A Vignette to Test Two Means

Christina Knudson, PhD

March 20, 2018

Overview

This vignette will demonstrate the process and functions used to answer the following research question: Is the mean weight percent nitrogen greater during laminated intervals?

In this vignette, we first cover the hypotheses, the test statistic used, and the intuition behind the test. Next, we demonstrate the `Bering` functions necessary to conduct the test. Then, we demonstrate how to use the test output to visualize our results. Finally, we discuss the computational cost of this procedure.

Hypotheses, test statistics, and intuition

The null hypothesis is that the mean weight percent nitrogen does not differ between laminated and non-laminated intervals. The alternative hypothesis is that the mean weight percent nitrogen during laminated intervals is greater than the mean weight percent nitrogen during non-laminated intervals.

Our test statistic is the mean during laminated intervals minus the mean during non-laminated intervals. Using Monte Carlo, we simulate 10000 data-sets and test statistics under the null hypothesis. To simulate each data-set, we randomly assign each measurement to one of two states (laminated or non-laminated) using a Markov chain with empirical transition probabilities calculated using the original data set.

If the test statistic is relatively large compared to the distribution of simulated test statistics, we will reject the null hypothesis and conclude that the mean weight percent nitrogen differs between laminated and non-laminated intervals. In particular, our p-value is the proportion of test statistics greater than or equal to the original data's test statistic.

This vignette uses `mockdata`, which has real weight percent nitrogen measurements, real ages, and simulated laminations. We will update the `Bering` package and this vignette to include the true data if/when our manuscript is published.

Exploring the data

Begin by invoking the `Bering` package and `mockdata` data.

```r
library(Bering)
data(mockdata)
```

At this point, I highly recommend conducting exploratory data analysis to become acquainted with the data. In particular, I recommend using the `summary` command (in base R) to understand the data.

```r
summary(mockdata)
```

Our data frame `mockdata` has three variables: `age`, `wtN`, and `lam`. Information on these variables can be found by typing the following into the `R` console:

```r
?mockdata
```

After completing exploratory data analysis, we ensure that the measurements are in order by age. The code below demonstrates one way to perform this check. TRUE indicates the measurements are indeed in order.

```r
sum(mockdata$age == sort(mockdata$age)) == length(mockdata$age)
```

```
## [1] TRUE
```

Conducting the test

To prepare for the hypothesis test, we calculate two empirical probabilities:

- the empirical probability of a laminated measurement following a nonlaminated measurement
- the empirical probability of a laminated measurement following another laminated measurement.

The `peas` function from `Bering` calculates these probabilities using a single input: the vector indicating which measurements are laminated.

```r
mypea <- peas(mockdata$lam)
(pswitchL <- mypea$pswitchL)
```

```
## [1] 0.3047493
```

```r
(pstayL <- mypea$pstayL)
```

```
## [1] 0.3113772
```

Next, we calculate the test statistic (the mean weight percent nitrogen during laminated intervals minus the mean weight percent nitrogen during non-laminated intervals) of the original data. The code below demonstrates one way to calculate this statistic.

```r
mean0 <- mean(mockdata[mockdata$lam == 0, 2]) #mean  for nonlam
mean1 <- mean(mockdata[mockdata$lam == 1, 2]) #mean  for lam
(teststat <- mean1 - mean0)
```

```
## [1] -0.001508251
```

Finally, we calculate the p-value using the `righttailtest` function from `Bering`. (We use `righttailtest` rather than `lefttailtest` or `twotailtest` because our alternative hypothesis is that the mean weight percent nitrogen during laminated intervals is greater than the mean weight percent nitrogen during non-laminated intervals.) The first argument for `righttailtest` is the vector of weight percent nitrogen measurements, the second is the test statistic from the original data set, the third is the empirical probability of a laminated measurement following a nonlaminated measurement, the fourth is the empirical probability of a laminated measurement following another laminated measurement, and the fifth is the Monte Carlo sample size. The Monte Carlo sample size represents the number of test statistics that will be simulated under the null hypothesis (no difference between the means).

```r
out <- righttailtest(mockdata$wtN, teststat, pswitchL, pstayL, m = 10^4)
out$pvalue
```

