

# MTH 3220 Lab 5

Due Thu., Oct. 10

## 1 Part A: Checking Assumptions in One-Factor ANOVA

### 1.1 Blood Coagulation Times Data Set (Cont'd)

The data below, used in **Labs 3** and **4**, are coagulation times (seconds) for samples of blood drawn from 24 animals receiving four different **diets**, A, B, C, and D.

Diet (Treatment)			
A	B	C	D
62	63	68	56
60	67	66	62
63	71	71	60
59	64	67	61
59	65	68	63
63	66	68	64
62	64	66	63
61	67	68	59

The data are in the file **blood.txt** *stacked* in the order in which the blood samples were drawn.

For a *one-factor ANOVA F test* of the hypotheses

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_a : \text{At least two } \mu_i \text{'s are different}$$

the **assumptions** are that in the model

$$X_{ij} = \mu_i + \epsilon_{ij},$$

the **random errors**  $\epsilon_{ij}$  are iid  $N(0, \sigma)$ .

1. Use `read.table()` to read the data from **blood.txt** into an R *data frame* named, say, `my.data`.
2. In **Lab 3** we carried out a **one-factor ANOVA F test** of the hypotheses

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_a : \text{At least two } \mu_i \text{'s are different.}$$

by typing:

```
my.anova <- aov(CoagTime ~ Diet, data = my.data)
summary(my.anova)
```

The object `my.anova` is a *list*. To verify, type:

```
is.list(my.anova)
```

A *list* is like a *vector*, but its elements can be any R objects. To see the names of the objects stored in `my.anova`, type:

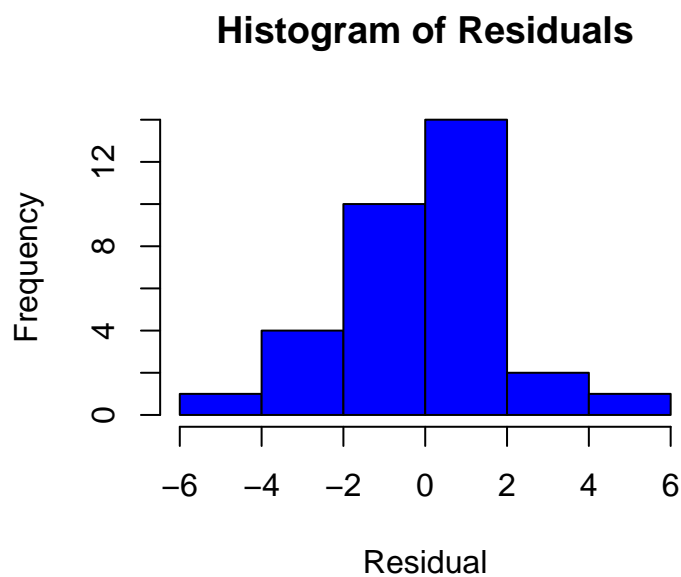
```
names(my.anova)
```

3. The operator `$` is used to extract elements from a *list*. For example,

```
my.anova$residuals
```

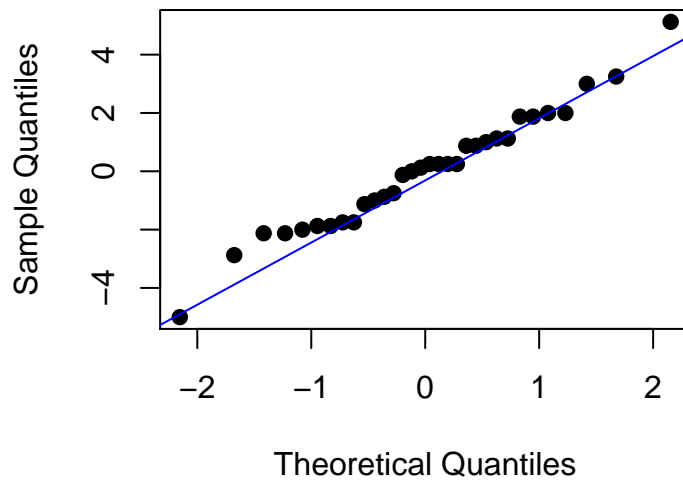
will extract the *residuals*.

Make a **histogram** of the **residuals** using `hist()`. Your plot should look similar to this:



4. Now make a **normal probability plot** of the **residuals** using `qqnorm()` followed by `qqline()`. Your plot should look similar to this:

## Normal Q-Q Plot



5. The **fitted values** are obtained via:

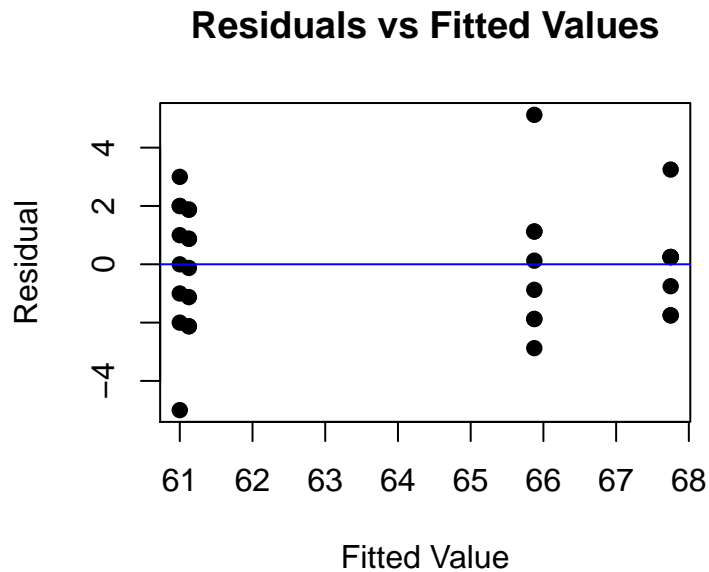
```
my.anova$fitted.values
```

Make a plot of the **residuals** (**y axis**) versus the **fitted values** (**x axis**) using `plot()`

Then add a horizontal line to the plot at **y = 0** by typing:

```
abline(h = 0, col = "blue")
```

Your plot should look similar to the one below.



## 2 Part B: Equivalence Between Two-Sample $t$ Test and ANOVA $F$ Test

### 2.1 Blood Coagulation Times Data Set (Cont'd)

Let's pretend there were only **two diets** in the blood coagulation study, diets A and B.

1. Create the subset of `my.data` that only contains diets A and B, for example by typing:

```
ab_only <- subset(my.data, subset = Diet %in% c("A", "B"))
```

2. Carry out a **two-sided pooled two-sample  $t$  test** of

$$H_0 : \mu_1 = \mu_2$$

$$H_a : \mu_1 \neq \mu_2$$

by typing:

```
t.test(CoagTime ~ Diet,
       data = ab_only,
       var.equal = TRUE,
       alternative = "two.sided")
```

3. Now carry out **one-factor ANOVA  $F$  test** of the *same* hypotheses by typing:

```
my.anova <- aov(CoagTime ~ Diet, data = ab_only)
summary(my.anova)
```