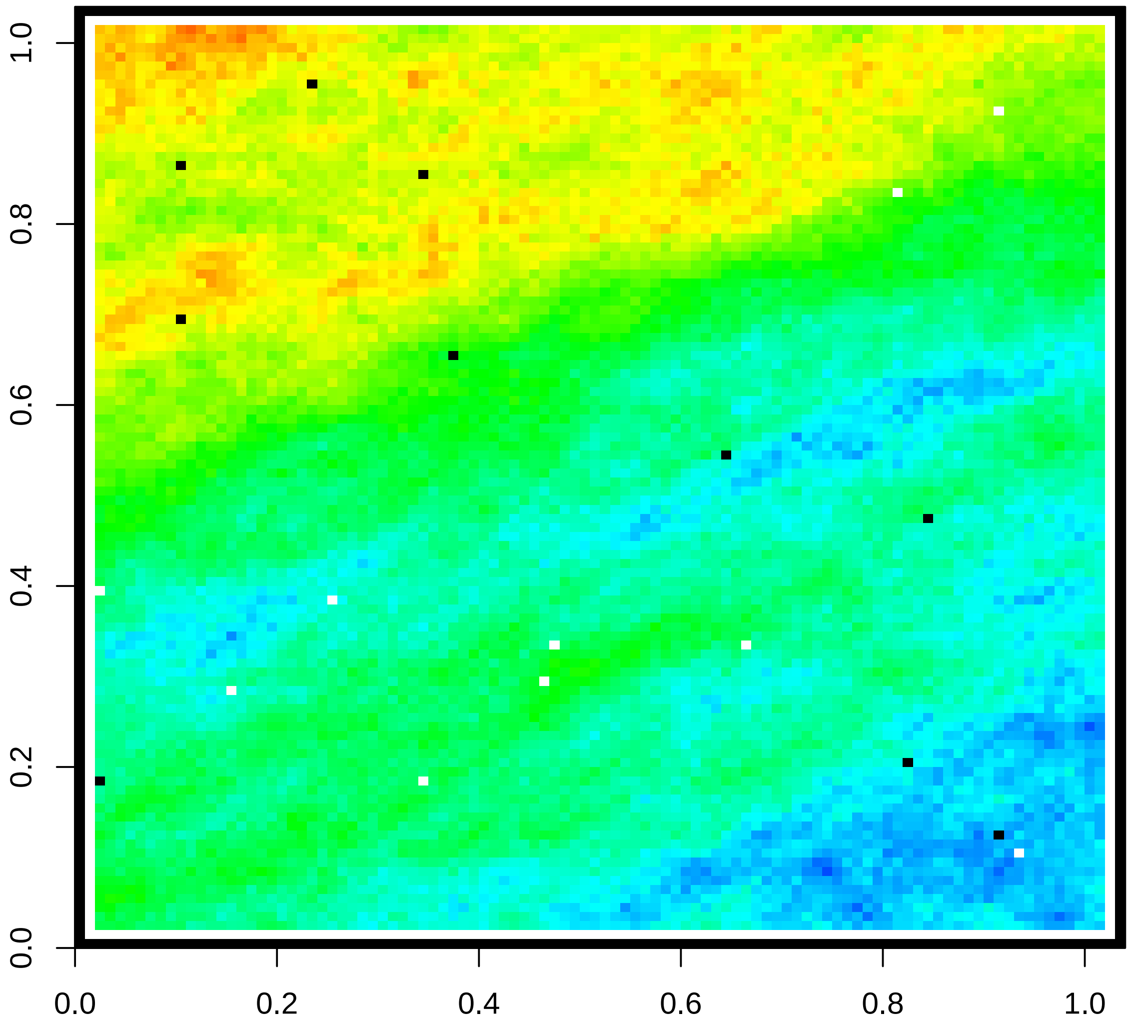
Control:



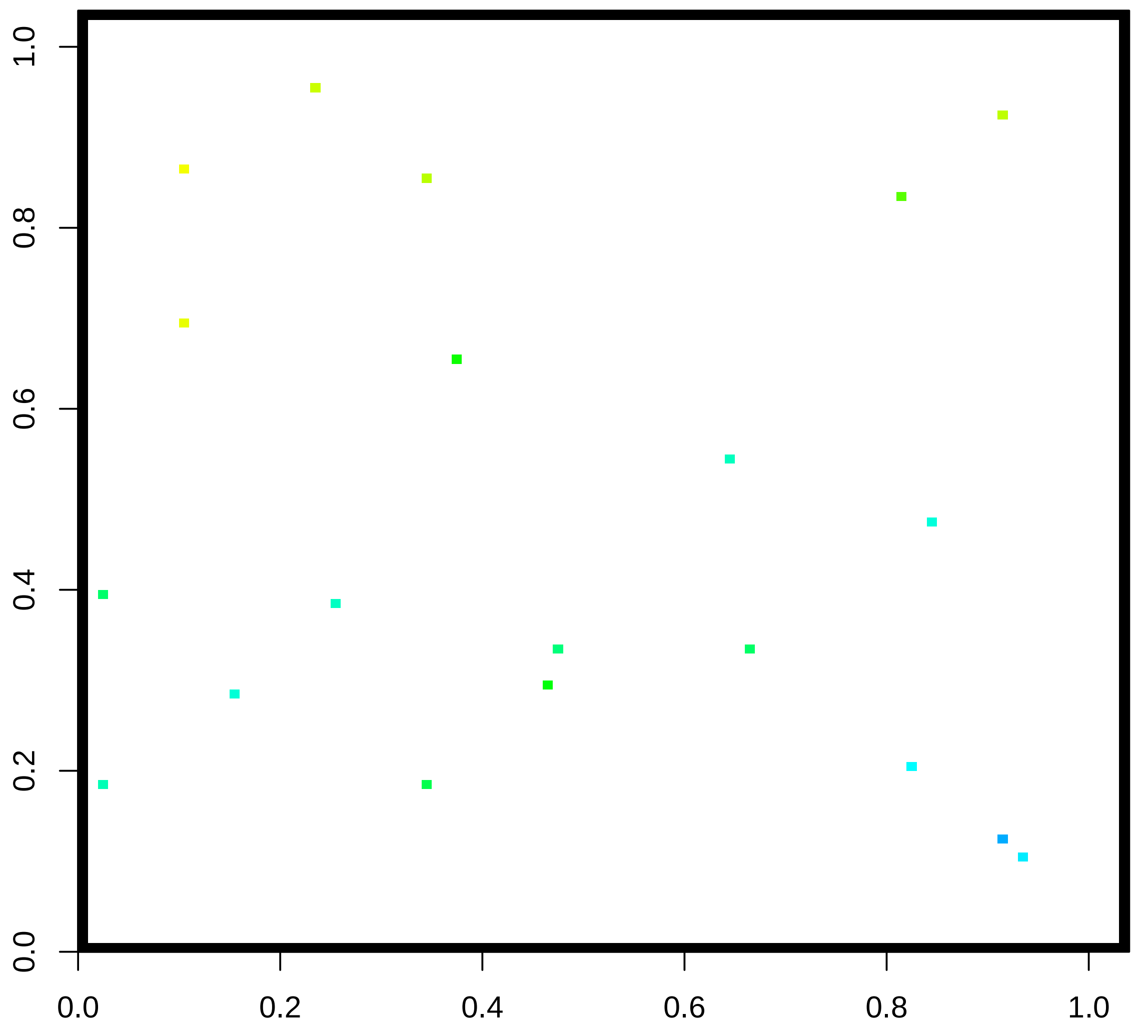
Treatment:



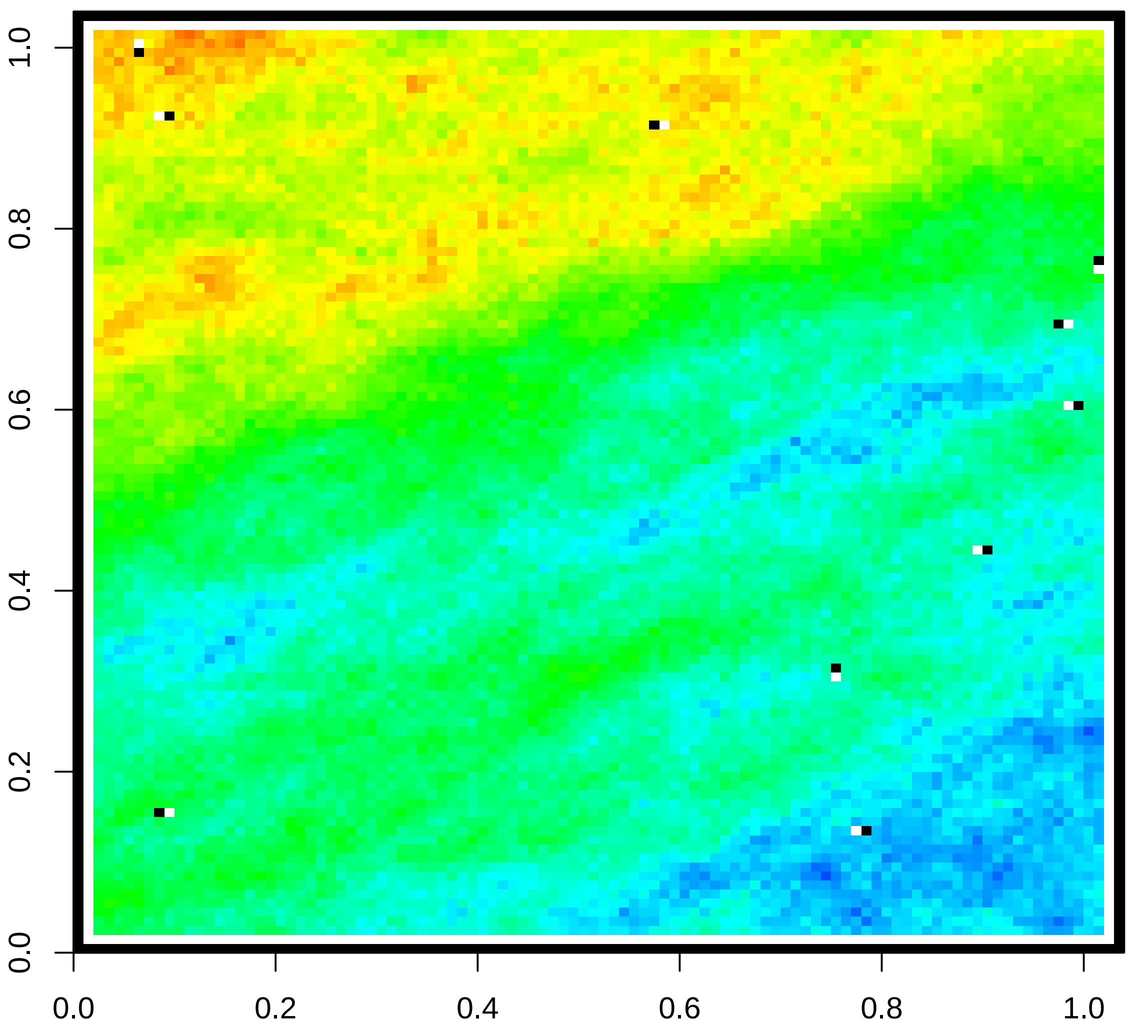
We could randomly select 10 control locations and 10 treatment locations:



