STA305/1004 in-class problem

Nathan Taback

2019-10-23

Answer the questions based on the following experimental design.

A study at UofT recruited twenty-one students to complete a thirty minute survey on their diet and eating habits at the end of an academic year. Students were paid $10 to complete the survey and answer a few questions. The data below shows their weight gain from September to April classified by the frequency that students ate fast food. In group A students reported eating fast food once per month; the students in group B reported eating fast food twice per month; and the students in group C reported eating fast food four times per month.

	A	B	C
	1.02	1.44	0.73
	-0.32	0.40	1.11
	0.27	-0.30	3.63
	0.08	2.27	-0.88
	0.51	-0.17	1.21
	0.34	1.07	1.22
	1.30	1.18	2.40
Treatment Average	0.46	0.84	1.35
Treatment SD	0.55	0.92	1.40

The researchers analyzed the data using R.

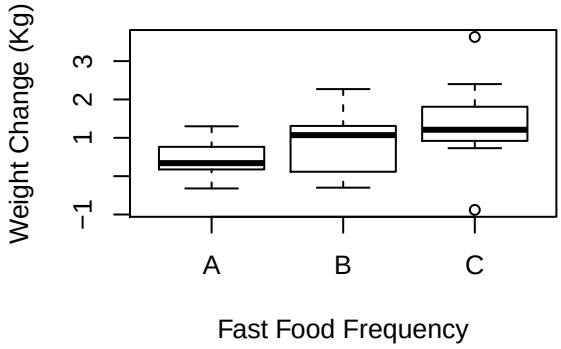
```
surveydat <- read.csv("surveydat.csv")
aovsurvey <- aov(wtchange~grp,data=surveydat)
summary(aovsurvey)
```

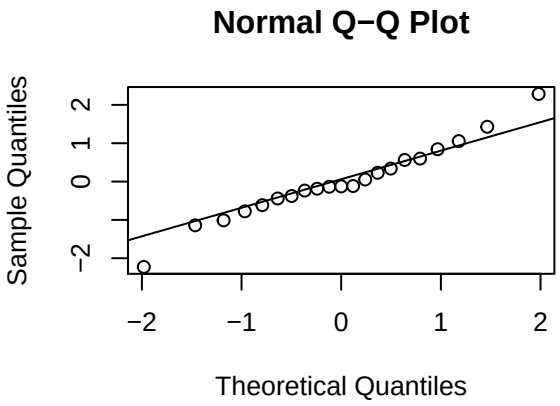
```
##             Df Sum Sq Mean Sq F value Pr(>F)
## grp          2   2.78   1.390   1.341  0.287
## Residuals   18  18.66   1.037
```

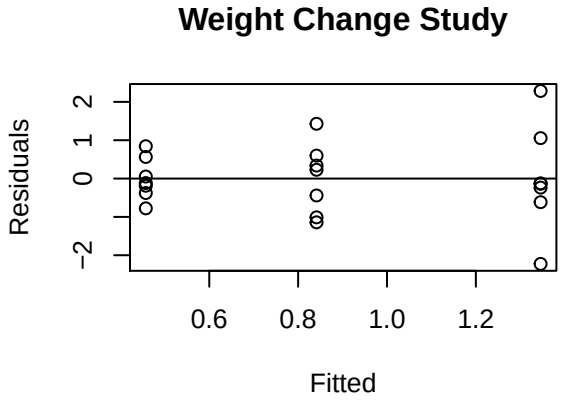
```
mod1 <- lm(wtchange~grp,data=surveydat)
summary(mod1)$coefficients
```

```
##              Estimate Std. Error   t value  Pr(>|t|)
## (Intercept) 0.4571429  0.3848706 1.1877832 0.2503565
## grpB        0.3842857  0.5442893 0.7060322 0.4892064
## grpC        0.8885714  0.5442893 1.6325353 0.1199371
```

```r
par(mfrow = c(2, 2)) # arrange plots in two columns
boxplot(wtchange~grp,data = surveydat,ylab="Weight Change (Kg)", xlab="Fast Food Frequency")
qqnorm(aovsurvey$residuals);qqline(aovsurvey$residuals)
plot(aovsurvey$fitted.values, aovsurvey$residuals,ylab="Residuals", xlab="Fitted",
     main="Weight Change Study"); abline(h=0)
```







Questions

(a) Is this study and experiment or observational study? What are the treatments?

(b) Would it have been feasible for the researcher to randomized students to the treatments? What randomization scheme (assigning the subjects to the treatments) could the researcher use to accomplish the randomization?

(c) What are the null and alternative hypotheses that the researchers are testing in the ANOVA table? Is there evidence to reject the null hypothesis? Explain.

(d) Are the statistical assumptions used to calculate the P-value in the ANOVA table satisfied? Explain.

(e) Interpret the least squares estimates in `mod1`. Are you surprised that the P-values for the parameters are not small? Briefly explain.

(f) The researcher is convinced that the results of the study would have provided strong evidence that eating fast food four times per month causes students to gain weight, if the sample size in each group was larger. Is this a valid statement? Explain.