BIOS 662

Homework 6 Solution

November, 2018

Part 1

(a) Analysis Plan

Standard analysis of variance methodology will be used. Triglyceride level Y will be modeled by

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

where the index $i = 1,2,3,4,5$ denotes the groups, receiving 0, 75, 150, 300, or 600 mg/kg of peppermint extract, respectively, the index j denotes the j^{th} rat within the i^{th} dosage group, μ_i denotes the population mean triglyceride level in the i^{th} dosage group, and the ε_{ij} are assumed to be independent and identically distributed as $N(0,\sigma^2)$. The primary interest is in pairwise comparisons between groups, to determine which groups differ from one another. The group sizes are unequal, varying from 7 to 10, so it is more appropriate to use the Scheffé or Bonferroni method to adjust for multiple comparisons. We'll use Scheffé's method for the primary analysis.

Standard ANOVA diagnostics will be used to assess the fit of the model above. In the event of violations of the assumptions of ANOVA (in particular, homogeneity of variance or normality), the Box-Cox family of transformations will be used to find the transformation of the data that minimizes the MSE. We recognize that the sample sizes are rather small so that only quite large departures from the assumptions are likely to be detectable.

(b) Analyses

Figure 1 has boxplots of the data for each group, with the individual points overlaid and Table 1 has corresponding summary statistics.

Group	N	Median	Mean	Std Dev
1	10	251.5	244.2	17.87
2	10	241.0	238.1	10.30
3	7	230.0	228.1	8.55
4	10	220.0	220.5	6.67
5	9	210.0	209.9	5.09

Table 1: Summary statistics for data from peppermint extract study

The boxplots, summary statistics and diagnostics for the ANOVA model suggest that the homogeneity of variance assumption is questionable. In particular, looking at the

