Exercise set 5

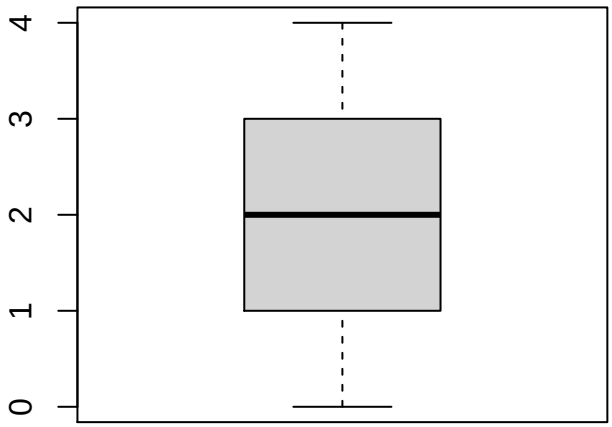
Stat 251

(1) To see if eating breakfast before work can help with productivity, a randomized experiment was conducted with 10 people. To test for a difference, the total amount of work hours for each person was recorded for the hours they worked with and again without breakfast.
   (a) Is there sufficient evidence that there is a significant difference in working with or without breakfast? Conduct a hypothesis test; state the test statistic, *pvalue*, *df*, results, and conclusion in context
   (b) State the kind of error *in context*
   (c) Estimate the true mean difference with 95% confidence; interpret.

```
breakf
```

```
     1 2 3 4 5 6  7 8 9 10
w    8 7 9 5 9 8 10 7 6  9
wo   6 5 5 4 7 7  7 5 6  5
diff 2 2 4 1 2 1  3 2 0  4
```

```
boxplot(diff)
```

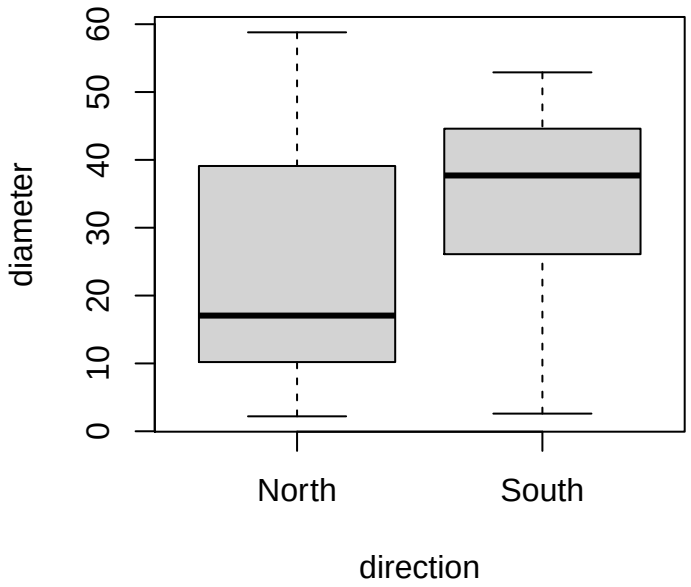



```
t.test(hours~eats,data=food,paired=T)
```

```
    Paired t-test

data:  hours by eats
t = 5.1612, df = 9, p-value = 0.0005942
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 1.179562 3.020438
sample estimates:
mean of the differences
                    2.1
```

(2) *Trees.* A study of 584 longleaf pine trees in the Wade Tract in Thomas County, Georgia had several purposes. To see if there is a difference in their sizes (in diameters) in two separate areas of the Wade Tract (northern and southern areas), a random sample of 30 trees from the northern area and 30 trees from the southern area was taken
   (a) Estimate the true difference in mean tree sizes between the northern and southern parts of the Wade Tract with 95% confidence; interpret
   (b) Is there a significant difference in the mean diameter of trees in the north versus the trees in the south? Conduct hypothesis test; state the test statistic, *pvalue*, *df*, results, and conclusion in context
   (c) State the kind of error *in context*

```
boxplot(diameter~direction,data=trees)
```




```
t.test(diameter~direction,data=trees)
```

```
    Welch Two Sample t-test

data:  diameter by direction
t = -2.6286, df = 55.725, p-value = 0.01106
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -19.090199  -2.576468
sample estimates:
mean in group North mean in group South
          23.70000            34.53333
```

