14.12: On the Relationship Between ANOVA and the Student t Test

There's one last thing I want to point out before finishing. It's something that a lot of people find kind of surprising, but it's worth knowing about: an ANOVA with two groups is identical to the Student t-test. No, really. It's not just that they are similar, but they are actually equivalent in every meaningful way. I won't try to prove that this is always true, but I will show you a single concrete demonstration. Suppose that, instead of running an ANOVA on our `mood.gain ~ drug` model, let's instead do it using `therapy` as the predictor. If we run this ANOVA, here's what we get:

```
summary( aov( mood.gain ~ therapy, data = clin.trial ))
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## therapy      1  0.467  0.4672   1.708   0.21
## Residuals   16  4.378  0.2736
```

Overall, it looks like there's no significant effect here at all but, as we'll see in Chapter @ref(anova2 this is actually a misleading answer! In any case, it's irrelevant to our current goals: our interest here is in the F-statistic, which is F(1,16)=1.71, and the p-value, which is .21. Since we only have two groups, I didn't actually need to resort to an ANOVA, I could have just decided to run a Student t-test. So let's see what happens when I do that:

```
t.test( mood.gain ~ therapy, data = clin.trial, var.equal = TRUE )
```

```
##
##  Two Sample t-test
##
## data:  mood.gain by therapy
## t = -1.3068, df = 16, p-value = 0.2098
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.8449518  0.2005073
## sample estimates:
## mean in group no.therapy        mean in group CBT
##                0.7222222                1.0444444
```

Curiously, the p-values are identical: once again we obtain a value of p=.21. But what about the test statistic? Having run a t-test instead of an ANOVA, we get a somewhat different answer, namely t(16)=−1.3068. However, there is a fairly straightforward relationship here. If we square the t-statistic

```
1.3068 ^ 2
```

```
## [1] 1.707726
```

we get the F-statistic from before.