Stat 411/511

Levene's Test & Welch's t-Test

Feb 8 2012

Charlotte Wickham

stat511.cwick.co.nz

Midterm

In class Friday, 17th 1pm to 1.50pm.

I will assign no lab work next Friday.

There will be a homework due.

Wilcoxon Rank Sum

	Wilcoxon Rank Sum test
Null hypothesis*	The difference in population medians is zero. OR The treatment effect is zero.
Assumptions	•Equal population standard deviations •Independence of subjects within and between groups.
Robust to assumptions?	•Robust to inequality of variances if sample sizes are equal.
Resistant to outliers?	Resistant
Test statistic	Sum of the ranks in the smaller group

Equality of standard deviations

The **Wilcoxon Rank Sum** test, and the **two-sample *t*-test** both assume the population standard deviations are the same. Why not the paired *t*-test?

Levene's test provides a formal way to test this assumption.

Welch's *t*-test is an alternative two-sample *t*-test for when the population standard deviations are different.

Levene's Test

Levene's test provides a formal way to test the assumption of equal population standard deviations.

Null: The standard deviation of population A is the same as the standard deviation of population B, $\sigma_1 = \sigma_2$

Alternative: The standard deviation of population A is **not** the same as the standard deviation of population B, $\sigma_1 \neq \sigma_2$

Levene's test procedure

Do a **two sample t-test** on

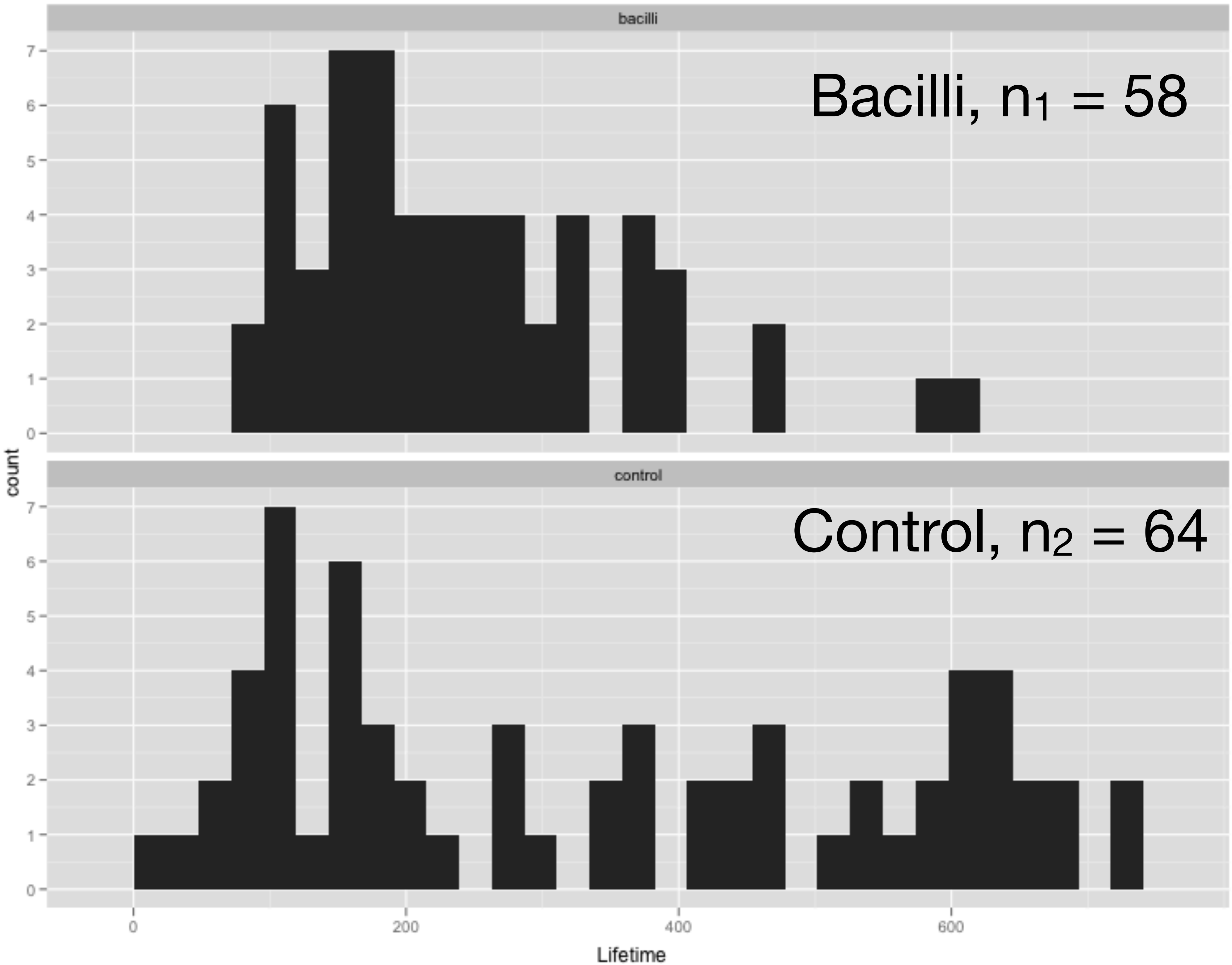
$Z_1 = (Y_1 - \overline{Y}_1)^2$, and $Z_2 = (Y_2 - \overline{Y}_2)^2$

