Linear Models Foundations Handout

1 Initialization

```
> library(NCStats)
```

2 2-sample t-Test

You must change the directory to where the following file is located.

```
> setwd("C://aaaWork//Class Materials//MTH207//Lecture/Handouts//")
> sdir <- read.table("SexDirection.txt",head=TRUE)
> str(sdir)

'data.frame':   60 obs. of 2 variables:
 $ abserr: int 13 13 38 59 58 8 130 68 23 5 ...
 $ sex   : Factor w/ 2 levels "female","male": 2 2 2 2 2 2 2 2 2 2 ...

> sdir$f.sex <- factor(sdir$sex,levels=c("male","female"))
> str(sdir)

'data.frame':   60 obs. of 3 variables:
 $ abserr: int 13 13 38 59 58 8 130 68 23 5 ...
 $ sex   : Factor w/ 2 levels "female","male": 2 2 2 2 2 2 2 2 2 2 ...
 $ f.sex : Factor w/ 2 levels "male","female": 1 1 1 1 1 1 1 1 1 1 ...

> hist(abserr~f.sex,data=sdir,xlab="Absolute Error")
```
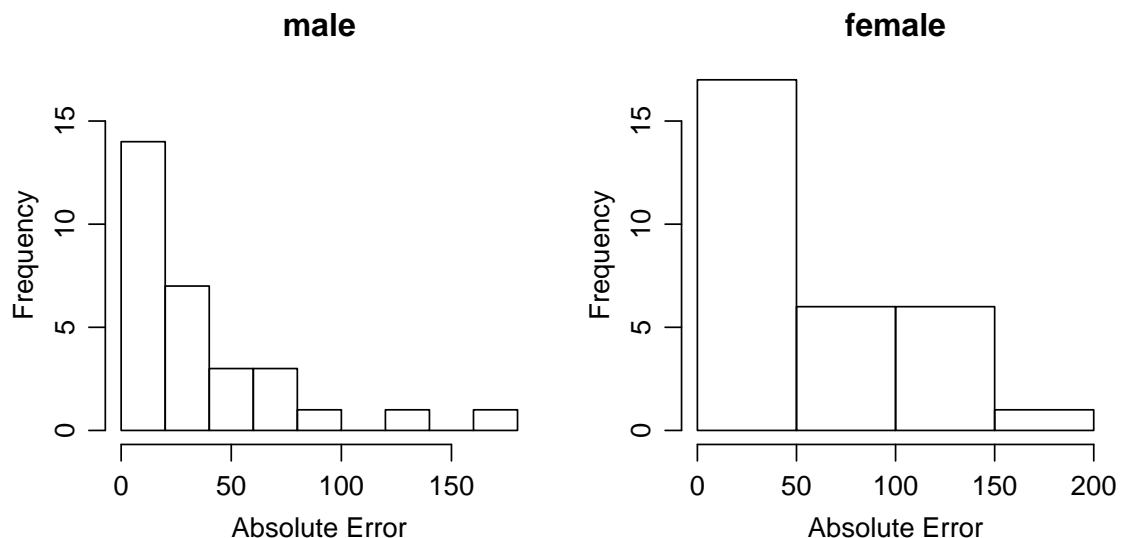
```
> Summarize(abserr~f.sex,data=sdir,digits=2)

    f.sex  n Mean St. Dev. Min. 1st Qu. Median 3rd Qu. Max.
1   male 30 37.6 38.48654    3   11.50   22.5   58.75  167
2 female 30 55.8 48.25964    3   15.75   35.0   88.25  176

> leveneTest(abserr~f.sex,data=sdir)

Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  1  2.1692 0.1462
      58

> t.test(abserr~f.sex,data=sdir,var.equal=TRUE)

        Two Sample t-test

data:  abserr by f.sex
t = -1.6149, df = 58, p-value = 0.1118
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -40.758823   4.358823
sample estimates:
  mean in group male mean in group female
              37.6                 55.8
```

3 2-Sample t-Test as a Linear Model

This example begins by assuming that you are using the same data as shown in Section 2.

```
> lm1 <- lm(abserr~f.sex,data=sdir)
> anova(lm1)

Analysis of Variance Table

Response: abserr
          Df Sum Sq Mean Sq F value Pr(>F)
f.sex      1   4969  4968.6   2.608 0.1118
Residuals 58 110496  1905.1

> summary(lm1)

Call:
lm(formula = abserr ~ f.sex, data = sdir)

Residuals:
   Min     1Q Median     3Q    Max
-52.80 -29.00 -18.10  22.85 129.40

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   37.600      7.969   4.718 1.54e-05
f.sexfemale   18.200     11.270   1.615    0.112
```
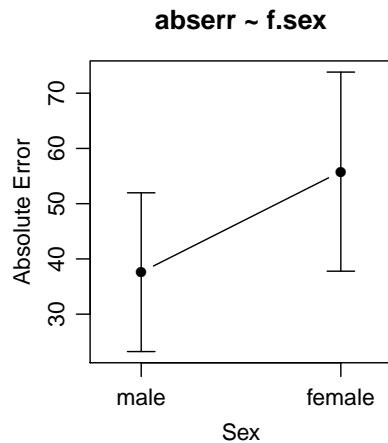
2

```
Residual standard error: 43.65 on 58 degrees of freedom
Multiple R-squared: 0.04303,        Adjusted R-squared: 0.02653
F-statistic: 2.608 on 1 and 58 DF,  p-value: 0.1118

> confint(lm1)

                 2.5 %    97.5 %
(Intercept) 21.648503 53.55150
f.sexfemale -4.358823 40.75882

> fitPlot(lm1,xlab="Sex",ylab="Absolute Error")
```
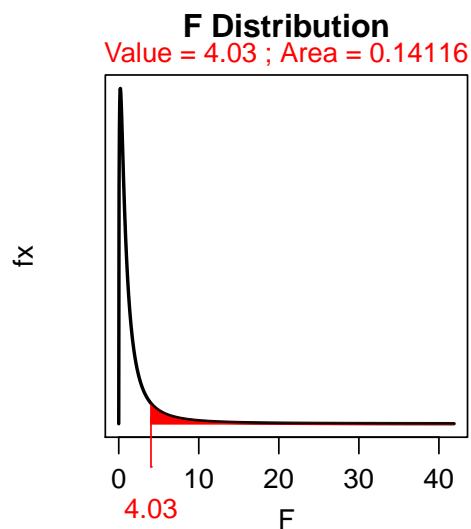
abserr ~ f.sex

4   F Distribution Calculations

```
> distrib(4.03,distrib="f",df1=3,df2=3,lower.tail=FALSE)
```

F Distribution
Value = 4.03 ; Area = 0.14116


```
> distrib(4.03,distrib="f",df1=3,df2=3,lower.tail=FALSE,plot=FALSE)

[1] 0.1411648
```