STAT 120 HW 7

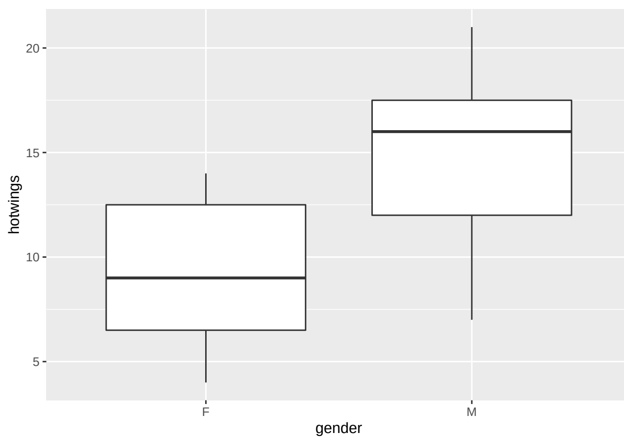
Victor Huang

5/17/2021

Section 7.6: Some parts ask you to create quantile-normal plots split by levels of a factor. The code to do this is right above the exercises (also see the example on pages 76-77).

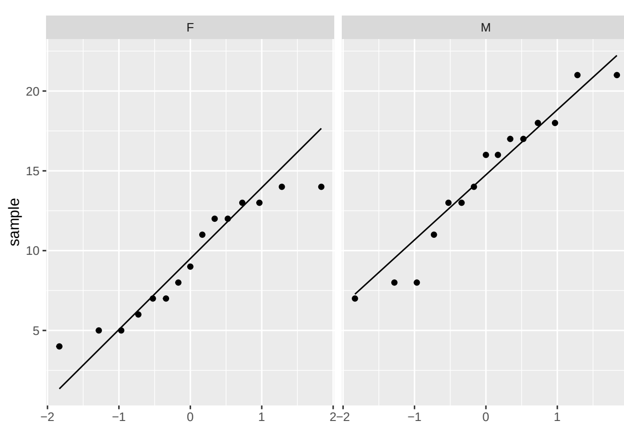
2

a.



b.

```
## # A tibble: 2 x 3
##   gender `mean(hotwings)` `sd(hotwings)`
##   <chr>              <dbl>          <dbl>
## 1 F                   9.33           3.56
## 2 M                  14.5            4.50
```



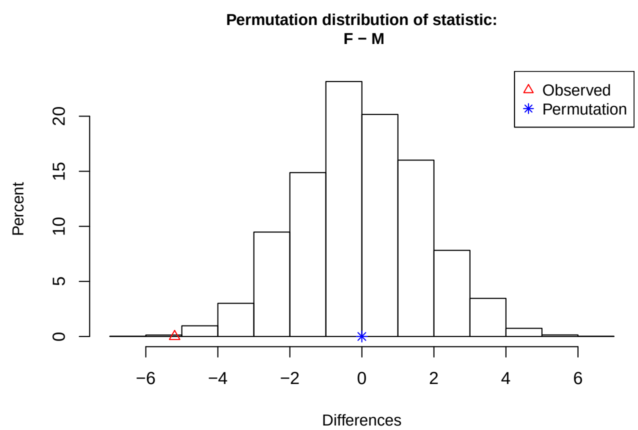
From the graph, one can see that there are some clear outliers to the rough line prescribed by the graph. As such, a normal distribution may not be appropriate and normality may not be met.

c.

```
##
##  Welch Two Sample t-test
##
## data:  hotwings by gender
## t = -3.5094, df = 26.584, p-value = 0.001619
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.242507 -2.157493
## sample estimates:
## mean in group F mean in group M
##        9.333333       14.533333
```

We can conclude that there is a difference in the average number of hotwings consumed between males and females. This is assuming that the data does not match a normal distribution.

d.




```
##
##  ** Permutation test **
##
##  Permutation test with alternative: two.sided
##  Observed statistic
##   F :  9.33333    M :  14.53333
##  Observed difference: -5.2
##
##  Mean of permutation distribution: -0.00151
##  Standard error of permutation distribution: 1.7436
##  P-value:  0.0026
##
##  *-------------*
```

As one can see, we reach the same conclusion

3

a.

```
##
##  One Sample t-test
##
## data:  facesymmetry$l.pupil
## t = 69.095, df = 30, p-value < 2.2e-16
```

```
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  31.62929      Inf
## sample estimates:
## mean of x
##  32.42581
```

From the data displayed above, we fid a 95 confidence interval between 31.62929 to Inf.

b.