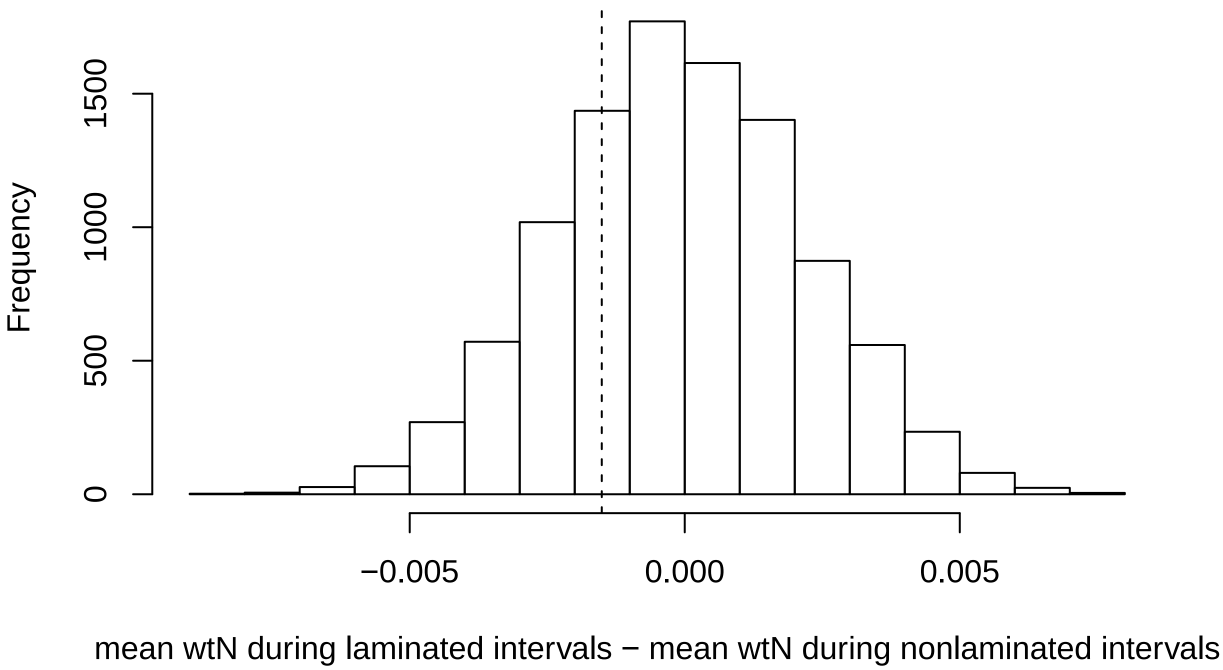
```
## [1] 0.7374263
```

Because the p-value from this test (using mock data) is quite large, we do not reject our null hypothesis. In other words, the mean weight percent nitrogen is not significantly higher during laminated intervals. Due to the nature of the mock data, this is not surprising.

Visualizing the results

To better understand our results, we can visualize the simulated test statistics and compare them to the original data's test statistic. We first create a histogram using the simulated test statistics (returned by our test) and then we draw a vertical dashed line at our original data's mean.

```r
hist(out$distrib, main = NULL,
     xlab="mean wtN during laminated intervals - mean wtN during nonlaminated intervals")
abline(v = teststat, lty = 2)
```



This plot makes it clear that our original data's test statistic is not at all unusual compared to the test statistics simulated under the hypothesis that sea levels and laminations are unrelated. Due to the nature of the mock data, this is not surprising.

Computational cost

Compiling this vignette on a 'normal' computer (Windows 10 with 8 GB of RAM) took about 30 seconds.

Alternative hypotheses

This vignette demonstrated the analysis for testing the alternative hypothesis that the mean weight percent nitrogen during laminated intervals is *greater* than the mean weight percent nitrogen during non-laminated intervals. Two functions similar to the demonstrated `righttailtest` are `lefttailtest` and `twotailtest`. All three functions have the same arguments and same argument order. The difference between the functions is the alternative hypothesis (and corresponding p-value calculation).

If we instead wish to test the alternative hypothesis that the mean weight percent nitrogen during laminated intervals is *smaller* than the mean weight percent nitrogen during non-laminated intervals, we should use `lefttailtest` rather than `righttailtest`. The p-value is calculated as the proportion of test statistics no larger than our original data's test statistic.

If we wish to test the alternative hypothesis that the mean weight percent nitrogen during laminated intervals *differs* from the mean weight percent nitrogen during non-laminated intervals, we should use `twotailtest`. The p-value is calculated as the proportion of test statistics as extreme as or more extreme than our original data's test statistic.