Module 8 Assignment

Alex Schlesener

2023-10-15

Contents

Instructions . . . 1
- Import Data . . . 1

Examine the Data . . . 2
- Conduct One-Way Analysis of Variance (R) . . . 3

A) Analyze Contrasts . . . 4
- Define and Test Contrasts in R . . . 5
- Summary of Results . . . 5

B) Calculate by Hand . . . 6
- RQ1 by Hand . . . 6

C) Pairwise Comparison . . . 6
- Tukey's HSD . . . 7
- Summary of Results . . . 7

Instructions

You have data from participants who were on one of four types of diets (A, B, C, D). After three months, the amount of weight loss was recorded in lb. In Module 7, you already completed the one-way ANOVA using hand calculations and you used SPSS and R. You conducted this analysis at $\alpha = .05$. Recall that you rejected the null hypothesis. In Module 7, you also wrote a brief summary including the results of your statistical test and you reported an index of effect size.

Import Data

```
Diet <- read.csv("Data/weightloss.csv")
str(Diet)
```

```
## 'data.frame':    23 obs. of  2 variables:
##  $ weightloss: int  12 10 13 9 13 9 14 17 22 14 ...
##  $ diet      : int  1 1 1 1 1 1 2 2 2 2 ...
```

After utilizing the ‘str()’, we can see that the data frame, “Diet”, has 23 rows and 2 columns (variables). The two variables are Weightloss and Diet. Both variables are vectors of integers; however, Diet should be treated as a ‘factor’. We will convert ‘Diet’ to a factor and replace the integer vector to its appropriate code: “1” = “A”, “2” = “B”, “3” = “C”, and “4” = “D”.

```
Diet$diet <- factor(Diet$diet, labels = c("A", "B", "C", "D"))
str(Diet)
```

```
## 'data.frame':    23 obs. of  2 variables:
##  $ weightloss: int  12 10 13 9 13 9 14 17 22 14 ...
##  $ diet      : Factor w/ 4 levels "A","B","C","D": 1 1 1 1 1 1 2 2 2 2 ...
```

```
levels(Diet$diet)
```

```
## [1] "A" "B" "C" "D"
```

```
summary(Diet)
```

```
##    weightloss    diet
##  Min.   : 9.00   A:6
##  1st Qu.:12.50   B:5
##  Median :14.00   C:7
##  Mean   :14.78   D:5
##  3rd Qu.:17.50
##  Max.   :22.00
```

Examine the Data

```
# Calculate the Descriptive Statistics by Condition
aggregate(weightloss ~ diet, data = Diet, mean)
```

```
##   diet weightloss
## 1    A       11.0
## 2    B       16.4
## 3    C       18.0
## 4    D       13.2
```

```
aggregate(weightloss ~ diet, data = Diet, sd)
```

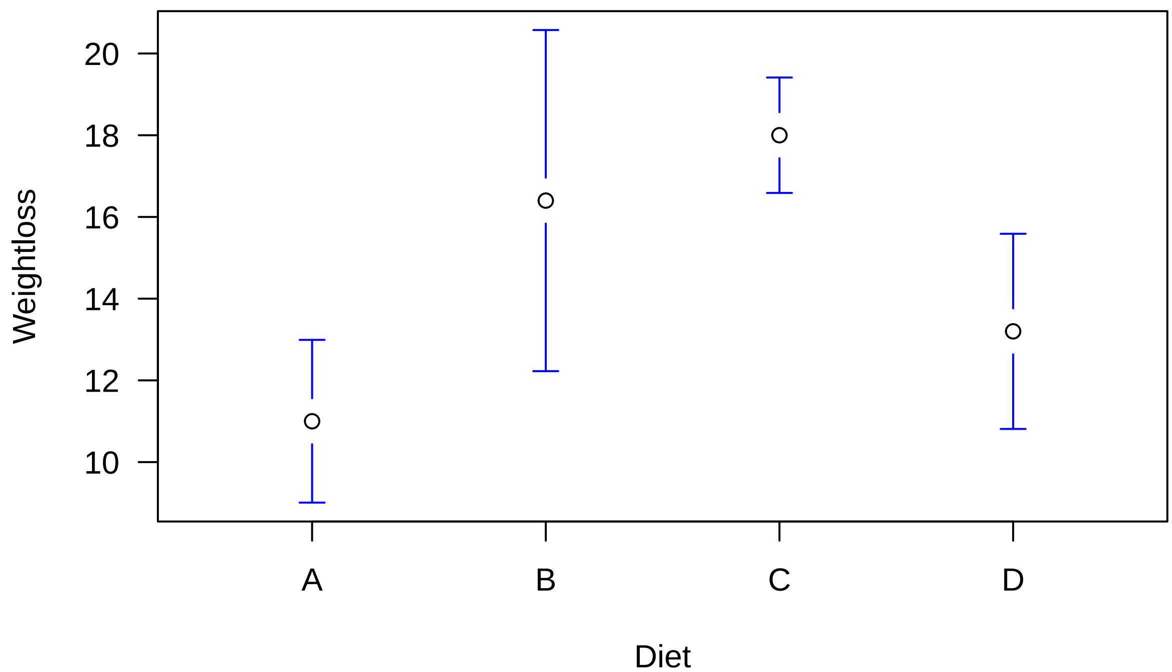
```
##   diet weightloss
## 1    A   1.897367
## 2    B   3.361547
## 3    C   1.527525
## 4    D   1.923538
```

```
# Plot each mean with the confidence interval using the gplots library.
library(gplots)
```

```
##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##     lowess
```

```
plotmeans(weightloss ~ diet, Diet, las = 1,
          ylab = "Weightloss",
          xlab = "Diet",
          n.label = F,
          connect = F)
```