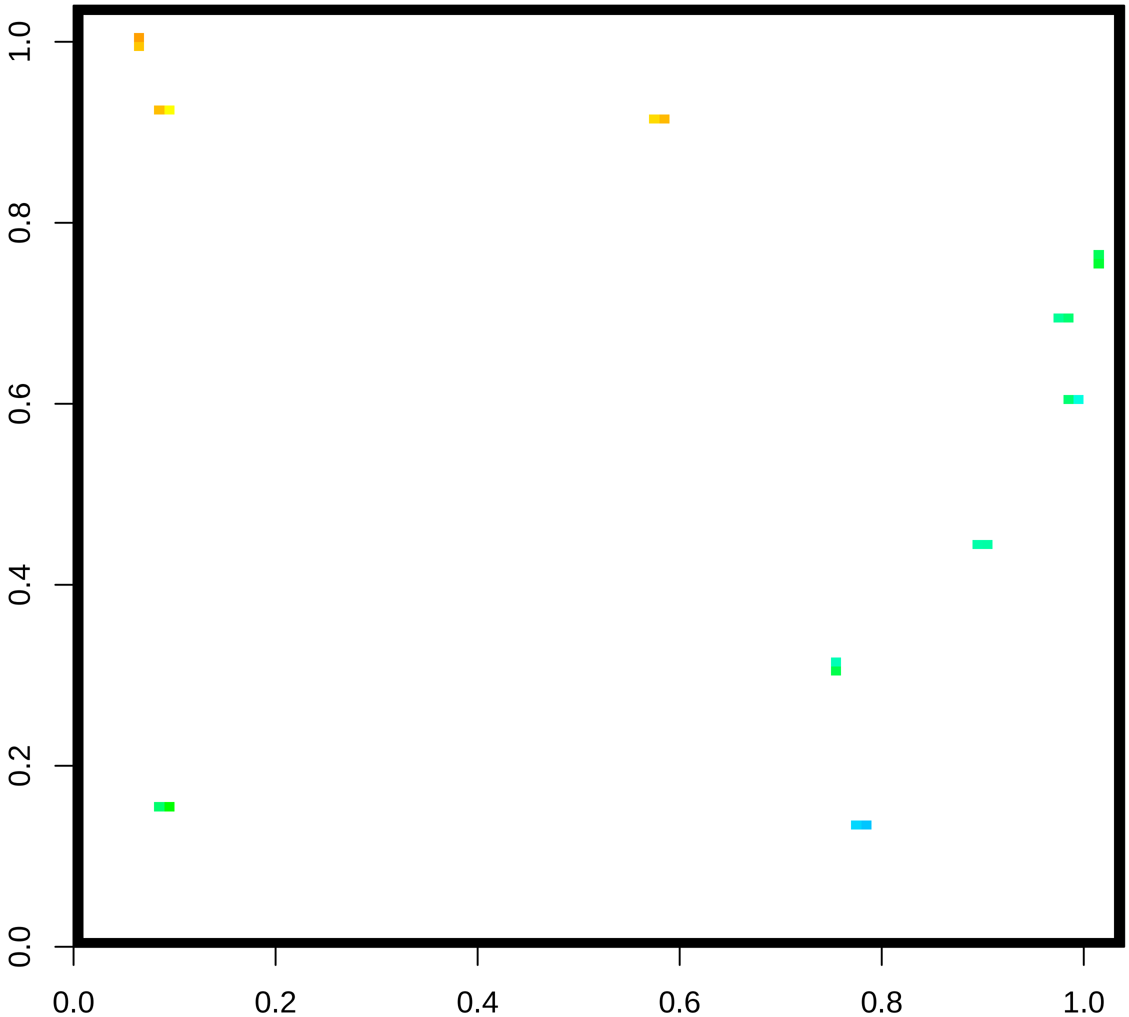
Only the sampled locations shown:



In a paired experimental design, we look at the difference between treatment and control from data that is naturally paired:



Only the sampled locations shown:



Paired		
Treatment	Control	Difference
39.6	48.5	-8.9
45.5	53.4	-7.9
140.9	166.0	-25.1
149.5	175.3	-25.8
192.2	196.5	-4.3
162.6	162.0	0.6
46.6	62.2	-15.6
122.0	148.2	-26.2
134.7	143.3	-8.6
149.7	158.7	-9.0

$$\bar{d} = -13.08$$

$$s_d = 9.60$$

$$n = 10$$

$$SE_{\bar{d}} = 3.04$$

$\bar{d} = -13.08 \quad s_d = 9.60 \quad n = 10$

$$SE_{\bar{d}} = \frac{9.6}{\sqrt{10}} = 3.04$$

$$t = \frac{\bar{d} - \mu_0}{SE_{\bar{d}}} = \frac{-13.08}{3.04} = -4.31$$

$$df = 9$$

$$P\text{-value} = 0.00197$$

We reject the null hypothesis that our treatment has no effect on the biomass of bark beetles. We compared 10 geographically-paired plots. Treated areas had an average of 13.08 kg less bark beetles (standard error of the difference = 3.04). This difference is too large to be explained by sampling error ($P < 0.002$).