Since we get the right and left eye measurements from the same person, this would be a matched-pair setting.

```
##
##  One Sample t-test
##
## data:  eyesdiff
## t = 0.19309, df = 30, p-value = 0.4241
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  -0.5277127        Inf
## sample estimates:
##  mean of x
## 0.06774194
```

From the data showed above, we get a 95% confidence interval from -0.5277 to Inf.

c.

This is an independent sample setting

```
##
##  Welch Two Sample t-test
##
## data:  eyesdiff by eye
## t = -0.55731, df = 22.911, p-value = 0.7086
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -1.684289       Inf
## sample estimates:
## mean in group 1 mean in group 2
##      -0.1055556       0.3076923
```

From the data above, we see that on symmetry does indeed depend on eye dominance.

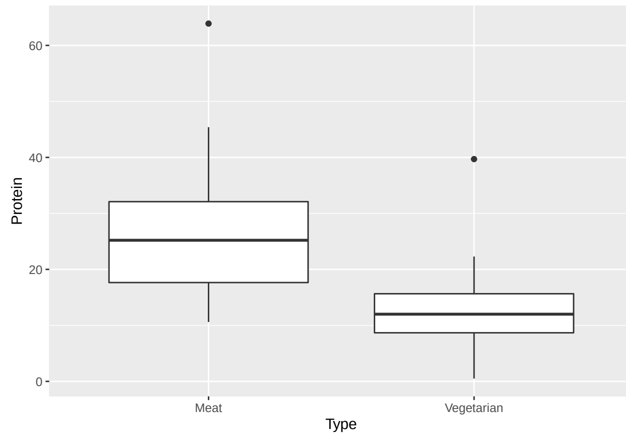
d.

```
##
##  Welch Two Sample t-test
##
## data:  eyesdiff by hand
## t = -0.7774, df = 7.6238, p-value = 0.7698
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -1.691085       Inf
## sample estimates:
## mean in group 1 mean in group 2
##     0.003703704     0.500000000
```

Likewise, we see a 95% confidence interval from -1.691 to inf. As such, we also see that symmetry does indeed depend on hand dexterity (left or right handed).

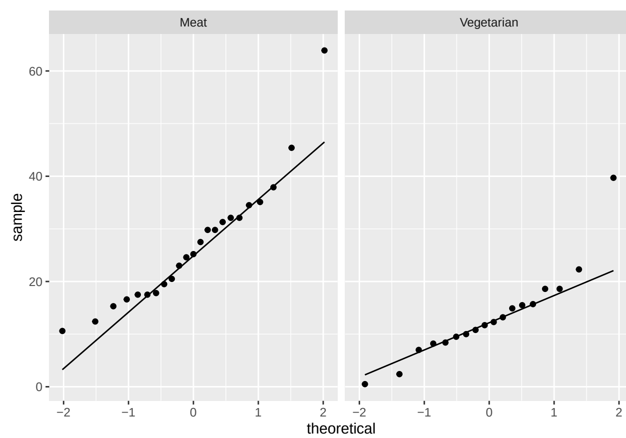
4

a.



b.

```
## # A tibble: 2 x 3
##   Type       `mean(Protein)` `sd(Protein)`
##   <chr>                <dbl>         <dbl>
## 1 Meat                  27.0         12.0
## 2 Vegetarian            13.3          8.59
```



From the graph, one can see that there are some clear outliers to the rough line prescribed by the graph. As such, a normal distribution may not be appropriate and normality may not be met.

