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
- I will interweave R code to illustrate ideas

```
# this is a comment
this_is_some_code <- this_is_a_function(this_is_an_argument)
print(this_is_some_code)
#> [1] "this is output"
```

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

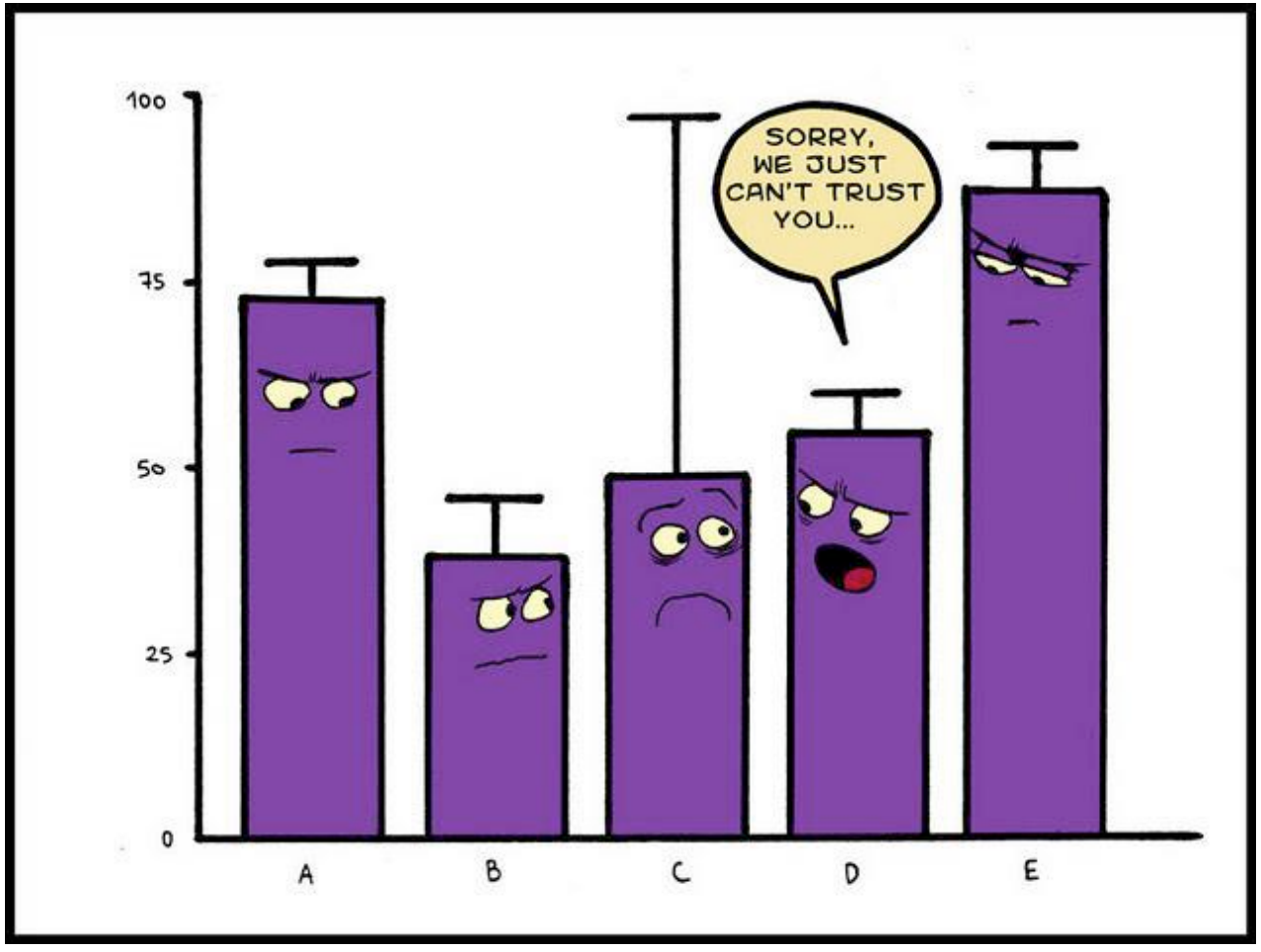
- This presentation focuses on 3.