Unpaired			Paired	
Treatment	Control		Treatment	Control
173.9	202.0		39.6	48.5
187.3	119.9		45.5	53.4
168.0	74.1		140.9	166.0
77.1	167.1		149.5	175.3
151.2	64.7		192.2	196.5
101.5	174.8		162.6	162.0
147.9	79.0		46.6	62.2
146.0	183.6		122.0	148.2
140.8	67.5		134.7	143.3
124.3	165.5		149.7	158.7

Unpaired:
t = 0.5929, df = 18, P-value = 0.5606
95 % confidence interval of the difference in means:
(-30.5, 54.4)

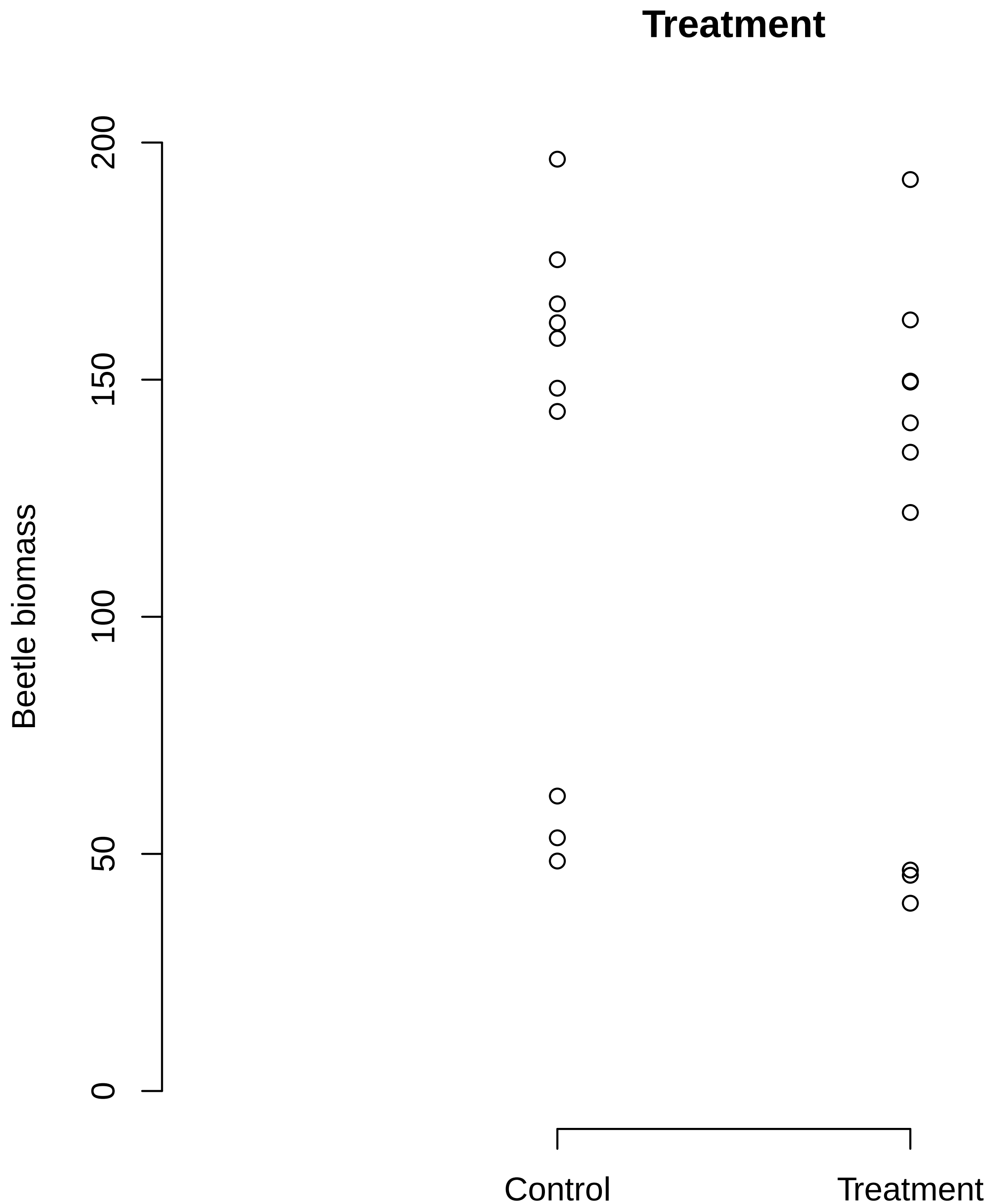
Paired: t = -4.3071, df = 9, P-value = 0.00197
95 % confidence interval of the difference in means:
(-19.95, -6.21)