Conduct One-Way Analysis of Variance (R)

```
oneway.diet <- lm(weightloss ~ diet, data=Diet)
summary(oneway.diet)
```

```
##
## Call:
## lm(formula = weightloss ~ diet, data = Diet)
##
## Residuals:
```

```
##    Min     1Q Median     3Q    Max
##   -2.4   -1.7    0.0    0.9    5.6
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.0000     0.8983  12.245 1.84e-10 ***
## dietB         5.4000     1.3325   4.053 0.000679 ***
## dietC         7.0000     1.2242   5.718 1.64e-05 ***
## dietD         2.2000     1.3325   1.651 0.115153
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.2 on 19 degrees of freedom
## Multiple R-squared:  0.6666, Adjusted R-squared:  0.6139
## F-statistic: 12.66 on 3 and 19 DF,  p-value: 8.881e-05
```

```
anova(oneway.diet)
```

```
## Analysis of Variance Table
##
## Response: weightloss
##           Df Sum Sq Mean Sq F value    Pr(>F)
## diet       3 183.91  61.304  12.661 8.881e-05 ***
## Residuals 19  92.00   4.842
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Check Homogeneity of Variance
library(car)
```

```
## Loading required package: carData
```

```
leveneTest(weightloss ~ diet, data=Diet, center=mean)
```

```
## Levene's Test for Homogeneity of Variance (center = mean)
##       Df F value Pr(>F)
## group  3  1.7052 0.1998
##       19
```

A) Analyze Contrasts

Provide contrast coefficients that test three mutually orthogonal comparisons and test your contrasts in R. Summarize the results of your planned comparisons.

Conduct comparisons of the means using the following RQs:

1. Is there a difference in weight loss between B and D?
2. Is there a difference in weight loss between A and C?
3. Is there a difference in weight loss between A and the other diet options (B, C & D)?

Define and Test Contrasts in R

```r
# Define the Contrasts
# c(A, B, C, D)
c1 <- c(0, 1, 0, -1) # RQ1
c2 <- c(1, 0, -1, 0) # RQ2
c3 <- c(3, -1, -1, -1) # RQ3

# Test the Contrasts
contrasts(Diet$diet) <- cbind(c1, c2, c3)

# Check the Contrasts
contrasts(Diet$diet)
```

```
##   c1 c2 c3
## A  0  1  3
## B  1  0 -1
## C  0 -1 -1
## D -1  0 -1
```

```r
# Refit the model
oneway.diet3 <- lm(weightloss ~ diet, Diet)
summary(oneway.diet3)
```

```
##
## Call:
## lm(formula = weightloss ~ diet, data = Diet)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##   -2.4   -1.7    0.0    0.9    5.6
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  14.6500     0.4634  31.615  < 2e-16 ***
## dietc1        1.6000     0.6959   2.299  0.03300 *
## dietc2       -3.2000     1.0844  -2.951  0.00821 **
## dietc3       -0.1500     0.4634  -0.324  0.74970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.2 on 19 degrees of freedom
## Multiple R-squared:  0.6666, Adjusted R-squared:  0.6139
## F-statistic: 12.66 on 3 and 19 DF,  p-value: 8.881e-05
```

Summary of Results

To assess whether the type of diet had an impact on weight loss, we conducted a one-way analysis of variance. Weight loss was the dependent variable and type of diet (with four levels) was the independent variable. Diet type had statistical significance, F(3, 19) = 12.66, $p < .001$. In addition, the effect size ($\eta^2 = .67$) indicated that 66.7% of the variance in weight loss can be explained by the type of diet.