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

Percent Green Cover

Create Data

- Use a simulation to compare ANOVA and GVANOVA
  - So helpful because we know the truth!
  - Study several scenarios without having to design an experiment, collect data, etc.

Table 1: Treatment Means, Standard Deviations (StDev), and Replicates

Treatment	Mean	StDev	Replicates
A	50	5.0	8
B	50	2.0	8
C	58	1.0	8
D	60	0.5	8

Create Data

- Study pairwise differences between treatments at $\alpha = 0.05$
  - For each comparison, compute a test statistic and a p-value
  - Reject a true null hypothesis (Type I error, false positive)
  - Do not reject a true alternative hypothesis (Type II error, false negative)
  - Family-wise error rate is $\alpha \rightarrow$ multiple comparison adjustment (Bonferroni)
  - Degree of freedom adjustments (Sattherwaite)

Create Data

```
set.seed(1130)
data <- create_data(treatments = c("A", "B", "C", "D"),
                    means = c(50, 50, 58, 60),
                    stdevs = c(5, 2, 1, 0.5),
                    replicates = c(8, 8, 8, 8))
head(data, n = 9)
#>   treatments response
#> 1          A 43.98231
#> 2          A 54.94049
#> 3          A 45.64911
#> 4          A 50.33370
#> 5          A 45.03723
#> 6          A 54.45938
#> 7          A 47.62064
#> 8          A 40.34577
#> 9          B 50.07621
```