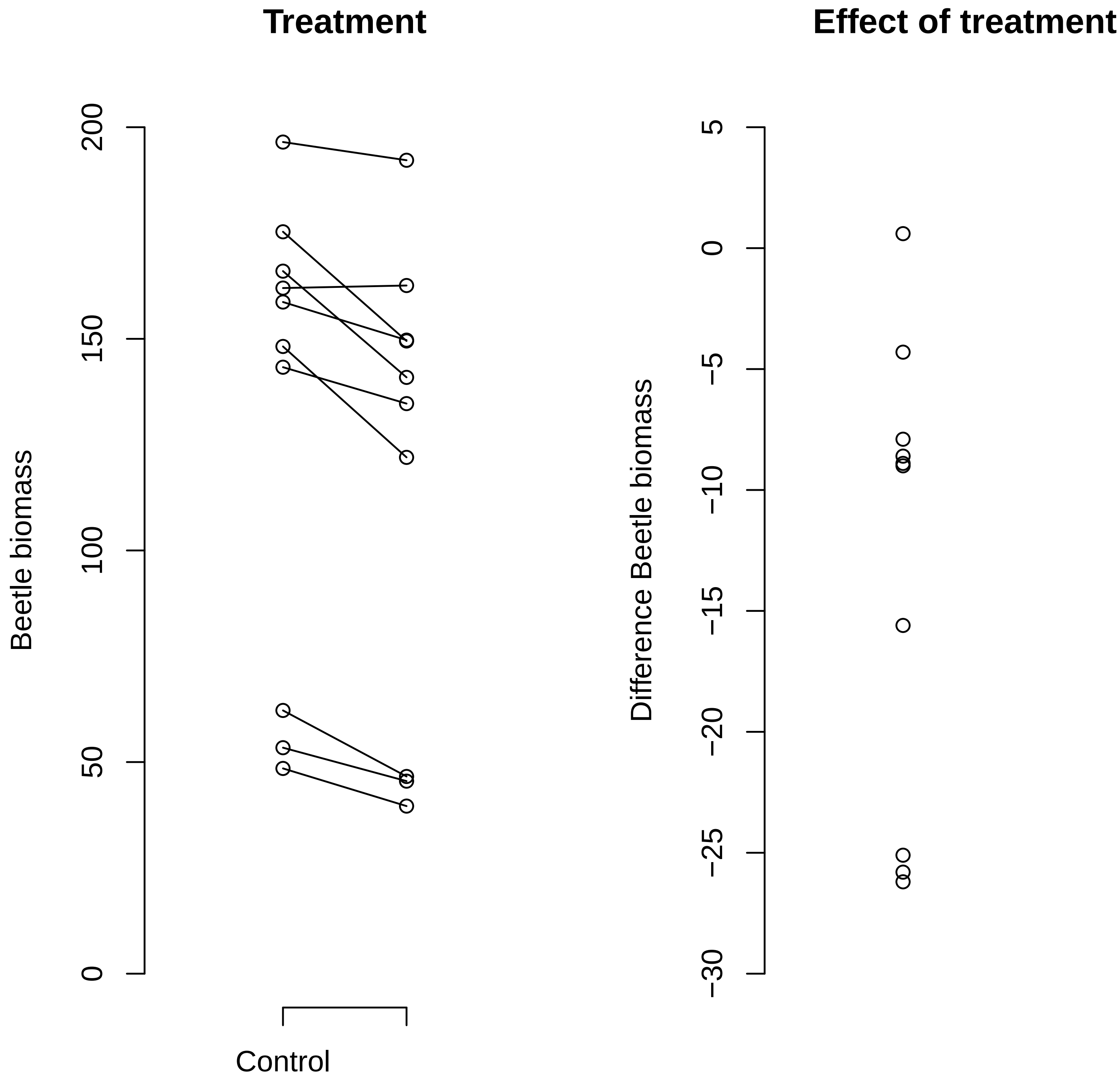
95 % confidence interval of the difference in means:

$$\bar{d} - \left(t_{0.05(2),9}\right) SE_{\bar{d}} < \mu_t - \mu_c < \bar{d} + \left(t_{0.05(2),9}\right) SE_{\bar{d}}$$

$$-13.08 - (2.26)3.04 < \mu_t - \mu_c < -13.08 + (2.26)3.04$$

$$-19.95 < \mu_t - \mu_c < -6.21$$


Treatment
Beetle biomass
0
50
100
150
200
Control
Treatment


Treatment
Effect of treatment
Beetle biomass
200
150
100
50
0
Control
Difference Beetle biomass
5
0
−5
−10
−15
−20
−25
−30

Ignoring pairing:

$$\bar{Y}_T = 118.33 \qquad \sigma_T^2 = 2978.991 \qquad n_T = 10$$

$$\bar{Y}_C = 131.41 \qquad \sigma_C^2 = 3023.659 \qquad n_C = 10$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s_p^2 = \frac{(9)2978.991 + (9)3023.659}{18} = 3001.325$$

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$SE_{\bar{Y}_T - \bar{Y}_C} = \sqrt{3001.325 \left(\frac{1}{10} + \frac{1}{10}\right)} = 24.50$$

$$t = \frac{\bar{Y}_T - \bar{Y}_C}{SE_{\bar{Y}_1 - \bar{Y}_2}} = \frac{-13.08}{24.50} = -0.53$$

$$df = 18 \qquad P \approx 0.3$$

Consider paired and unpaired analyses of the same (paired) data:

Paired		**Unpaired**
Reject H_0 $t = \frac{\bar{d}}{SE_{\bar{d}}}$ $df = 9$		Do not reject H_0 $t = \frac{\bar{Y}_T - \bar{Y}_C}{SE_{\bar{Y}_1 - \bar{Y}_2}}$ $df = 18$
$t = \frac{\bar{d}}{\left(\frac{s_d}{\sqrt{n}}\right)}$ $t = \frac{-13.08}{3.04}$		$t = \frac{\bar{Y}_T - \bar{Y}_C}{\sqrt{s_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$ $t = \frac{-13.08}{24.50}$

Paired		
Treatment	Control	Difference
39.6	48.5	-8.9
45.5	53.4	-7.9
140.9	166.0	-25.1
149.5	175.3	-25.8
192.2	196.5	-4.3
162.6	162.0	0.6
46.6	62.2	-15.6
122.0	148.2	-26.2
134.7	143.3	-8.6
149.7	158.7	-9.0

$\bar{Y}_T = 141.80 \qquad \sigma_T^2 = 1116.4 \qquad n_T = 10$

$\bar{Y}_C = 129.82 \qquad \sigma_C^2 = 2966.5 \qquad n_C = 10$

$$\bar{Y}_T = 141.80 \qquad \sigma_T^2 = 1116.4 \qquad n_T = 10$$

$$\bar{Y}_C = 129.82 \qquad \sigma_C^2 = 2966.5 \qquad n_C = 10$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$$s_p^2 = \frac{(9)1116.4 + (9)2966.5}{18} = 2041.5$$

$$SE_{\bar{Y}_1 - \bar{Y}_2} = \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

$$SE_{\bar{Y}_T - \bar{Y}_C} = \sqrt{2041.5 \left(\frac{1}{10} + \frac{1}{10} \right)} = 20.21$$

$$t = \frac{\bar{Y}_T - \bar{Y}_C}{SE_{\bar{Y}_1 - \bar{Y}_2}} = \frac{11.98}{20.21} = 0.593$$

$$df = 18$$

If you are a doctor who wants to reduce the level of cholesterol in the bloodstream of a patient, should you prescribe exercise or drug XYZ?

(fake) data from a (fake) study:

- randomly select trial subjects,
- measure their cholesterol level,
- randomly assign them to a treatment (exercise or XYZ),
- treat for 6 months
- measure their cholesterol level after the treatment,
- **report the difference** in cholesterol level for each patient.

