

Transcript of Mick Crawley's R course 2010 Imperial College London, Silwood Park

Emanuel G Heitlinger

Disclaimer: The following document is a private transcript of Mick Crawley's R-course. I am a participant in this course and my writeup has in no way been approved by Mick Crawley (from whom the ideas behind the code and teaching concepts are) or any of his staff.

Interlude: The perfect error bar

What should error bars tell us?

They should be overlapping if there is no significant difference between the means, non-overlapping if there is significant difference between the means.

What bars are used and what is wrong about it?

In a t-test (assuming normal errors and a equal variance of the two samples) we are using the difference between the two means over the the standard error of the difference:

$$\frac{\bar{x} - \bar{y}}{SE_{difference}}$$

Or more verbosely:

$$\frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

This value (the test statistic) is compared to tables, if it is bigger than the value from tables, we accept the **alternate hypothesis**: significant difference.

Bars of length one standard error (1 s.e.)

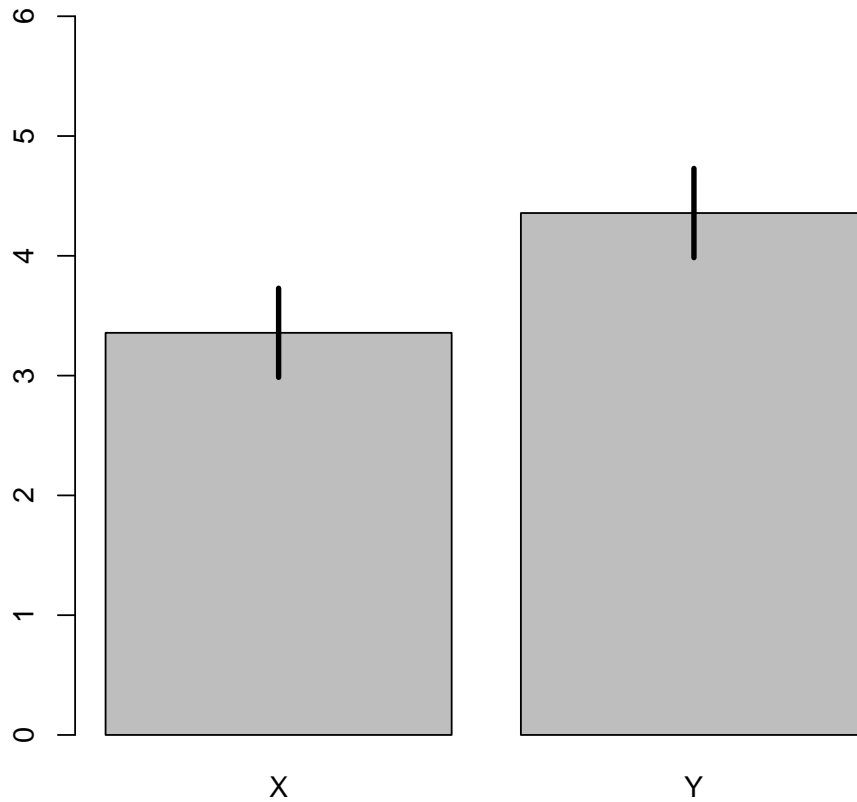
Remember, that the standard error of the mean is:

$$SE_{\bar{y}} = \sqrt{\frac{s^2}{n}}$$

Error bars derived from standard error do not incorporate the value from tables. Instead just 2 individual standard errors are used. 2 is often close to the value from the table, but it is multiplied with the wrong quantity: The standard error of one sample being always lower than the standard error of the difference (about 1.4 times) . Error bars derived from the former can therefore be too short. They could fail indicating non-significance as they could be non-overlapping and still the means non-significantly different.

```
> x <- c(1, 2, 3, 3, 2, 3, 4, 5, 3, 2, 5, 4, 4, 6)
> y <- x + 1
> barplot(c(X = mean(x), Y = mean(y)), ylim = c(0, 6), main = "means with error")
> lines(c(0.7, 0.7), c(mean(x) - sqrt(var(x)/length(x)), mean(x) +
+   sqrt(var(x)/length(x))), lwd = 3)
> lines(c(1.9, 1.9), c(mean(y) - sqrt(var(y)/length(y)), mean(y) +
+   sqrt(var(y)/length(y))), lwd = 3)
```

means with error bars of 1 s.e.