A **small** p-value is **evidence** that the populations have **different** standard deviations, and we need an alternative to the standard t-test.

Why this transform? $\hat{\sigma} = s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2}$

```
qplot(Lifetime, data = ex0211) + facet_wrap(~ Group, ncol = 1)
```

The data are survival times (in days) of guinea pigs that were randomly assigned either to a control group or to a treatment group that received a dose of tubercle bacilli.



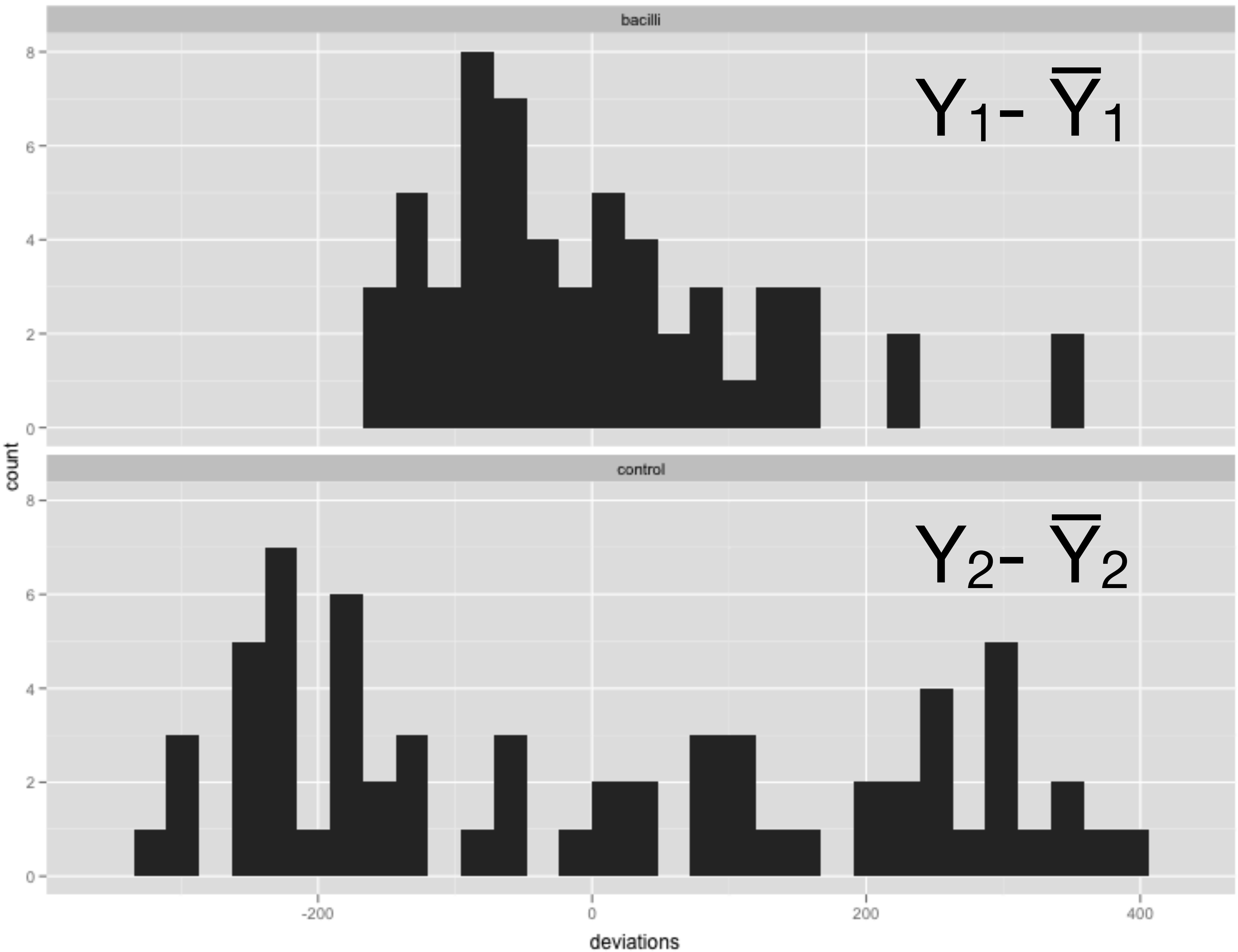
We are interested in whether giving guinea pigs TB, changes their life span. We would like to do a two sample t-test.