H_0: $\mu_e = 0$
H_A: $\mu_e \neq 0$

H_0: $\mu_d = 0$
H_A: $\mu_d \neq 0$

H_0: $\mu_e = \mu_d$
H_A: $\mu_e \neq \mu_d$

Change in cholesterol level by treatment (measured in mg/dL):

Exercise	Drug XYZ
-2.3	-5.2
-0.4	-2.6
5.0	-1.2
-13.8	0.9
-12.3	-7.2
	-4.1
	-9.2
	1.5
	-4.1
	-3.7

Exercise:

$n_e = 5$ $\bar{Y}_e = -4.76$ $s_e = 8.045$

Drug XYZ:

$n_d = 10$ $\bar{Y}_d = -3.49$ $s_d = 3.34$

Is each treatment effective?

H_0: exercise does not affect cholesterol level
$\mu_e = 0.0$
H_A: $\mu_e \neq 0.0$

H_0: drug XYZ does not affect cholesterol level
$\mu_d = 0.0$
H_A: $\mu_d \neq 0.0$

Exercise:
$n_e = 5$ $\quad$ $\bar{Y}_e = -4.76$ $\quad$ $s_e = 8.045$

Drug XYZ:
$n_d = 10$ $\quad$ $\bar{Y}_d = -3.49$ $\quad$ $s_d = 3.34$

$$t = \frac{\bar{Y} - \mu_0}{\left(\frac{s}{\sqrt{n}}\right)} =$$

$$t_e = \frac{-4.76}{\left(\frac{8.045}{\sqrt{5}}\right)} = -1.323$$

$$t_{0.05(2),4} = 2.78$$

$$t_d = \frac{-3.49}{\left(\frac{3.34}{\sqrt{10}}\right)} = -3.3061$$

$$t_{0.05(2),9} = 2.26$$

$$t_{0.01(2),9} = 3.25$$

We do not reject the null hypothesis that exercise does not have a mean effect on cholesterol ($P > 0.2$) based on these data ($n = 5$, mean change = -4.65mg/dL).

We found evidence that drug XYZ lowers cholesterol ($n = 10$, mean change = -3.49mg/dL). This drop is greater than would be expected if it were the result sampling error ($P < 0.01$).

We found evidence that drug XYZ is more effective than exercise at lowering cholesterol. The drug had a significant effect on cholesterol ($P < 0.01$), while exercise had no contribution that could not be explained by sampling error($P > 0.2$).

Exercise:

$n_e = 5 \qquad \bar{Y}_e = -4.76 \qquad s_e = 8.045$

Drug XYZ:

$n_d = 10 \qquad \bar{Y}_d = -3.49 \qquad s_d = 3.34$

$$t = \frac{\bar{Y}_e - \bar{Y}_d}{SE_{\bar{Y}_e - \bar{Y}_d}}$$

$$SE_{\bar{Y}_e - \bar{Y}_d} = \sqrt{s_p^2 \left(\frac{1}{n_e} + \frac{1}{n_d} \right)}$$

$$s_p^2 = \frac{(n_e - 1)s_e^2 + (n_d - 1)s_d^2}{n_e + n_d - 2}$$

$$s_p^2 = \frac{4(64.72) + 9(11.14)}{5 + 10 - 2} = 27.62$$

$$SE_{\bar{Y}_e - \bar{Y}_d} = \sqrt{s_p^2 \left(\frac{1}{n_e} + \frac{1}{n_d} \right)}$$

$$SE_{\bar{Y}_e - \bar{Y}_d} = \sqrt{27.62 \left(\frac{1}{5} + \frac{1}{10} \right)} = 2.88$$

$$t = \frac{-4.76 + 3.49}{2.88} = \frac{-1.27}{2.88} = 0.44$$

$$t_{0.05(2),13} = 2.16$$

Avoid indirect comparisons:
If the mean effect of A is significantly different from 0, but the mean effect of B is not significantly different from 0, we **cannot** conclude that the effect of A is larger than the effect of B.

We have to A vs B using a two sample t-test.

Change in cholesterol level before and after execise treatment (measured in mg/dL):

Before	After	Difference
260.5	258.2	-2.3
263.1	262.7	-0.4
258.9	263.9	5.0
271.7	257.9	-13.8
287.6	275.3	-12.3

To do a paired t-test we simply treat the difference as variable with and do a one-sample t-test!