Here are three means, which have non-overlapping standard error bars. Are they significantly different? **No!**

```
> t.test(x, y, var.equal = TRUE)
```

```
Two Sample t-test
```

```
data: x and y
```

```
t = -1.8998, df = 26, p-value = 0.06861
```

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

```
95 percent confidence interval:
```

```
-2.0819977 0.0819977
```

```
sample estimates:
```

```
mean of x mean of y
```

```
3.357143 4.357143
```

Bars of the length of a confidence interval

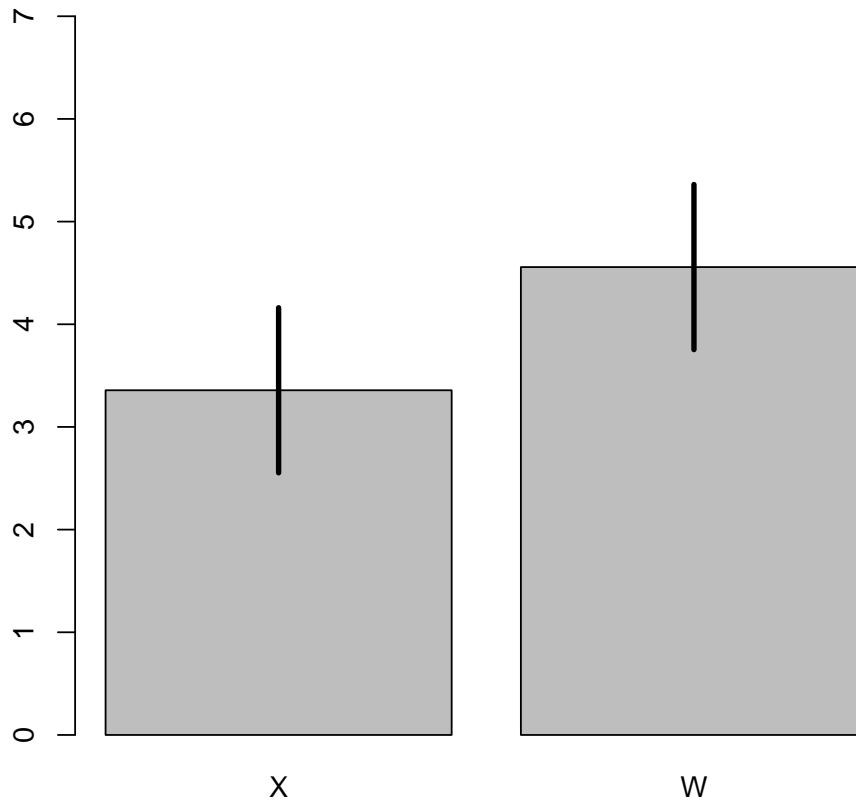
These are given as follows:

$$CI_{95\%} = t_{(\alpha, d.f.)} \sqrt{\frac{s^2}{n}}$$

To generate error bars based on confidence intervals the t-value (from the table) is multiplied with each one of the standard errors. The value from table enters the comparison twice here and inflates the error bars. The fact that we are still using the wrong standard error doesn't save us. It is about 1.4 times smaller than the correct one (of the difference). The one t (from table) too much is always bigger than 1.4, still inflating the the bars too much.

```
> w <- x + 1.2
> barplot(c(X = mean(x), W = mean(w)), ylim = c(0, 7), main = "Error bars of 95% CI",
+        lines(c(0.7, 0.7), c(mean(x) - qt(0.975, length(x) - 1) * sqrt(var(x)/length(x)),
+          mean(x) + qt(0.975, length(x) - 1) * sqrt(var(x)/length(x))),
+          lwd = 3)
> lines(c(1.9, 1.9), c(mean(w) - qt(0.975, length(w) - 1) * sqrt(var(w)/length(w)),
+          mean(w) + qt(0.975, length(w) - 1) * sqrt(var(w)/length(w))),
+          lwd = 3)
```