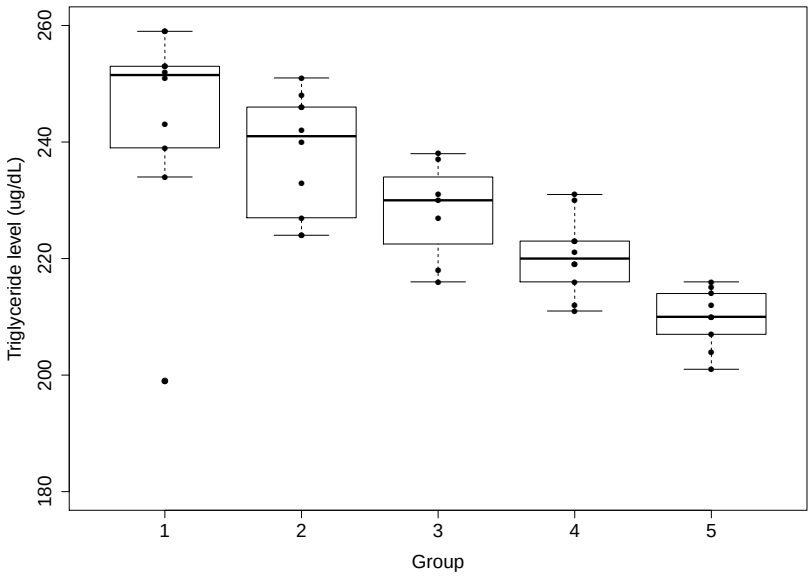

Figure 1: Triglyceride level by peppermint extract dosage group

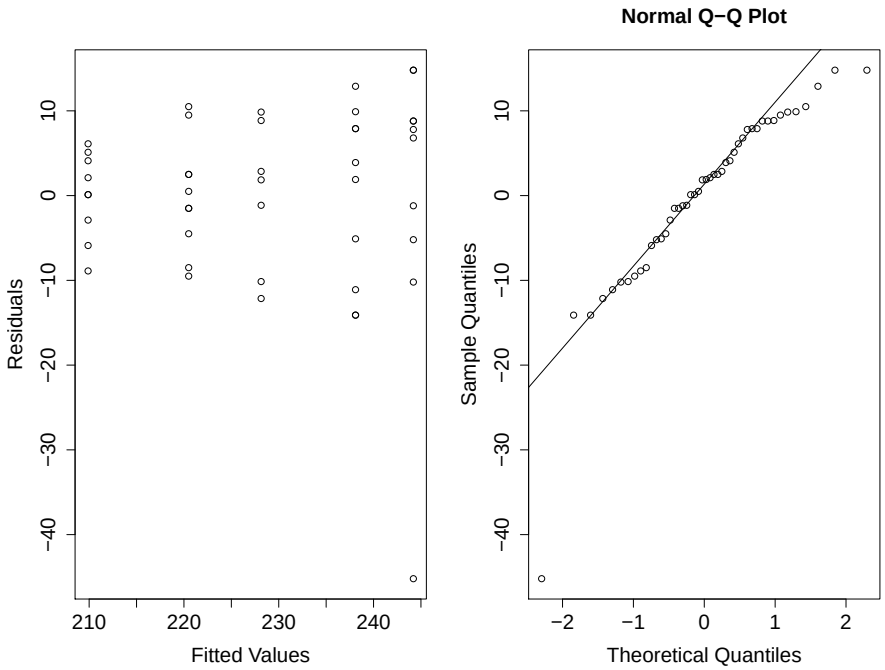

Figure 2: Residual plots from fitted ANOVA model

standard deviations in Table 1 and the residual plot (left panel of Figure 2) it appears that the variance decreases with increasing peppermint extract dose. The trend in the standard deviations is influenced substantially by a single observation, the 199 for rat R15. Omitting this observation reduces the standard deviation for group 1 from 17.86 to 8.70 (and increases that group's mean from 244.2 to 249.2). However, even without excluding this observation, the modified Levene test ($p = 0.38$, see below) does not reject the hypothesis of homogeneity of variance.

```
proc anova;
   class group;
   model trg = group;
   means group / hovtest=bf;

  Brown and Forsythe's Test for Homogeneity of trg Variance
       ANOVA of Absolute Deviations from Group Medians

                       Sum of        Mean
Source        DF     Squares      Square    F Value    Pr > F

group          4       307.6     76.9122       1.09    0.3759
Error         41      2902.5     70.7928
```

In the right panel of Figure 2, the QQ-plot for the residuals from the fitted model suggests that there is departure from normality in the tails of the distribution of the residuals. Pearson's correlation coefficient from the QQ-plot is 0.93 (see below). Here $n = 46$ and as $r = 0.93$ is smaller than the critical value of 0.97 for $n = 40$ on page 9 of the "ANOVA, Part III" overheads, the normality assumption is questionable.

```
> group<-as.factor(group)
> fit<-aov(trg~group)
> par(mfcol=c(1,2))
> plot(fit$fitted.values,fit$residuals)
> qq<-qqnorm(fit$residuals)
> qqline(fit$residuals)
> cor.test(qq$x,qq$y)

        Pearson's product-moment correlation

data:  qq$x and qq$y

sample estimates:
      cor
0.9255097
```

We are not given information about whether some of the rats were from the same litter so we don't have enough information to determine if the independence assumption is satisfied.

Because it appears that the normality assumption is violated and the sample size in each group is rather small for the Central Limit Theorem to help, we will use the Box-Cox

method to investigate potential transformations. From the SAS code and output below, $\lambda = 0.6$ minimizes the MSE. The 95% confidence interval for λ extends from -3.0 to 4.2, indicating a large amount of uncertainty about the most appropriate value for λ. Because a square root transformation ($\lambda = 0.5$) is close to the optimal value, we'll try this to see whether it improves the normality of the residuals. Using this transformation, both parts of Figure 3 are very similar to those in Figure 2, Pearson's correlation coefficient from the QQ-plot is 0.92, little changed from the 0.93 using the untransformed data and the modified Levene test ($p = 0.44$) again does not indicate lack of homogeneity of variance.

(The results of these diagnostic checks are somewhat surprising. For this example I did not have the original data. I saw summary statistics in a journal article and generated data to yield similar summary statistics. I generated the residuals from the normal distribution, with different variances for the 5 groups, yet the diagnostics indicate that normality is questionable rather than homogeneity of variance. So, violation of one of the assumptions may manifest as violation of one of the other assumptions.)

```
data hw6;
   set hw6;

grp1=0; grp2=0;grp3=0;grp4=0;grp5=0;
if group=1 then grp1=1;
   else if group=2 then grp2=1;
   else if group=3 then grp3=1;
   else if group=4 then grp4=1;
   else if group=5 then grp5=1;


%boxcox(resp=trg,model=grp2 grp3 grp4 grp5,lopower=-2,hipower=2,
           npower=41,data=hw6);
```

