HBR Maples Miniproject

Siena Blier

2025-09-11

Introduction

I am citing a paper (Juice et al. 2006). And another one Patykowski et al. (2018).

Methods

I got the data from (Juice and Fahey 2019).

```
clean.data <- read.csv("../02_outdata/hbr-maples_clean-dataset.csv")
```

We checked that the assumptions of a t-test were met.

1. Stem dry mass and leaf dry mass are both **continuous** variables.

```
head(clean.data$stem_dry_mass)
```

```
## [1] 0.0300 0.0338 0.0248 0.0194 0.0180 0.0246
```

```
head(clean.data$leaf_dry_mass)
```

```
## [1] 0.0453 0.0476 0.0423 0.0397 0.0204 0.0317
```

2. We assume that Juice and Fahey (2019) carried out **random sampling**.
3. The assumption of **homoscedasticity** is not respected, so we use a t-test with **unequal variance**

```
# Check equal variance for stem dry mass
aggregate(stem_dry_mass~watershed, data=clean.data, var)
```

```
##   watershed stem_dry_mass
## 1 Reference   0.000593267
## 2        W1   0.001062190
```

```
leveneTest(clean.data$stem_dry_mass, clean.data$watershed)
```

```
## Warning in leveneTest.default(clean.data$stem_dry_mass, clean.data$watershed):
## clean.data$watershed coerced to factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value   Pr(>F)
## group   1  9.2034 0.002593 **
##       357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Check equal variance for leaf dry mass
aggregate(leaf_dry_mass~watershed, data=clean.data, var)
```

```
##   watershed leaf_dry_mass
## 1 Reference   0.001092108
## 2        W1   0.001114441
```

```r
leveneTest(clean.data$leaf_dry_mass, clean.data$watershed)
```

```
## Warning in leveneTest.default(clean.data$leaf_dry_mass, clean.data$watershed):
## clean.data$watershed coerced to factor.
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value   Pr(>F)
## group   1  9.6759 0.002017 **
##       357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

4. The data are not **normally distributed**, but because we're dealing with a fairly large sample size (n = 359), we will still conduct the t-test.

```r
# Check normality of stem dry mass for reference samples

clean.data.ref <- clean.data %>%
  filter(watershed == "Reference")

shapiro.test(clean.data.ref$stem_dry_mass)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  clean.data.ref$stem_dry_mass
## W = 0.81542, p-value = 8.84e-14
```

```r
shapiro.test(clean.data.ref$leaf_dry_mass)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  clean.data.ref$leaf_dry_mass
## W = 0.59272, p-value < 2.2e-16
```

```r
clean.data.w1 <- clean.data %>%
  filter(watershed == "W1")

shapiro.test(clean.data.w1$stem_dry_mass)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  clean.data.w1$stem_dry_mass
## W = 0.84849, p-value = 2.155e-12
```

```r
shapiro.test(clean.data.w1$leaf_dry_mass)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  clean.data.w1$leaf_dry_mass
## W = 0.94138, p-value = 1e-06
```

```r
rm(clean.data.ref, clean.data.w1)
```

Results

A first look at the distribution of the data suggested there may be differences in stem and leaf dry mass between reference and calcium-treated seedlings.

```r
# Histogram of stem dry mass
ggplot(clean.data, aes(x = stem_dry_mass, fill = watershed)) +
  geom_histogram(alpha = 0.6, position = "identity", bins = 30) +
  labs(x = "Value", y = "Frequency") +
  theme_classic()

