Homework #08

Robert Campbell

12 Apr 2021

Chapter 08

Problem 14

98% Confidence interval: [-1.037, -0.556] P value is < 0.05, so the mean differs significantly from 0. This means that the two nurses' interpretations of the data differed significantly, effecting the values.

```
react <- ISwR::react
xbar <- mean(react)
s <- sd(react)
stderror <- s/sqrt(length(react))
tcritical <- -qt(0.01,(length(react)-1))
xbar - tcritical * stderror  # lower
```

```
## [1] -1.036587
```

```
xbar + tcritical * stderror  # upper
```

```
## [1] -0.5562269
```

```
t.test(react,conf.level = 0.98)
```

```
##
##  One Sample t-test
##
## data:  react
## t = -7.7512, df = 333, p-value = 1.115e-13
## alternative hypothesis: true mean is not equal to 0
## 98 percent confidence interval:
##  -1.0365875 -0.5562269
## sample estimates:
##  mean of x
## -0.7964072
```

Problem 16

Hypothesis: Children's estimate of an object's weight will deviate from 0, meaning they will make inaccurate estimates.

The only accurate estimates were for the 200g object, all others were significantly different.

```r
child_est<- fosdata::weight_estimate %>% filter(age!= 'adult')
t.test(child_est$mean100, mu=100)
```

```
##
##  One Sample t-test
##
## data:  child_est$mean100
## t = 9.003, df = 59, p-value = 1.13e-12
## alternative hypothesis: true mean is not equal to 100
## 95 percent confidence interval:
##  158.0713 191.2621
## sample estimates:
## mean of x
##  174.6667
```

```r
t.test(child_est$mean200, mu=200)
```

```
##
##  One Sample t-test
##
## data:  child_est$mean200
## t = 1.5712, df = 59, p-value = 0.1215
## alternative hypothesis: true mean is not equal to 200
## 95 percent confidence interval:
##  196.0338 232.9662
## sample estimates:
## mean of x
##     214.5
```

```r
t.test(child_est$mean300, mu=300)
```

```
##
##  One Sample t-test
##
## data:  child_est$mean300
## t = -5.1213, df = 59, p-value = 3.497e-06
## alternative hypothesis: true mean is not equal to 300
## 95 percent confidence interval:
##  235.0999 271.5668
## sample estimates:
## mean of x
##  253.3333
```

```r
t.test(child_est$mean400, mu=400)
```

```
##
##  One Sample t-test
##
## data:  child_est$mean400
## t = -6.6477, df = 59, p-value = 1.065e-08
## alternative hypothesis: true mean is not equal to 400
```

```
## 95 percent confidence interval:
##  339.0696 367.2637
## sample estimates:
## mean of x
##  353.1667
```

Problem 17

Population is Mexican-American adults in California, with a sample size of 102. P-value = (1.743 x 10^-4) < 0.05, so the blood pressure significantly differs from the normal systolic blood pressure of 120 mm Hg.

```
obesity<-ISwR::bp.obese
t.test(obesity$bp, mu=120)
```

```
##
##  One Sample t-test
##
## data:  obesity$bp
## t = 3.8986, df = 101, p-value = 0.0001743
## alternative hypothesis: true mean is not equal to 120
## 95 percent confidence interval:
##  123.4479 130.5914
## sample estimates:
## mean of x
##  127.0196
```

Problem 19

Natural null hypothesis is that the population's ratio of actual weight to ideal weight will not deviate from the mean of 1.

P-value is much less that 0.05, meaning that the population's mean ratio significantly differs from the null hypothesis.

```
obesity<-ISwR::bp.obese
t.test(obesity$obese, mu=1)
```

```
##
##  One Sample t-test
##
## data:  obesity$obese
## t = 12.262, df = 101, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 1
## 95 percent confidence interval:
##  1.262395 1.363684
## sample estimates:
## mean of x
##  1.313039
```

Problem 25

a. Pain felt by a laser shot on the hand will not significantly change based on the aesthetic quality of an observed painting.

b. There is no difference in appearance of objects between normal viewing and viewing by bending over and looking between your legs.

c. The time to urinate will differ based off of species.

Problem 26

a. The mean of the sample is greater than the null hypothesis mean in all measurements.
b. The quality of sleep did not see significant improvement. This can be seen by the P-value being greater than 0.05 (P = 0.27).

c. Reported sleep disturbance showed the most significance, having the smallest P-value of less than 0.01.