Box-Cox Power (lambda)	Log Likelihood	Root mean squared error	0.95 Confidence Interval
0.0	-109.931	10.9113	*+
0.1	-109.916	10.9079	*
0.2	-109.905	10.9051	*
0.3	-109.896	10.9030	*
0.4	-109.890	10.9015	*
0.5	-109.886	10.9007	*+
0.6	-109.885	10.9005	<
0.7	-109.887	10.9010	*
0.8	-109.892	10.9021	*
0.9	-109.899	10.9038	*
1.0	-109.909	10.9062	*+

(The output above has been edited to show just part of the range of values of λ.)

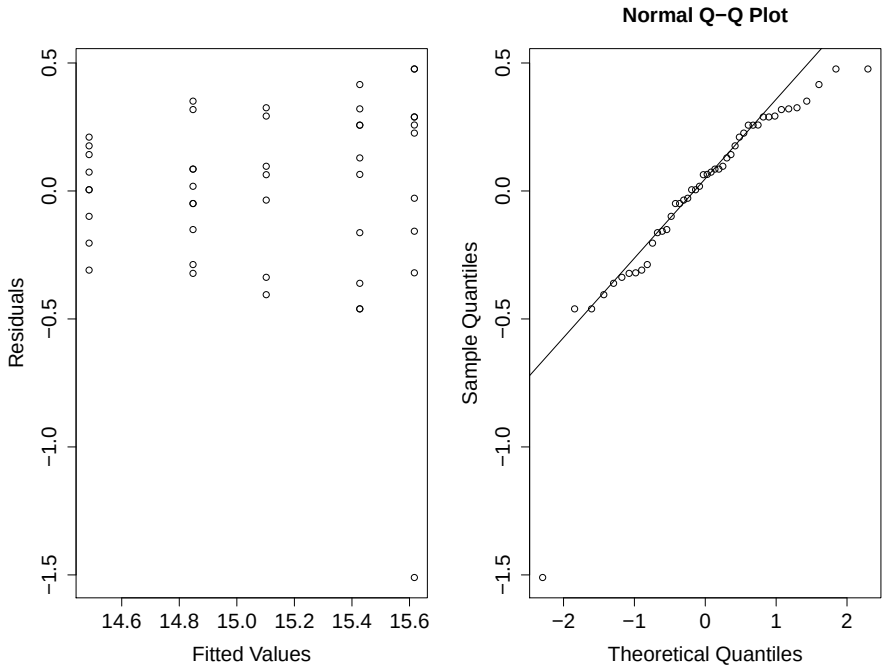

Figure 3: Residual plots from fitted ANOVA model after square root transformation

We now investigate pairwise differences in the means, using the untransformed data, with Scheffé's method for adjusting for the multiple comparisons. SAS code and output are presented below. Groups 1 and 2 differ significantly from groups 4 and 5 and group 3 differs significantly from group 5 but not from the other three groups. These comparisons are presented schematically in Figure 4.

Results using Bonferroni or Tukey's method are similar except that with those methods groups 1 and 3 are significantly different. The corresponding schematic is in Figure 5.

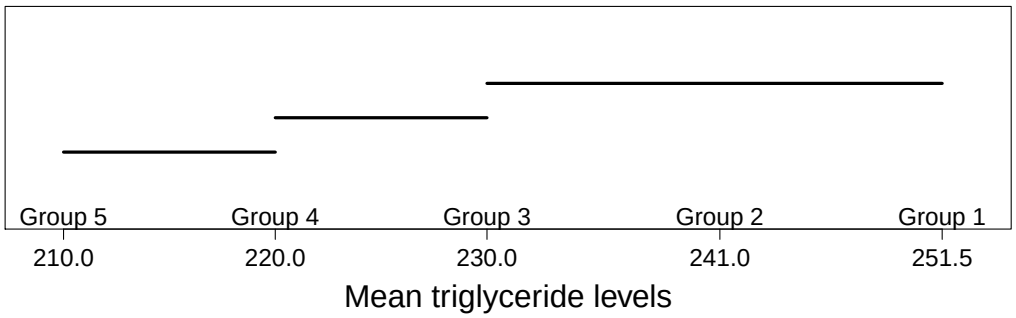

Figure 4: Using Scheffé's method; lines join groups that do not differ significantly