# Histogram of leaf dry mass
ggplot(clean.data, aes(x = leaf_dry_mass, fill = watershed)) +
  geom_histogram(alpha = 0.6, position = "identity", bins = 30) +
  labs(x = "Value", y = "Frequency") +
  theme_classic()
```

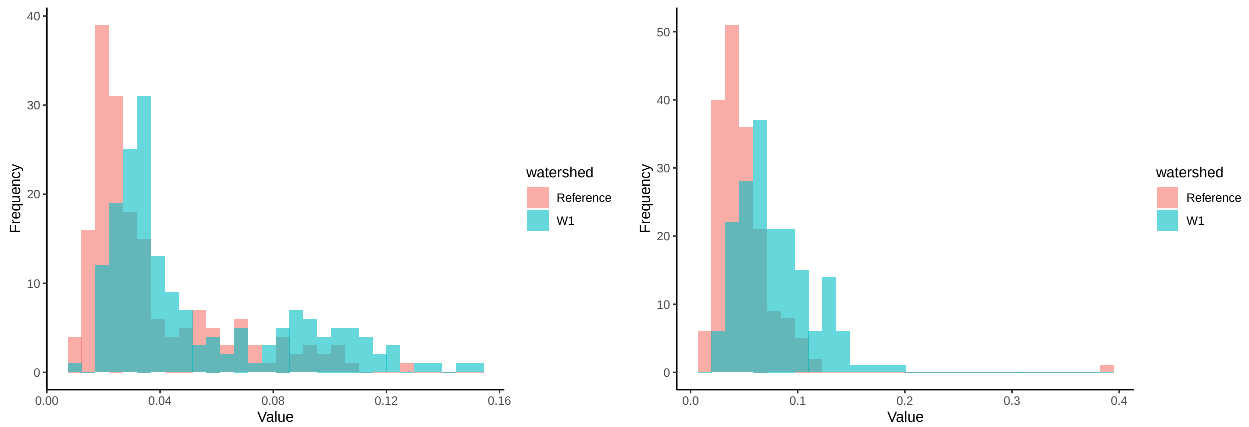

Figure 1: Distribution of stem dry mass (left) and leaf dry mass (right) of seedlings in reference (red) and calcium treated (blue) samples (n = 359).

We found that stem and leaf dry mass differed significantly between calcium treated and reference samples.

```r
# Test for differences in stem dry mass
t.test(stem_dry_mass ~ watershed, data = clean.data, var.equal = FALSE) # Use Welch's Two Sample t-test
```

```
##
##  Welch Two Sample t-test
##
## data:  stem_dry_mass by watershed
## t = -5.4508, df = 331.41, p-value = 9.81e-08
## alternative hypothesis: true difference in means between group Reference and group W1 is not equal to
## 95 percent confidence interval:
##  -0.02251850 -0.01057521
## sample estimates:
## mean in group Reference        mean in group W1
##              0.03676592              0.05331278
```

```r
# Test for differenecs in leaf dry mass
t.test(leaf_dry_mass ~ watershed, data = clean.data, var.equal = FALSE) # Use Welch's Two Sample t-test
```

```
##
##  Welch Two Sample t-test
##
## data:  leaf_dry_mass by watershed
## t = -8.0462, df = 356.99, p-value = 1.278e-14
## alternative hypothesis: true difference in means between group Reference and group W1 is not equal to
## 95 percent confidence interval:
##  -0.03510551 -0.02131523
## sample estimates:
## mean in group Reference        mean in group W1
##              0.04953296              0.07774333
```

Discussion

A citation for the discussion Kobe et al. (2002).

References

Juice, S. M., T. J. Fahey, T. G. Siccama, C. T. Driscoll, E. G. Denny, C. Eagar, N. L. Cleavitt, R. Minocha, and A. D. Richardson. 2006. Response of Sugar Maple to Calcium Addition to Northern Hardwood Forest. Ecology 87:1267–1280.

Juice, S., and T. Fahey. 2019. Health and mycorrhizal colonization response of sugar maple (Acer saccharum) seedlings to calcium addition in Watershed 1 at the Hubbard Brook Experimental Forest. Environmental Data Initiative.

Kobe, R. K., G. E. Likens, and C. Eagar. 2002. Tree seedling growth and mortality responses to manipulations of calcium and aluminum in a northern hardwood forest. Canadian Journal of Forest Research 32:954–966.

Patykowski, J., J. Kołodziejek, and M. Wala. 2018. Biochemical and growth responses of silver maple (Acer saccharinum L.) to sodium chloride and calcium chloride. PeerJ 6:e5958.