

Reporting statistical result: t-test example

Nathan Brouwer brouwern@gmail.com @lobrowR

October 18, 2017

Introduction

This vignette shows an example of reporting the results from a 2-sample t-test using data on the impact of invasive trout on salmon survival.

Outline of tasks

- Load the data into R
- Create a boxplot of the raw data
- State the relevant statistical null (H_0) and alternative (H_a) hypotheses
- Carry out an appropriate t-test
- Report the appropriate results in a full sentence as it would appear in a report or scientific paper

```
#load the data
## If I were loading the data by hand it would look something like this
## don't forget the commas!
salmon <- data.frame(survival = c(0.83,0.87,0.82,
                                0.84,0.81,0.84,
                                0.72, 0.84,0.75,
                                0.79,0.89,0.87),
                    brook.trout.PRES.ABS =
                      c("present","present","present",
                        "absent","absent","absent",
                        "absent","absent","absent"))

#Look at data
##I am using the salmon vs. brook trout data as an example
head(salmon)
```

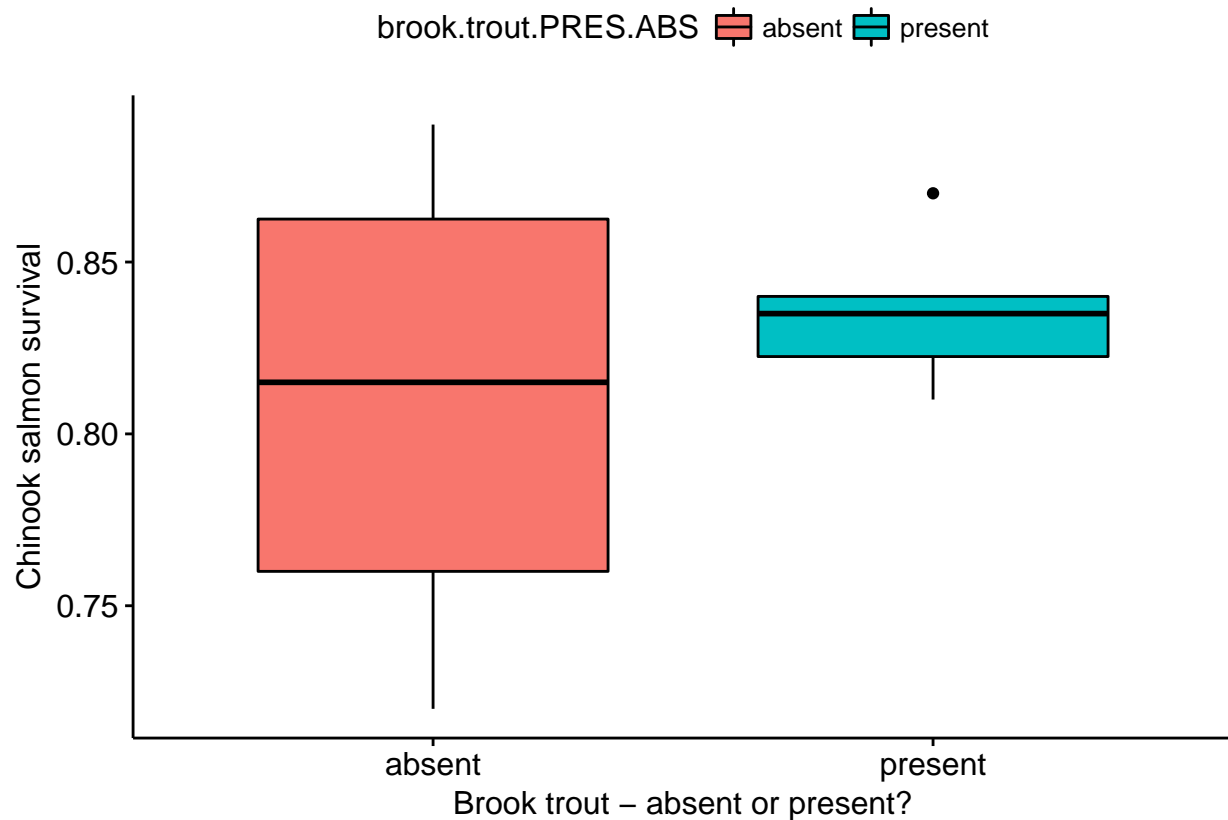
```
##   survival brook.trout.PRES.ABS
## 1    0.83          present
## 2    0.87          present
## 3    0.82          present
## 4    0.84          present
## 5    0.81          present
## 6    0.84          present
```

```
#Plot raw data
library(ggplot2)
library(ggpubr)
```

```
## Loading required package: magrittr
```

```
ggboxplot(data = salmon,
          y = "survival",
          x = "brook.trout.PRES.ABS",
          fill = "brook.trout.PRES.ABS",
```

```
xlab = "Brook trout - absent or present?",
ylab = "Chinook salmon survival")
```



```
#Do t-test
t.test(survival ~ brook.trout.PRES.ABS,
       data = salmon)
```

```
##
## Welch Two Sample t-test
##
## data: survival by brook.trout.PRES.ABS
## t = -0.86344, df = 5.9267, p-value = 0.4215
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.09606084 0.04606084
## sample estimates:
## mean in group absent mean in group present
## 0.810 0.835
```

State the hypotheses

Ho: The survival rates of Chinook salmon are the same whether brook trout are present or absent Ha: The persence of brook trout changes survival rates of salmon.

Report the results

For the real data, the results could be reported like this: “There was no evidence that the mean survival of salmon when brook trout are present (mean = 0.81) is different than when brook trout are absent (mean = 0.84; 2-sample t-test: $p = 0.44$, $t = 0.82$, $n = 12$ streams, $df = 6$).”

Normally I would also report the standard errors (SE) around the means, but for this exercise we will ignore it.

Alternative results

What if the results really looked like this?

```
##
##  Welch Two Sample t-test
##
## data:  fake.surv by salmon$brook.trout.PRES.ABS
## t = 3.5938, df = 5.725, p-value = 0.01241
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.03173718 0.17227394
## sample estimates:
##  mean in group absent mean in group present
##                0.8142305                0.7122249
```

The results could be reported like this: “Survival of chinook salmon in streams where brook were present (mean = 0.71) was significantly lower than when brook trout were absent (mean = 0.81) with a mean difference of 0.10 (95% CI: 0.03-0.17; 2-sample t-test $p = 0.012$, $t = 3.6$, $n = 12$ stream, $df = 5.73$)”