Create Data

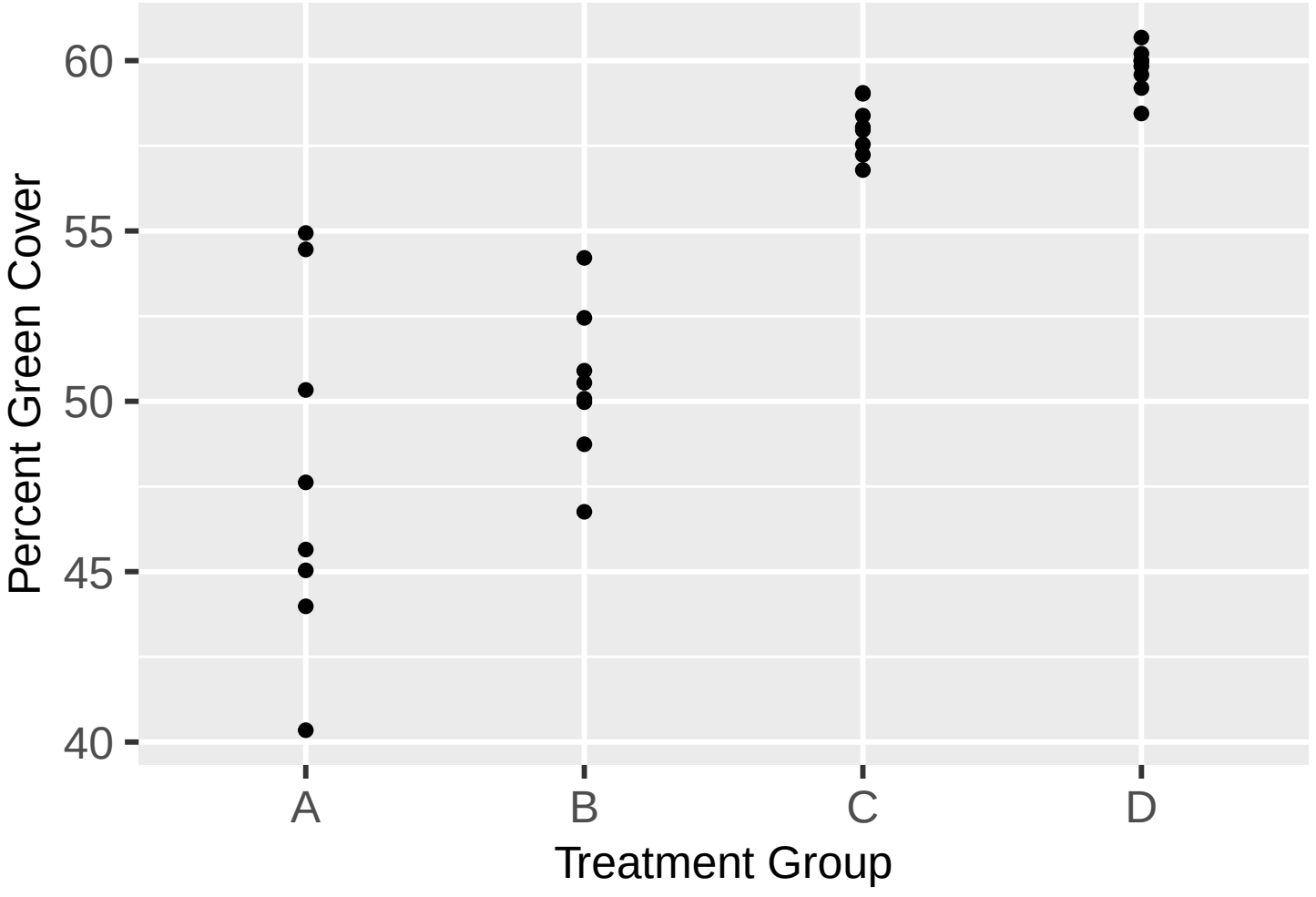
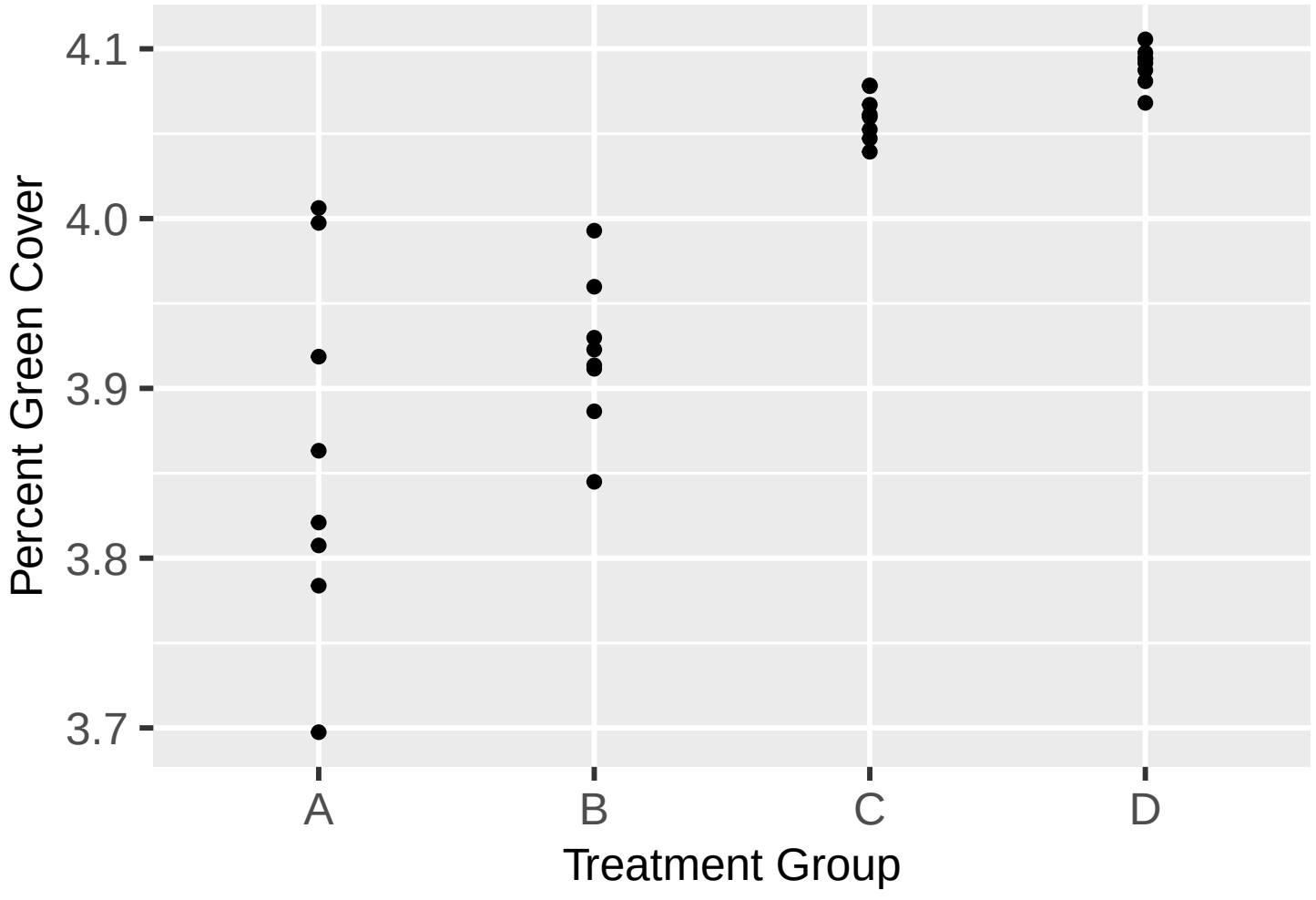