First we need to check the assumptions.

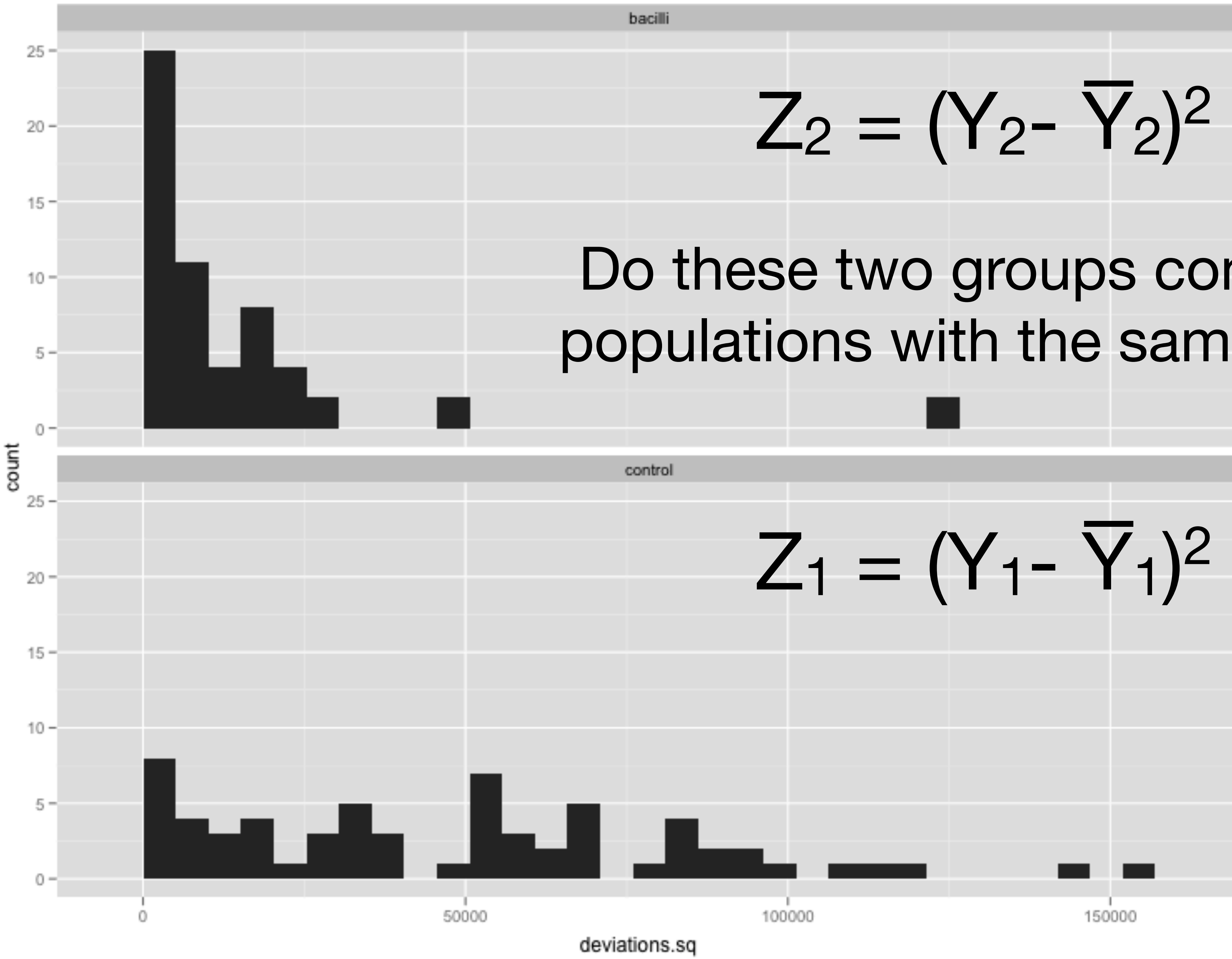
Large sample, so normality is probably ok.