c.

```
##
##  Welch Two Sample t-test
##
## data:  Protein by Type
## t = 4.2531, df = 38.769, p-value = 0.0001286
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   7.161098 20.154361
## sample estimates:
```

```
##       mean in group Meat mean in group Vegetarian
##                 26.95217                 13.29444
```

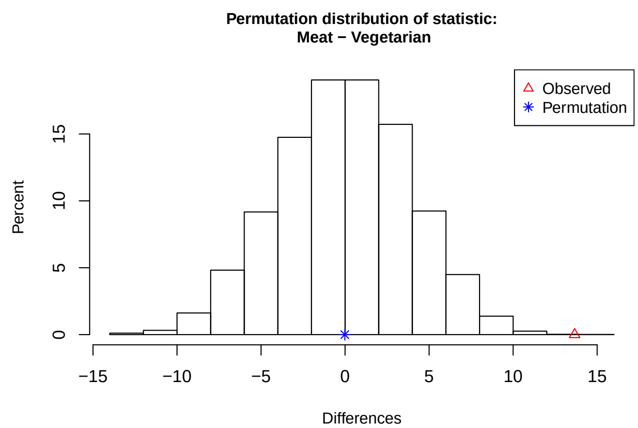
As we can clearly see from the data, the mean for meat protein is 26.95217 whereas vegetarian protein is only 13.2944. As such, we conclude that the means are not the same

d.

```
##
##  Welch Two Sample t-test
##
## data:  Protein by Type
## t = 5.7115, df = 35.712, p-value = 1.734e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   8.725291 18.337811
## sample estimates:
##       mean in group Meat mean in group Vegetarian
##                 25.27273                 11.74118
```

Even after getting rid of the outliers, we still come to the same conclusion as before.

e.




```
##
##  ** Permutation test **
##
##  Permutation test with alternative: two.sided
##  Observed statistic
##   Meat :  26.95217    Vegetarian :  13.29444
##  Observed difference: 13.65773
##
##  Mean of permutation distribution: -0.01624
##  Standard error of permutation distribution: 3.92014
##  P-value:  4e-04
##
##  *-------------*
```

Looking at the data we see that we can reach the same conclusion as above.

Section 7.1

7.4

$p_1 = 280$

$p_2 = 40$

$p_3 = 40$

$p_4 = 40$

7.11

a.

$(210 + 732 + 396 + 125 + 213 + 324)(0.35) = 700$

b.

$\frac{(732-700)^2}{700} = 1.46$

c.

$df = 6 - 1 = 5$

7.14

$\chi^2 = \frac{(Actual-Expected)^2}{Expected} = \frac{(66-39.67)^2}{39.67} + \frac{(39-39.67)^2}{39.67} + \frac{(14-39.67)^2}{39.67} = 34.0980$

Looking at our χ^2 value, we can say that there are proportions that aren't the $\frac{1}{3}$ outcome we expect, we find that the win count for rock and scissors don't match the one-third win rate

7.16

$df = 5 - 1 = 4$

Confidence Level: 5%

$\chi^2 = \frac{(Actual-Expected)^2}{Expected} = \frac{(391-410)^2}{410} + \frac{(257-280)^2}{280} + \frac{(156-130)^2}{130} + \frac{(89-70)^2}{70} + \frac{(107-110)^2}{110} = 13.2087$

p-value = 0.010

Summarizing our data, we can conclude that the data given by the article is not accurate.

7.26

$\chi^2 = \frac{(Actual-Expected)^2}{Expected} = \frac{(357.588-345)^2}{357.588} + \frac{(209.088-197)^2}{209.088} + \frac{(148.5-170)^2}{170} + \frac{(115.236-126)^2}{115.236} + \frac{(93.852-101)^2}{93.852} + \frac{(79.596-72)^2}{79.596} + \frac{(68.904-69)^2}{68.904} + \frac{(60.588-51)^2}{60.588} + \frac{(54.648-57)^2}{54.648} = 9.0876$

p-value = 0.3350

Summarizing out data, we see that our data is not consistent with Benford's Law.