

## Question 2.4

- a. The `anova()` results for the one-way ANOVA are shown in Table 1.

Table 1. Analysis of variance table for the egg length by bird species.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
species	2	0.410	0.20496	0.1992	0.8202
Residuals	41	42.182	1.02883		
Total	43	42.592			

- $df_{Residuals}$ , or  $df_{Within}$ , is one less than the number of groups (host bird species;  $I = 3$ ) in the analysis.
- $df_{species}$ , or  $df_{Among}$ , is the total number of cuckoo eggs ( $n = 44$ ) minus the total number of groups ( $I = 3$ ) in the analysis.
- $MS_{Residuals}$ , or  $MS_{Within}$ , is the variability in cuckoo egg length around the mean cuckoo egg length of each host bird species group or the variability unexplained by using a separate mean for each group (i.e., the full model).
- $MS_{species}$ , or  $MS_{Among}$ , is the variability that is explained by using separate means (full model) that was not explained when only one mean (simple model) was used to represent all three groups.
- The F test statistic is the ratio of the variability explained by using separate means to the variability not explained by using a single mean to represent all three groups. Alternatively, the F test statistic is the ratio of the variability among means to the variability among individuals within groups.

- b. The `summary()` results of the one-way ANOVA are shown in Table 2.

Table 2. Coefficients from the one-way ANOVA fit of egg length by bird species.

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	23.09000	0.26189	88.165	<2e-16
speciesSparrow	0.03143	0.37693	0.083	0.934
speciesWagtail	-0.18667	0.37037	-0.504	0.617
---				
Residual standard error: 1.014 on 41 degrees of freedom				
Multiple R-squared: 0.009624, Adjusted R-squared: -0.03869				
F-statistic: 0.1992 on 2 and 41 DF, p-value: 0.8202				

- The “Intercept” coefficient is the sample mean cuckoo egg length in the pipet (reference group) nests. Thus, the mean length of cuckoo eggs in the pipet nests is 23.09 mm.
- The “speciesSparrow” estimate is the difference in the sample mean cuckoo egg length in the sparrow and pipet nests. In this case, the mean cuckoo egg length in the sparrow nests is 0.03 mm longer than the mean cuckoo egg length in the pipet nests.
- The “speciesWagtail” estimate is the difference in the sample mean cuckoo egg length in the wagtail and pipet nests. In this case, the mean cuckoo egg length in the wagtail nests is 0.19 mm *shorter* than the mean cuckoo egg length in the pipet nests.
- The “intercept” p-value is used to determine if the mean cuckoo egg length in the pipet (reference group) nests differs from zero. In this case, a  $p < 0.00005$  suggests that the mean length of cuckoo eggs in pipet nests is different from zero.
- The “speciesSparrow” p-value is used to determine if the difference in the mean cuckoo egg length in the sparrow and pipet nests is different from zero. In this case, a  $p = 0.9340$  suggests that this

difference is not different than zero, suggesting that there is no difference in the mean cuckoo egg length between the sparrow and pipet nests.

- The “speciesWagtail” p-value is used to determine if the difference in the mean cuckoo egg length in the wagtail and pipet nests is different from zero. In this case,  $p = 0.6170$  suggests that this difference is not different than zero suggesting that there is no difference in the mean cuckoo egg length between the wagtail and pipet nests.

c. No, the overall p-value ( $p = 0.8202$ ) (Table 1) is greater than 0.05. There is no statistical difference in mean cuckoo length among the three species of host birds.

d. Not needed because no difference was identified.

## Notes from Professor

- When asked to assess differences among “all” group means then you must use the overall p-value found in the ANOVA table. The p-values found in the coefficients table do not answer this question because these p-values do not tell all pairs.
- Don’t say “the groups differ” or the “the mean differs among groups.” Be explicit ... say “the mean length of cuckoo eggs differs among the three host species nests.”
- Don’t say “first group” or “second group” when those groups have specific names. For example, say “the intercept represents the mean length of cuckoo eggs in the tree pipet group.”

R commands

```
> cuc <- read.csv("https://raw.githubusercontent.com/droglenc/NCDData/master/Cuckoos.csv")
> lm1 <- lm(length~species,data=cuc)
> anova(lm1)
> summary(lm1)
> confint(lm1)
```