Evidence that the population standard deviations are not equal, **test this!**

Subtract the group average from each group



Then square



$$Z_2 = (Y_2 - \bar{Y}_2)^2$$

Do these two groups come from populations with the same mean?

$$Z_1 = (Y_1 - \bar{Y}_1)^2$$

A variable I made

```
> t.test(deviations.sq ~ Group, data =
guinea_pigs, var.equal = TRUE)

   Two Sample t-test

data:  deviations.sq by Group
t = -6.1143, df = 120, p-value = 1.247e-08
```

very small p-value

```
alternative hypothesis: true difference in
means is not equal to 0

95 percent confidence interval:
 -46244.16 -23620.52

sample estimates:
mean in group bacilli mean in group control
             13667.90              48600.24
```

We have **convincing** evidence the population standard deviation of the treated guinea pigs **is not the same as** the control guinea pigs, (Levene's test, two-sided p-value < 0.0001)

We shouldn't do the usual *t*-test.

Welch's *t*-test

Welch's *t*-test

Allows different population standard deviations.

Instead of a pooled estimate of one standard deviation, we use each sample to estimate its own populations standard deviation.

Leads to a **different standard error** on the difference in averages, and a **different degrees of freedom.**

Same *t*-statistic

The two-sample *t*-ratio:

$$\frac{(\overline{Y_2} - \overline{Y_1}) - (\mu_2 - \mu_1)}{\boxed{\mathrm{SE}_{\overline{Y_2} - \overline{Y_1}}}}$$

Different estimate of standard error

$$SE_{\overline{Y_2} - \overline{Y_1}} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

No exact theory, a *t*-distribution with a special degrees of freedom (see pg 97 in Sleuth) is a good approximation.