Problem 27

P-value = 0.6463 > 0.05, there is no significant difference from between male and female blood pressure in the sample group.
It would not be appropriate to generalize this to the general American population. It is a single ethnic group, in a single geographic region, in a "small town", and does not account for the diversity in each of those larger categories that make up the general American population.

```
obesity<-ISwR::bp.obese
t.test(bp ~ sex, data=obesity)
```

```
##
##  Welch Two Sample t-test
##
## data:  bp by sex
## t = 0.46033, df = 98.535, p-value = 0.6463
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -5.443341  8.731743
## sample estimates:
## mean in group 0 mean in group 1
##        127.9545        126.3103
```

Problem 28

There is significant evidence, with the p-value = 0.01308. With Fish Oil's mean reduction in diastolic blood pressure having a positive value while regular oil has a negative value showing that fish oil provides greater reduction in diastolic blood pressure on average.

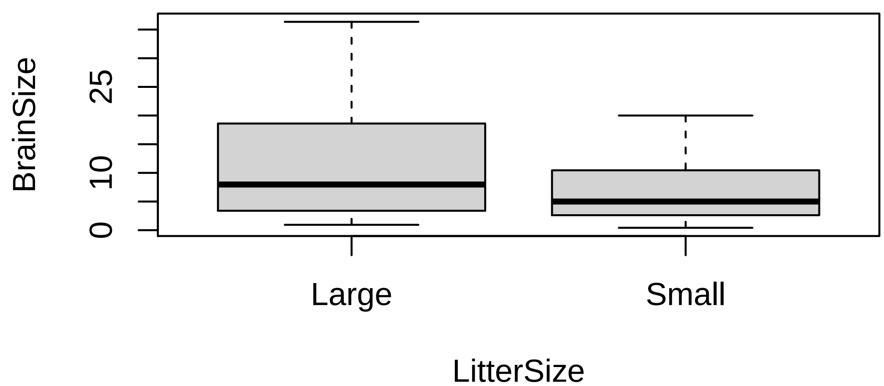
```
oil<-Sleuth3::ex0112
t.test(BP ~ Diet, data=oil)
```

```
##
##  Welch Two Sample t-test
##
## data:  BP by Diet
## t = 3.0621, df = 9.2643, p-value = 0.01308
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   2.039893 13.388678
## sample estimates:
##    mean in group FishOil mean in group RegularOil
##                 6.571429                -1.142857
```

Problem 29

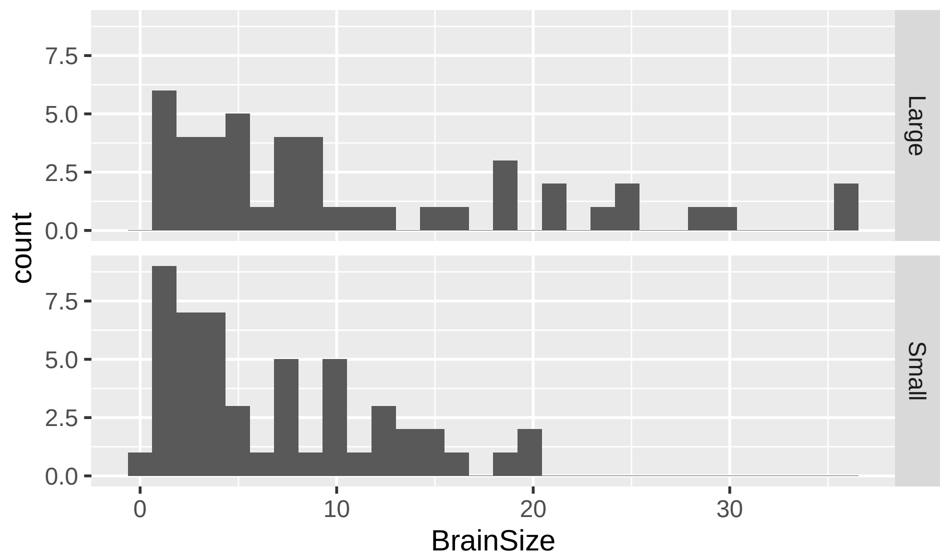
No, the initial boxplot looks lopsided and the initial histograms look more exponential than normal.
After calculating the logs, the data does appear more normal.
The difference in mean of the log of brainsize between large and small litter size is not significant as P-value= 0.05225 is not less than 0.05, though it is close.

```
size <- Sleuth3::ex0333
boxplot(BrainSize ~ LitterSize, data=size)
```