(3) A 2010 Pew Research foundation poll[1] indicated that among 1099 college graduates, 33% watch the Daily Show. Meanwhile, 22% of the 1100 people with a high school degree (but no college degree) watch The Daily Show.
   (a) Is there sufficient evidence that the proportion of those US adults with a college degree is higher than the proportion who watch The Daily Show that do not have a college degree (but do have a high school degree)? Conduct hypothesis test; state the test statistic, *pvalue*, *df*, results, and conclusion in context
   (b) State the kind of error could have been made in context of the problem
   (c) Estimate the true difference in proportions with 95% confidence, interpret.

```
t.test(x,y,alternative='g')
```

```
    Welch Two Sample t-test

data:  x and y
t = 5.8006, df = 2160.8, p-value = 3.79e-09
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.0785725       Inf
sample estimates:
mean of x mean of y
0.3296903 0.2200000
```

```
t.test(x,y)
```

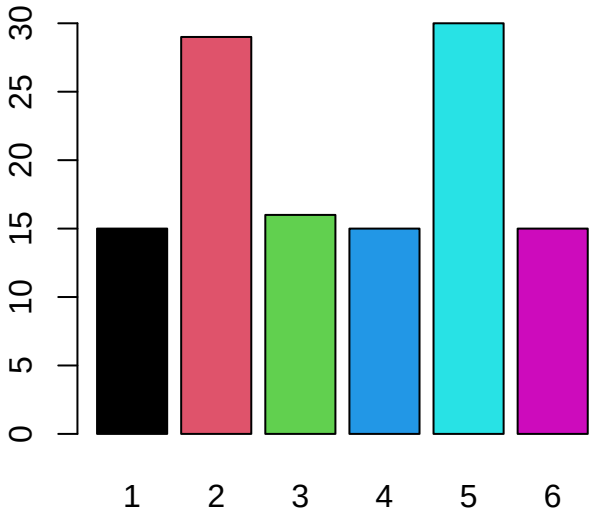
```
    Welch Two Sample t-test

data:  x and y
t = 5.8006, df = 2160.8, p-value = 7.58e-09
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.07260627 0.14677442
sample estimates:
mean of x mean of y
0.3296903 0.2200000
```

(4) A six-sided die is rolled 120 times to see if the die is a fair die.
   (a) Finish the following table
   (b) Is there evidence that the die is a fair die? Conduct hypothesis test; state the test statistic, *pvalue*, *df*, results, and conclusion in context
   (c) State the kind of error could have been made in context of the problem

Face value	Frequency	Expected
1	15	
2	29	
3	16	
4	15	
5	30	
6	15	

[1]The Pew Research Center, *Americans Spending More Time Following the News*, data collected June 8-28, 2010.

```
barplot(xtabs(observed~face),col=1:6)
```



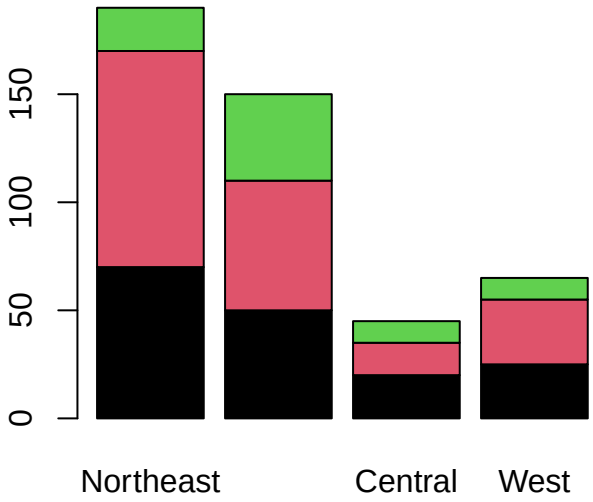

```
chisq.test(observed,p=p)
```

```
    Chi-squared test for given probabilities

data:  observed
X-squared = 13.6, df = 5, p-value = 0.01836
```

(5) A major food manufacturer is concerned that the sales for its skinny fries have been decreasing. As a part of a feasibility study, the company conducts research into the types of fries sold across the country to determine if the type of fries sold is independent of the area of the country. (a) Is there evidence that the preference of type of fries is related to the are of the country they live in? Conduct hypothesis test; state the test statistic, *pvalue*, *df*, results, and conclusion in context (b) State the kind of error could have been made in context of the problem

Type of Fries	*Northeast*	*South*	*Central*	*West*	**Total**
skinny	70	50	20	25	165
curly	100	60	15	30	205
steak	20	40	10	10	80
Total	190	150	45	65	450

```
barplot(fries,col=1:3)
```




```
chisq.test(fries)
```

```
    Pearson's Chi-squared test

data:  fries
X-squared = 18.837, df = 6, p-value = 0.004448
```