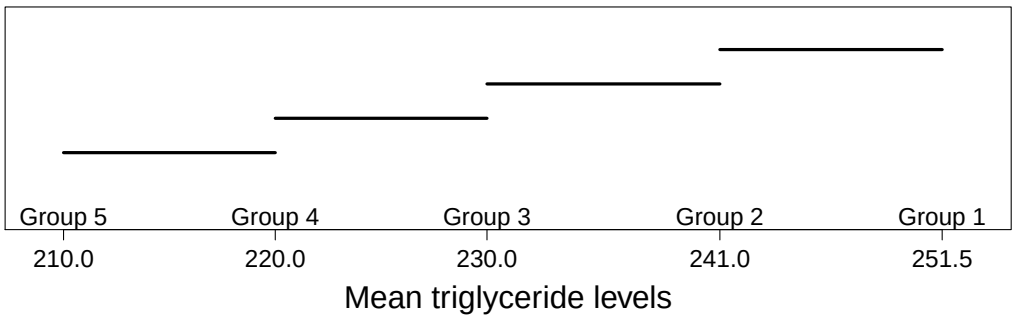

Figure 5: Using Bonferroni; lines join groups that do not differ significantly

```
proc anova data=hw6;
   class group;
   model trg = group;
   means group / scheffe;
```

Dependent Variable: trg

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	7144.55832	1786.13958	15.02	<.0001
Error	41	4876.74603	118.94503		
Corrected Total	45	12021.30435			

Scheffe's Test for trg

NOTE: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	41
Error Mean Square	118.945
Critical Value of F	2.59997

Comparisons significant at the 0.05 level are indicated by ***.

group Comparison	Difference Between Means	Simultaneous 95% Confidence Limits		
1 - 2	6.100	-9.629	21.829	
1 - 3	16.057	-1.275	33.390	
1 - 4	23.700	7.971	39.429	***
1 - 5	34.311	18.151	50.471	***
2 - 1	-6.100	-21.829	9.629	
2 - 3	9.957	-7.375	27.290	
2 - 4	17.600	1.871	33.329	***
2 - 5	28.211	12.051	44.371	***
3 - 1	-16.057	-33.390	1.275	
3 - 2	-9.957	-27.290	7.375	
3 - 4	7.643	-9.690	24.975	
3 - 5	18.254	0.529	35.979	***
4 - 1	-23.700	-39.429	-7.971	***
4 - 2	-17.600	-33.329	-1.871	***
4 - 3	-7.643	-24.975	9.690	
4 - 5	10.611	-5.549	26.771	
5 - 1	-34.311	-50.471	-18.151	***
5 - 2	-28.211	-44.371	-12.051	***
5 - 3	-18.254	-35.979	-0.529	***
5 - 4	-10.611	-26.771	5.549	

(c) Conclusions

Mean triglyceride levels appear to be included by dosage of peppermint extract, with higher doses associated with lower triglyceride levels. In comparing pairs of dosages, adjacent dosages generally do not have significantly different effects, though that may be because with the small sample sizes there is limited power to detect relatively small differences. Pairs of doages that are further apart generally do result in significant differences in triglyceride levels.

Part 2

Note that the problem specifically says to use your parametric ANOVA model to address these. This can be done using contrasts. For (i) we want to test

$$H_0 : \mu_1 - (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4 = 0 \quad \text{vs.} \quad H_A : \mu_1 \neq (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4$$

or, equivalently,

$$H_0 : \mu_1 = (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4 \quad \text{vs.} \quad H_A : \mu_1 \neq (\mu_2 + \mu_3 + \mu_4 + \mu_5)/4$$

and one way to approach (ii) is to test

$$H_0 : -2 \cdot \mu_1 - 1 \cdot \mu_2 + 0 \cdot \mu_3 + 1 \cdot \mu_4 + 2 \cdot \mu_5 = 0$$

vs.

$$H_A : -2 \cdot \mu_1 - 1 \cdot \mu_2 + 0 \cdot \mu_3 + 1 \cdot \mu_4 + 2 \cdot \mu_5 \neq 0$$

In each case H_0 is of the form $\sum_{i=1}^{5} c_i \mu_i = 0$, with $\sum_{i=1}^{5} c_i = 0$.

