Adjusting Standard ANOVA Methods to Account for Heterogeneous Variances With an Application to Turfgrass Management

Michael Dumelle

November 30, 2020

Section 1

Introduction

Background

- Slides available on my GitHub here
- PhD in Statistics from Oregon State University (2020)
- Research statistician at the Environmental Protection Agency

Background

- OSU Statistics Consulting Practicum
  - Encourage you to sign up!
  - Long format vs drop-in
  - Faculty are encouraged too - separate process
- Worked on several turfgrass projects with Alec Kowalewski and Clint Mattox
- Use Analysis of Variance (ANOVA) to study designed experiments
  - Are there *statistically significant differences* among treatment effects?
- One common problem: unequal variance / standard deviation within treatment groups
  - How can we use ANOVA to best understand our data when there is unequal variance?

Experiment Roadmap

1. Formulate a hypothesis
2. Choose an experimental design
3. Choose an analysis method
4. **Randomize** treatments
5. Collect data
6. Analyze data using ANOVA
   - $Y_i = \mu + \alpha_i + \epsilon_i$ (focus on one-way ANOVA)
   - Estimate treatment effects from the data
   - Do these estimates ($\hat{\alpha}$) suggest *statistically significant differences* among the true treatment effects (α)?
7. Report results

Section 2

Why ANOVA?

Properties

ANOVA has several attractive propeties:

1. Estimates of treatment effects equal the true treatment effects *on average*
   - But we only get to run the experiment once!
2. Treatment effect confidence intervals are as small as possible
3. Hypothesis tests have well known forms

But 2 and 3 rely on specific assumptions on the errors, ϵ

Assumptions on ϵ

1. Independence
   - Unit A does not depend on unit B
   - Dice roll, coin flip
2. Normality
3. Constant Variance
   - Variance homogeneity

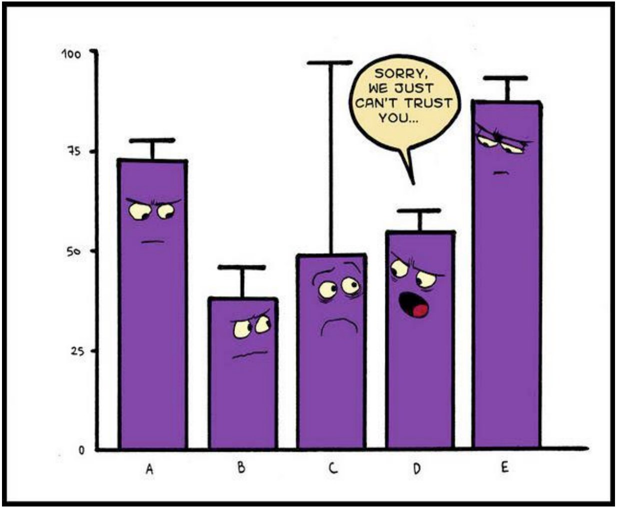
- This presentation focuses on 3.

Assumptions on ϵ

When the constant variance assumption is violated, the ANOVA

- Estimates of treatment effects still equal the true treatment effects *on average*
  - Great!
- Treatment effect confidence intervals are too lage
  - Inefficient use of resources
- Hypothesis tests don't have well known forms
  - Incorrect p-values $\rightarrow$ misleading conclusions, poor policy decisions

Assumptions



How Do I Know?

- Graphics! If it looks off, it probably is
- Ratio of largest and smallest variances
    - Suggestions of cutoff range from 1.5 to 9
- Statistical tests for constant variance
    - Levene's, Brown-Forsythe, several others
    - Come with their own assumptions

What Now?

So I know my data does not have constant variance, what now?

- Could transform the response, Y, so that the transformed Y satisfy standard assumptions
  - Can be very useful
  - Generally require a specific mean / variance relationship, $log_e(Y)$ often used

Poses pratical considerations:

- Challenging to find an appropriate transformation
- Difficult to interpret on original scale (usually of interest)
  - Significant difference between treatments on transformed scale DOES NOT imply the same on the original scale
- What else?

Section 3

GV-ANOVA

What is GV-ANOVA?

- Can use Generalized Variance ANOVA (GV-ANOVA) to directly model variances within groups
  - Separate variance for each treatment level
  - Does not require a mean / variance relationship
  - No transformation requirement
  - Requires the estimation of more variance parameters than when using a transformation
- Goal of this talk is to expose you to another possible way to handle the non constant variance problem
- Important to be aware of both approaches

Section 4

Application

Example

```r
anova_mod <- gls(response ~ trt, data = data)
anova(anova_mod)
```

```
## Denom. DF: 28
##             numDF  F-value p-value
## (Intercept)     1 10358.47  <.0001
## trt             3    30.27  <.0001
```

```r
emmeans(anova_mod, list(pairwise ~ trt), adjust = "bonferroni")
```

```
## $`emmeans of trt`
##  trt emmean   SE df lower.CL upper.CL
##  A     47.0 1.07 28     44.8     49.2
##  B     52.3 1.07 28     50.1     54.5
##  C     58.1 1.07 28     55.9     60.2
##  D     60.0 1.07 28     57.8     62.2
##
## Degrees-of-freedom method: df.error
## Confidence level used: 0.95
##
## $`pairwise differences of trt`
##  contrast estimate   SE df t.ratio p.value
##  A - B       -5.24 1.51 28 -3.473  0.0102
##  A - C      -11.03 1.51 28 -7.301  <.0001
##  A - D      -12.97 1.51 28 -8.589  <.0001
##  B - C       -5.78 1.51 28 -3.829  0.0040
##  B - D       -7.73 1.51 28 -5.116  0.0001
##  C - D       -1.94 1.51 28 -1.287  1.0000
##
## Degrees-of-freedom method: df.error
## P value adjustment: bonferroni method for 6 tests
```

Example

```
gvanova_mod <- gls(response ~ trt, weights = varIdent(form = ~ 1|trt), data = data)
anova(gvanova_mod)
```

```
## Denom. DF: 28
##             numDF   F-value p-value
## (Intercept)     1 230058.47  <.0001
## trt             3     67.27  <.0001
```

```
emmeans(gvanova_mod, list(pairwise ~ trt), adjust = "bonferroni")
```

```
## $`emmeans of trt`
##  trt emmean     SE   df lower.CL upper.CL
##  A    47.03 0.9674 6.97    44.74    49.32
##  B    52.27 1.8105 7.01    47.99    56.55
##  C    58.05 0.5745 7.00    56.69    59.41
##  D    60.00 0.1288 7.00    59.69    60.30
##
## Degrees-of-freedom method: satterthwaite
## Confidence level used: 0.95
##
## $`pairwise differences of trt`
##  contrast estimate    SE    df t.ratio p.value
##  A - B       -5.24 2.053 10.72  -2.555 0.1636
##  A - C      -11.03 1.125 11.37  -9.799 <.0001
##  A - D      -12.97 0.976  7.21 -13.289 <.0001
##  B - C       -5.78 1.900  8.40  -3.044 0.0905
##  B - D       -7.73 1.815  7.08  -4.256 0.0220
##  C - D       -1.94 0.589  7.70  -3.302 0.0686
##
## Degrees-of-freedom method: satterthwaite
## P value adjustment: bonferroni method for 6 tests
```

Section 5

Conclusions