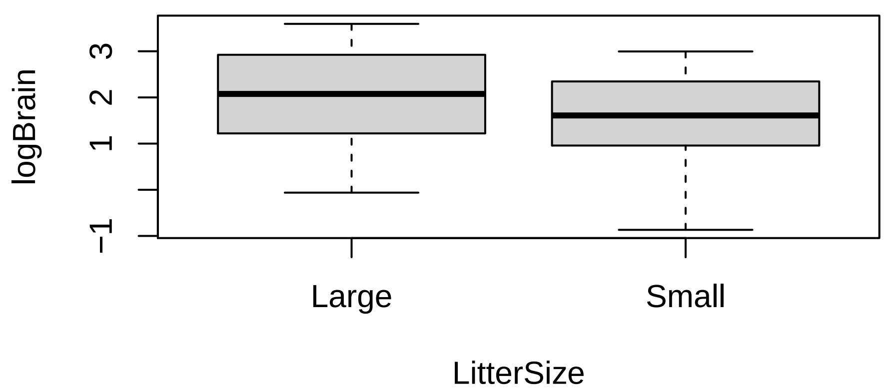



```
size %>% ggplot(aes(x=BrainSize)) + geom_histogram() + facet_grid(vars(LitterSize))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

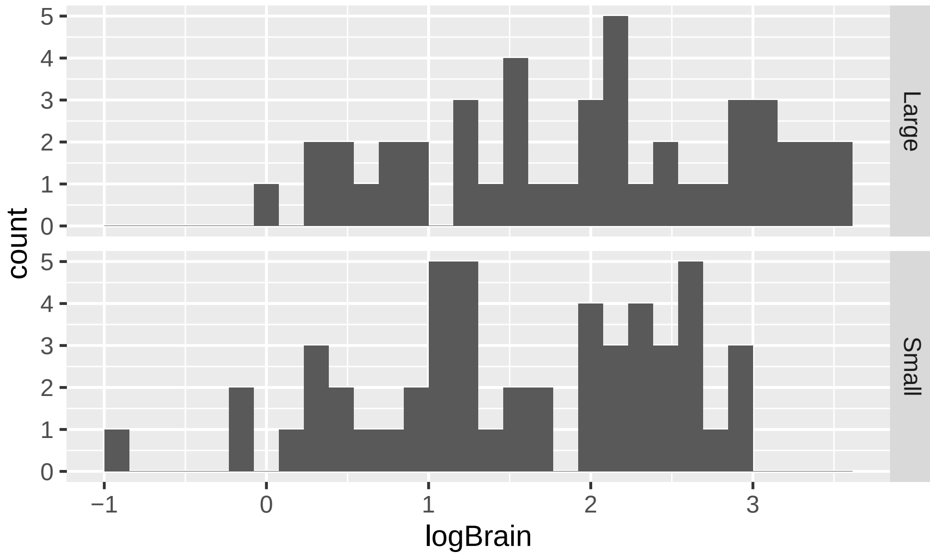



```
size<-mutate(size,logBrain=log(BrainSize))
boxplot(logBrain ~ LitterSize, data=size)
```




```
size %>% ggplot(aes(x=logBrain)) + geom_histogram() + facet_grid(vars(LitterSize))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



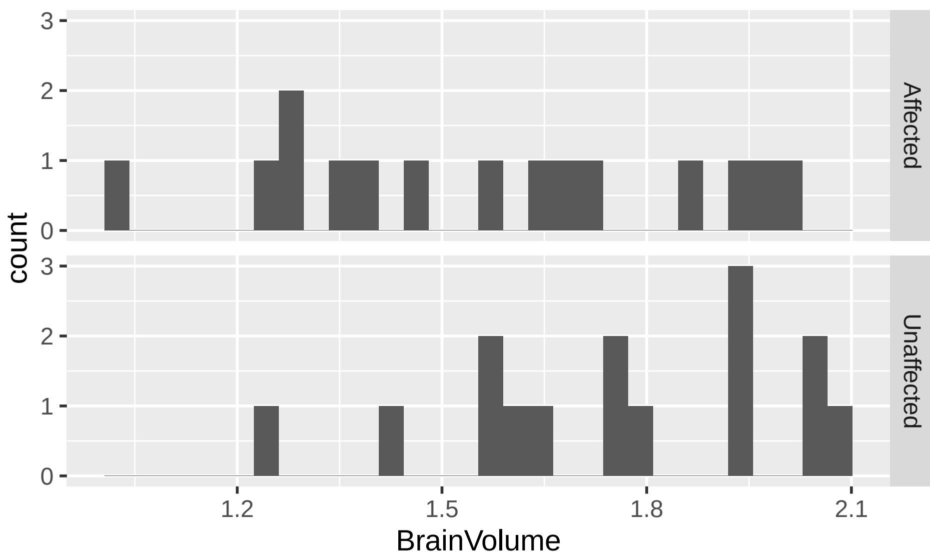

```
t.test(logBrain ~ LitterSize,data=size)
```

```
##
##  Welch Two Sample t-test
##
## data:  logBrain by LitterSize
## t = 1.9669, df = 90.684, p-value = 0.05225
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.003942989  0.797879510
## sample estimates:
## mean in group Large mean in group Small
##            1.949426            1.552458
```