Below is SAS code and corresponding output using contrasts to test these hypotheses. In both instances we reject H_0 and conclude that (i) the mean triclyceride level in rats on placebo differs significantly from that in rats given peppermint extract and (ii) there appears to be a linear relationship between group number and mean triclyceride level.

```
proc glm;
   class group;
   model trg = group;
   contrast 'Placebo vs. others' group 1 -0.25 -0.25 -0.25 -0.25;
   contrast 'Linear association' group -2 -1 0 1 2;
```

```
Dependent Variable: trg
                                 Sum of
Source                    DF        Squares    Mean Square    F Value    Pr > F
Model                      4     7144.55832     1786.13958      15.02    <.0001
Error                     41     4876.74603      118.94503
Corrected Total           45    12021.30435

Contrast                  DF    Contrast SS    Mean Square    F Value    Pr > F
Placebo vs. others         1    3129.040058    3129.040058      26.31    <.0001
Linear association         1    7117.919622    7117.919622      59.84    <.0001
```

How does using a contrast in an ANOVA model to test for a linear association differ from using a linear regression model? Below is SAS code and corresponding output from a regression model using group as the predictor variable. The F value is similar but not identical to that for the "Linear association" contrast. The difference arises because the ANOVA model fits 5 parameters (the 5 group means), leaving 41 degrees of freedom for the error whereas the regression model fits two parameters (intercept and slope), leaving 44 degrees of freedom for the error.

```
proc glm;
   model trg = group;

Dependent Variable: trg
                              Sum of
Source              DF       Squares    Mean Square    F Value    Pr > F
Model                1    7094.02783     7094.02783      63.35    <.0001
Error               44    4927.27652      111.98356
Corrected Total     45   12021.30435
```

Part 3

Here we fit a regression model using actual dose rather than group number. The regression model is

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

where X_i is the dose received by the i^{th} rat, Y_i is the triglyceride level of the i^{th} rat and the ε_i are iid $N(0, \sigma^2)$. SAS code and corresponding output are below. Dose is significantly associated with triglyceride levels. The estimated regression model is

$$\widehat{Y}_i = \widehat{\alpha} + \widehat{\beta} X_i.$$

From the output below, $\widehat{\beta} = -0.057$. That is, according to the model, mean triglyceride level decreases by 0.057 μg/dL for each 1 mg/kg increase in peppermint extract dose (or decreases by 5.7 μg/dL for each 100 mg/kg increase in peppermint extract dose). The corresponding 95% confidence interval is

$$\widehat{\beta} \pm t_{N-2,0.975}\text{SE}(\widehat{\beta}) = -0.057 \pm 2.02 \times 0.0076 = (-0.072, -0.041).$$

The predicted mean for the group with dose $= 0$ is $241.1 - 0.0566 \cdot 0 = 241$. Using SAS, the 95% confidence interval for the point on the line at dose $= 0$ is $(236.4, 245.8)$. This compares favorably with the actual mean for group 1, namely 244.2. On the other hand, omitting the outlier rat (R15) the estimate of the mean for dose $= 0$ becomes 243.1, with 95% confidence interval $(239.2, 247.0)$ whereas the mean for group 1 becomes 249.2.

(I hadn't stated explicitly whether the regression should have been done with the observations in the dose $= 0$ group included or removed. The above results are when those observations are included. Fitting the model without the dose $= 0$ observations the predicted mean for dose $= 0$ is 238.6 and the 95% confidence interval for the point on the line at dose $= 0$ is $(233.8, 243.3)$.)

```
proc glm;
   model trg = dose / solution clparm;

Dependent Variable: trg
                                 Sum of
Source                DF        Squares    Mean Square    F Value    Pr > F
Model                  1     6694.23245     6694.23245      55.29    <.0001
Error                 44     5327.07190      121.06982
Corrected Total       45    12021.30435

R-Square     Coeff Var      Root MSE      trg Mean
0.556864      4.814019      11.00317      228.5652

                            Standard
Parameter     Estimate        Error  t Value  Pr > |t|    95% Confidence Limits
Intercept  241.1084873   2.34039305   103.02    <.0001   236.3917350  245.8252395
dose        -0.0565677   0.00760740    -7.44    <.0001    -0.0718994   -0.0412360
```