No var.equal

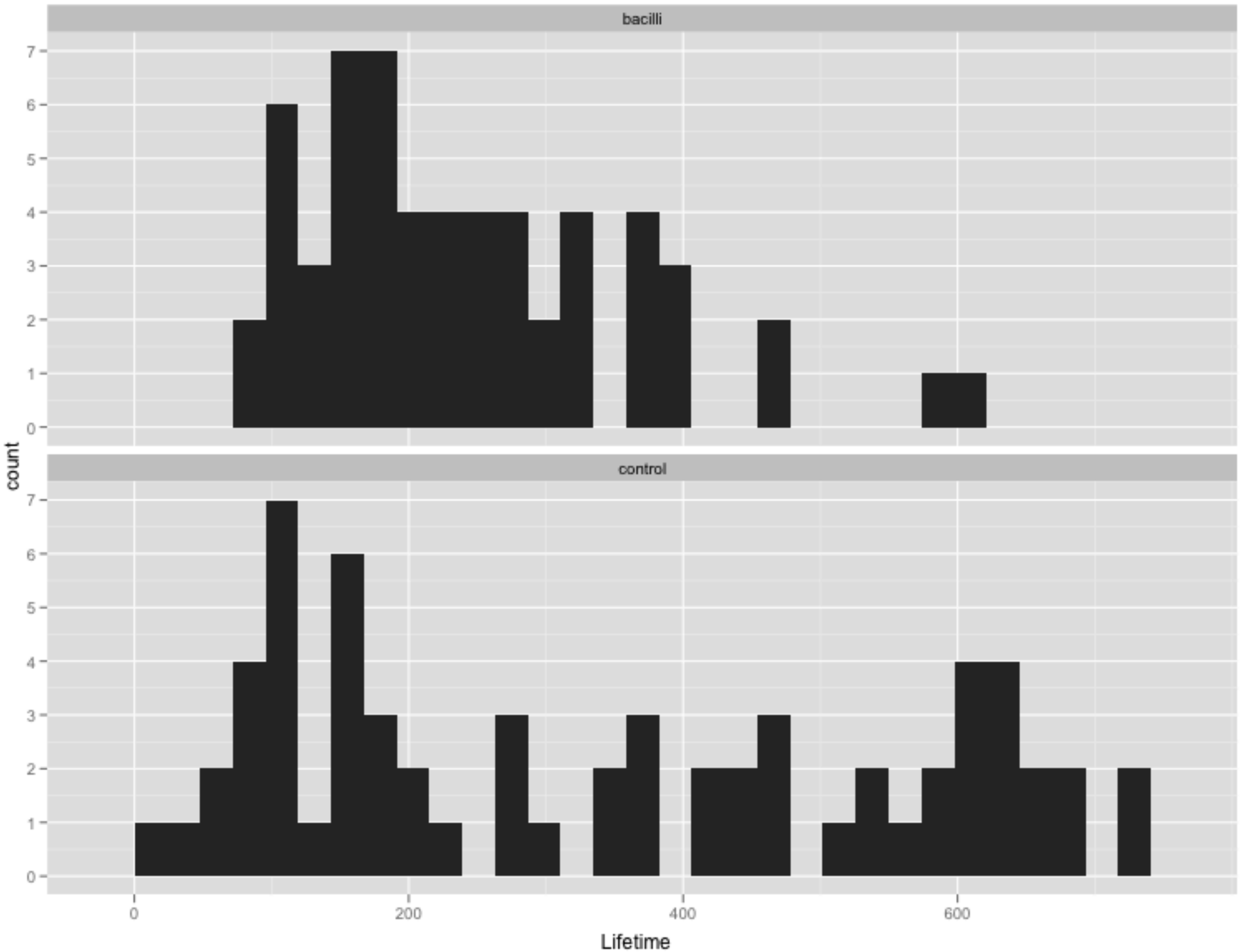
```
> t.test(Lifetime ~ Group, data = ex0211)

	Welch Two Sample t-test

data:  Lifetime by Group
t = -3.2296, df = 97.807, p-value = 0.001689
alternative hypothesis: true difference in means
is not equal to 0
95 percent confidence interval:
 -165.80689  -39.59289
sample estimates:
mean in group bacilli mean in group control
             242.5345              345.2344
```

Small p-value

We have convincing evidence that TB reduces the lifetime of guinea pigs (Welch's t-test, two-sided p-value = 0.002).



Caveat: Additive treatment effect isn't completely adequate.

The reduction in lifetime might depend on your lifetime to begin with.