Error bars of 95% confidence intervals



Here are three other means, which have overlapping 95% confidence interval bars. They are overlapping, but here is the t-test: **significant!**

```
> t.test(x, w, var.equal = TRUE)
```

```
Two Sample t-test
```

```
data: x and w
```

```
t = -2.2797, df = 26, p-value = 0.03107
```

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

```
95 percent confidence interval:
```

```
-2.2819977 -0.1180023
```

```
sample estimates:
```

```
mean of x mean of y
```

```
3.357143 4.557143
```

The perfect error bar

The error bar should have the length of half the **Least Significant Difference (LSD)**

The last significant difference is found inserting in the formula for the t-test statistic.

$$t_{(\alpha,d.f)} = \frac{LSD}{SE_{difference}}$$

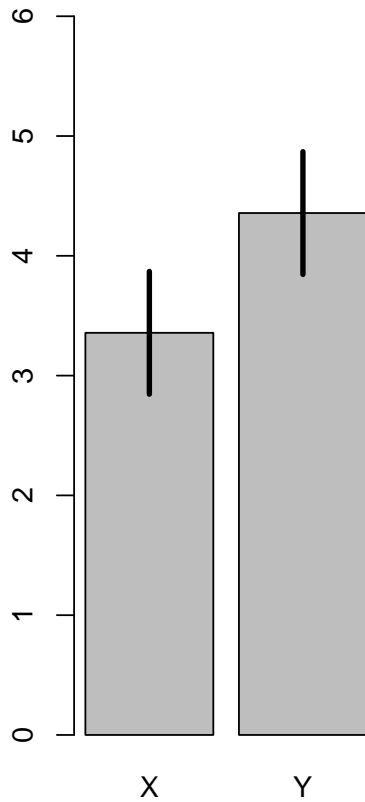
$$LSD = t_{(\alpha,d.f)} SE_{difference}$$

or more verbosely

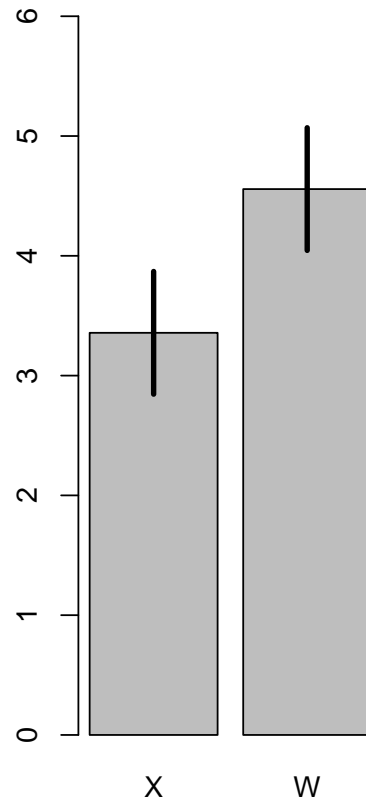
$$LSD = t_{(\alpha,d.f)} \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_A^2}{n_A} + \frac{s_B^2}{n_B}}}$$

```
> LSDline <- function(x, y, loc) {
+   lines(c(loc, loc), c(mean(x) - 2/qt(0.975, length(c(x, y)) -
+     2) * sqrt(var(x)/length(x) + var(y)/length(y)), mean(x) +
+     2/qt(0.975, length(c(x, y)) - 2) * sqrt(var(x)/length(x) +
+     var(y)/length(y))), lwd = 3)
+ }
> par(mfrow = c(1, 2))
> barplot(c(X = mean(x), Y = mean(y)), main = "error bars of length 1/2 LSD",
+   ylim = c(0, 6))
> LSDline(x, y, 0.7)
> LSDline(y, x, 1.9)
> barplot(c(X = mean(x), W = mean(w)), main = "error bars of length 1/2 LSD",
+   ylim = c(0, 6))
> LSDline(x, w, 0.7)
> LSDline(w, x, 1.9)
```

error bars of length 1/2 LSD



error bars of length 1/2 LSD



Remember the t-test's results: **Bingo!**

But what if you want to compare more than two means? What if assumptions of the t-test are violated? **Best do not use barplots at all: Much ink, low information content.** Boxplots are an alternative, they even are non-parametric!