Figure 1: Caption

Create Data



Data Summaries

```
group_summaries <- data %>%
  group_by(treatments) %>%
  summarize(grp_mean = mean(response),
            grp_stdev = sd(response))
```

Table 2: Treatment Group Summaries

Treatments	Group Mean	Group StDev
A	47.796	5.130
B	50.456	2.245
C	58.005	0.809
D	59.742	0.678

Are the Unequal Variances a Problem?

```
# ratio of largest and smallest stdev
max_stdev <- max(group_summaries$grp_stdev)
min_stdev <-  min(group_summaries$grp_stdev)
stdev_ratio <- max_stdev / min_stdev
stdev_ratio
#> [1] 7.563561
stdev_ratio > 3
#> [1] TRUE

# levene's test
leveneTest(response ~ treatments, data = data)
#> Levene's Test for Homogeneity of Variance (center = median)
#>       Df F value    Pr(>F)
#> group  3  7.3332 0.0008926 ***
#>       28
#> ---
#> Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Fit Models

```r
anova_mod <- gls(response ~ treatments, data = data)
anova_trtmeans <- emmeans(anova_mod, "treatments")
pairs(anova_trtmeans, adjust = "bonferroni")
#>  contrast estimate   SE df t.ratio p.value
#>  A - B       -2.66 1.42 28 -1.868  0.4340
#>  A - C      -10.21 1.42 28 -7.166  <.0001
#>  A - D      -11.95 1.42 28 -8.386  <.0001
#>  B - C       -7.55 1.42 28 -5.299  0.0001
#>  B - D       -9.29 1.42 28 -6.519  <.0001
#>  C - D       -1.74 1.42 28 -1.220  1.0000
#>
#> Degrees-of-freedom method: df.error
#> P value adjustment: bonferroni method for 6 tests
```

```
# SAS Code
proc mixed data=data;
  class treatments;
  model response = treatments;
  lsmeans treatments / diff adjust=BON;
run;
```

Example

```
gvanova_mod <- gls(response ~ treatments, weights = varIdent(form = ~ 1|treatments), data = data)
gvanova_trtmeans <- emmeans(gvanova_mod, "treatments")
pairs(gvanova_trtmeans, adjust = "bonferroni")
#>  contrast estimate    SE    df t.ratio p.value
#>  A - B       -2.66 1.980  9.59  -1.344 1.0000
#>  A - C      -10.21 1.836  7.35  -5.560 0.0043
#>  A - D      -11.95 1.829  7.25  -6.530 0.0017
#>  B - C       -7.55 0.844  8.78  -8.947 0.0001
#>  B - D       -9.29 0.829  8.26 -11.199 <.0001
#>  C - D       -1.74 0.373 13.59  -4.655 0.0024
#>
#> Degrees-of-freedom method: satterthwaite
#> P value adjustment: bonferroni method for 6 tests
```

```
# SAS Code
proc mixed data=data;
  class treatments;
  model response = treatments / ddfm=SAT; # this is different
  repeated / group = treatments; # this is different
  lsmeans treatments / diff adjust=BON;
run;
```

Section 5

Conclusions