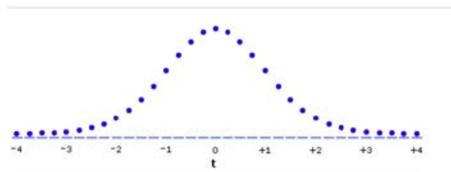
Minoo Ahmadi, Homework 1

Part I

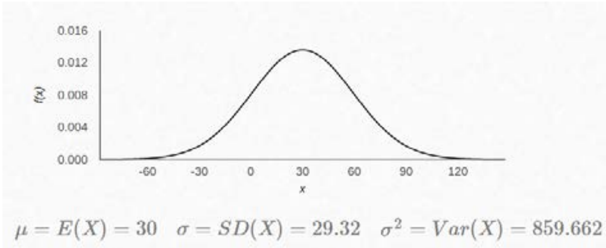
1. A mean of 35.17 (SD = 29.32) was found for the population. A median of 27.5 indicates a right-skewness in the data, which is probably due to the maximum value of 150 for the subject with ID = 27. The variance is 860.15.
2.
   a. H_0: $\mu = 30$, H_1: $\mu \neq 30$
   b. T-distribution: t(20)



   c. Using an online t-table, the critical t-value for a two-sided test at alpha level of .05 is 2.086.
   d. t(20) = 1.28, which is smaller than (doesn't exceed) the critical t-value.
   $$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} = \frac{39.04 - 30}{\frac{32.15}{\sqrt{21}}} = 1.28$$
   e. Using an online t-table: p>.05.
   f. $\%95CI = 39.04 \pm 2.086 \times \frac{32.15}{\sqrt{21}} = [24.41\ 53.68]$. We can be 95% confident that the population mean commute time is between 24.41 and 53.68 minutes. Note that this interval includes the mean of 30 and that's why we fail to reject the null.
   g. We failed to reject the null hypothesis.
   h. A two-sample t-test was performed to compare the sample mean commuting time in the sample to a hypothesized population mean of 30. Using an alpha level of 0.05, the sample mean of 39.04 (SD = 32.15) was found not to be statistically significantly different from this value (t(20) = 1.28, p > 0.05 (two-tailed)), suggesting that the mean commuting time in the sample is not significantly different from 30 minutes.

```
        One Sample t-test
data:  samp.data$commute
t = 1.2894, df = 20, p-value = 0.212
alternative hypothesis: true mean is not equal to 30
95 percent confidence interval:
 24.41011 53.68513
sample estimates:
mean of x
 39.04762
```

3. Since we know the population variance, we'll use a Z test. Here is the z distribution:



The z statistic:

$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{39.04 - 30}{\frac{29.32}{\sqrt{21}}} = 1.41$$

It’s less than 1.96 (2SD from the mean), so we fail to reject the null. We get the same result by looking at a Z table. The value we get for a z-score of 1.41 is .9207 which falls within the 95% interval.

%95CI = [26.50 51.59]

One-sample z-Test

data: samp.data$commute
z = 1.4137, p-value = 0.1575
alternative hypothesis: true mean is not equal to 30
95 percent confidence interval:
26.50391 51.59133
sample estimates:
mean of x
39.04762

Part II

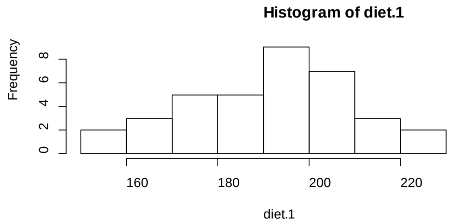
1. For diet 1:

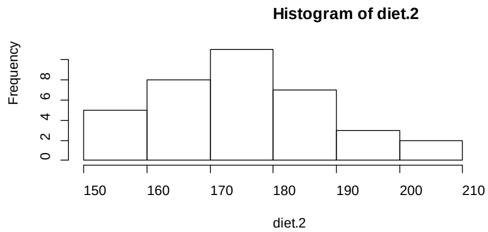
```
      median          mean       SE.mean CI.mean.0.95           var       std.dev 
194.00000000 193.00000000    2.90729788   5.90212847 304.28571429   17.44378727 
    coef.var 
  0.09038232 
```

For diet 2:

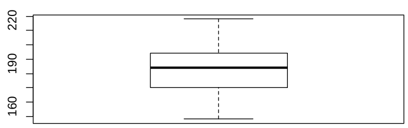
```
      median          mean       SE.mean CI.mean.0.95           var       std.dev 
174.50000000 176.00000000    2.31317826   4.69600152 192.62857143   13.87906954 
    coef.var 
  0.07885835 
```

Mean is slightly larger than median for diet 1 which indicates a potential left skewness while its slightly less than the median for the diet 2 which indicates a potential right skewness. The histograms show such trends as well. But overall they look normally distributed. No outliers observed for either of the diets.

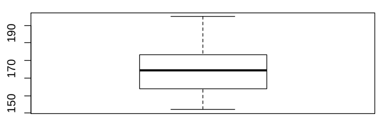




Boxplot for diet 1 (no outliers):



Boxplot for diet 2 (no outliers):



2. H0: $\mu_1 = \mu_2$, H1: $\mu_1 \neq \mu_2$
   An independent groups t-test was performed comparing the mean blood pressure for the diet 1 (M = 193, SD = 17.44) with that for the diet 2 (M = 176, SD = 13.87). Using an alpha level of 0.05, this test was found to be statistically significant, $t(66.63) = 4.57$, $p < 0.01$, indicating that diet 1 and diet 2 have different blood pressure outcomes and, to be more specific, diet 2 is associated with lower blood pressures.

```
        Welch Two Sample t-test
data:  hyptension by diet
t = 4.5757, df = 66.636, p-value = 2.125e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  9.58356 24.41644
sample estimates:
mean in group 1 mean in group 2 
            193             176 
```

3. This has already been answered above.