Since the type of diet was found to be related to weight loss, we evaluated the significance of three mutually orthogonal contrast coefficients. The difference in weight loss between those who used diet option B (M = 16.4, SD = 3.36) and option D (M = 16.4, SD = 1.92) was proven to be statistically significant (t(2.2, 19) = 1.60, p = 0.03), providing support for Hypothesis 1. Additionally, the difference in weight loss between those who used diet option A (M = 16.4, SD = 1.90) and C (M = 16.4, SD = 1.53) was proven to be statistically significant (t(2.2, 19) = -3.20, p = 0.01), providing evidence of Hypothesis 2. Hypothesis 3, which assumed that those who used option A (M = 16.4, SD = 1.90) would show difference in weight loss from the other options, did not reach statistical significance (t(2.2, 19) = -0.15, p = 0.75).

B) Calculate by Hand

Pick one of your contrasts and perform this calculation by hand and confirm that it is the same as the result obtained in part a. You do not need to calculate the means by hand or the MSW by hand because you already did this in Module 7. Instead, use the sample means and MSW from your output.

RQ1 by Hand

```
# Estimate the contrast of RQ1
# c1 <- c(0, 1, 0, -1) # RQ1
c1.con <- (0 * 11) + (1 * 16.4) + (0 * 18) + (-1 * 13.2)
c1.con
```

```
## [1] 3.2
```

```
# Calculate the Sum
c1.sum <- (((0)^2)/6) + (((1)^2)/5) + (((0)^2)/7) + (((-1)^2)/5)
c1.sum
```

```
## [1] 0.4
```

```
# Define the MSW (Mean Squares Within)
c1.MSW <- 61.304
c1.MSW
```

```
## [1] 61.304
```

```
# Calculate the Contrast
c1.Contrast <- (((c1.con)^2) / c1.sum) / c1.MSW
c1.Contrast
```

```
## [1] 0.417591
```

C) Pairwise Comparison

A researcher on your team would prefer not to conduct planned comparisons. Instead, this researcher wanted all pairwise comparisons. Pick a procedure, conduct the analysis in SPSS and R, and summarize what you found.

Tukey's HSD

```r
# Use Tukey's HSD to evaluate all pairwise comparisons.
library(emmeans)
THSD.diet <- emmeans(oneway.diet3, pairwise ~ diet)

# Summary of Tukey's HSD
summary(THSD.diet)
```

```
## $emmeans
##  diet emmean    SE df lower.CL upper.CL
##  A      11.0 0.898 19     9.12     12.9
##  B      16.4 0.984 19    14.34     18.5
##  C      18.0 0.832 19    16.26     19.7
##  D      13.2 0.984 19    11.14     15.3
##
## Confidence level used: 0.95
##
## $contrasts
##  contrast estimate   SE df t.ratio p.value
##  A - B        -5.4 1.33 19  -4.053  0.0035
##  A - C        -7.0 1.22 19  -5.718  0.0001
##  A - D        -2.2 1.33 19  -1.651  0.3755
##  B - C        -1.6 1.29 19  -1.242  0.6091
##  B - D         3.2 1.39 19   2.299  0.1334
##  C - D         4.8 1.29 19   3.725  0.0072
##
## P value adjustment: tukey method for comparing a family of 4 estimates
```

Summary of Results

To assess whether the type of diet had an impact on weight loss, we conducted a one-way analysis of variance. Weight loss was the dependent variable and type of diet (with four levels) was the independent variable. Diet type had statistical significance, F(3, 19) = 12.66, $p < .001$. In addition, the effect size ($\eta^2 = .67$) indicated that 66.7% of the variance in weight loss can be explained by the type of diet.

Since the type of diet was found to be related to weight loss, post hoc analyses were conducted using Tukey's HSD. Results from this procedure indicate that those who used diet option A ($M = 11$, $SD = 0.90$) lost significantly more weight than those who used option B ($M = 16.4$, $SD = 0.98$), $t(19)$ = -5.4, $p < 0.05$ and C ($M = 18$, $SD = 0.83$), $t(19)$ = -7.0, $p < 0.05$. Additionally, those who used option C ($M = 18$, $SD = 0.83$) lost significantly more weight than D ($M = 13.2$, $SD = 0.98$), $t(19) = 4.8$, $p < 0.05$.