Problem 30

Hypothesis: There is a significant difference in volume in the affected brain regions of schizophrenic patients. The sample data is inconclusive with P-Value = 0.05691 being greater than 0.05, but close.

```
vol<-Sleuth3::case0202
pvol<-vol %>% pivot_longer(col=Unaffected:Affected, names_to = "Diagnosis", values_to = "BrainVolume")
pvol %>% ggplot(aes(x=BrainVolume)) + geom_histogram() + facet_grid(vars(Diagnosis))
```




```
t.test(BrainVolume ~ Diagnosis, pvol)
```

```
##
##  Welch Two Sample t-test
##
## data:  BrainVolume by Diagnosis
## t = -1.9898, df = 26.775, p-value = 0.05691
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.403606427  0.006273094
## sample estimates:
##   mean in group Affected mean in group Unaffected
##                 1.560000                 1.758667
```

Problem 31

a. The tk-12 provides a significantly different temperature change from the control with a P-value of (5.609 x 10^-15), and provides a cooling afffect as all data values are negative while the control are positive.

```
moo<-fosdata::cows_small
t.test(moo$tk_12, moo$control, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  moo$tk_12 and moo$control
## t = -23.558, df = 18, p-value = 5.609e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.040784 -1.706584
## sample estimates:
## mean of the differences
##               -1.873684
```

b. The tk-0.75 provides a significantly different temperature change from the control with a P-value of (3.747 x 10^-14), and provides a cooling effect as all data values are negative while the control are positive.

```
t.test(moo$tk_0_75, moo$control, paired = TRUE)
```

```
##
##  Paired t-test
##
## data:  moo$tk_0_75 and moo$control
## t = -21.122, df = 18, p-value = 3.747e-14
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.917316 -1.570404
## sample estimates:
## mean of the differences
##                -1.74386
```

Problem 32

There is not a significant difference in task completion time based on gender as the P-value: 0.2033 much larger than the alpha: 0.01

```
child<-fosdata::child_tasks
t.test(day_night_completion_time_secs ~ gender, data=child)
```

```
##
##  Welch Two Sample t-test
##
## data:  day_night_completion_time_secs by gender
## t = -1.2891, df = 49.808, p-value = 0.2033
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.270679  0.931872
## sample estimates:
## mean in group Female   mean in group Male
##             26.44026             28.10967
```

Problem 21

a. The mean of 4 was within the confidence interval. (May not show that way after the knit)

```
x<-rnorm(10,mean=4, sd=1)
t.test(x,conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  x
## t = 17.069, df = 9, p-value = 3.655e-08
## alternative hypothesis: true mean is not equal to 0
```

```
## 95 percent confidence interval:
##  3.417182 4.461324
## sample estimates:
## mean of x
##  3.939253
```

b.

```
z<-replicate(10000, {x<-rnorm(10,mean=4, sd=1);
  y<-t.test(x,conf.level = 0.95)$conf.int;
  (4 > y[1]) && (4<y[2])})
mean(z)
```

```
## [1] 0.9493
```

Problem 23

a. The mean of 4 did appear in the confidence interval. (This may not match the R output after knitting)

```
x<-rexp(10,0.25)
t.test(x,conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  x
## t = 4.2843, df = 9, p-value = 0.002037
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  0.8181467 2.6486805
## sample estimates:
## mean of x
##  1.733414
```

b. The result is not .95 because the data is heavily skewed left.

```
z<-replicate(10000, {x<-rexp(10,0.25);
  y<-t.test(x,conf.level = 0.95)$conf.int;
  (4 > y[1]) && (4<y[2])})
mean(z)
```

```
## [1] 0.8998
```

c. This is closer to .95 because of the larger sample in each experiment iteration, which brings the sample mean closer to the true mean.

```
z<-replicate(10000, {x<-rexp(100,0.25);
  y<-t.test(x,conf.level = 0.95)$conf.int;
  (4 > y[1]) && (4<y[2])})
mean(z)
```

```
## [1] 0.9424
```

Problem 24

```r
pval<- replicate(10000,{x<-rnorm(20,0,1);
  y<-convolve(x, c(1, 1), type = "filter");
  t.test(y, mu=0)$p.value})
mean(pval>0.05)
```

```
## [1] 0.8319
```