HW4 KEY

42 points total, 2 points per problem part unless otherwise noted.

Q1 Rent (Sample Size)

1A. Empirical Rule: s = (3200-1600)/6 = 266.6666667.

1B. n = 46 gives 95%ME = 79.19

```
alpha <- 0.05
sd <- 266.67
n <- seq(40, 50, 1)
ME <- qt(1-(alpha/2), df = n-1)*sd/sqrt(n)
out <- data.frame(n, ME)
```

Q2 Zinc (Power)

2A. Power = 0.703 with n = 120.

2B. Higher

2C. Higher

2D. Lower

2E. Higher

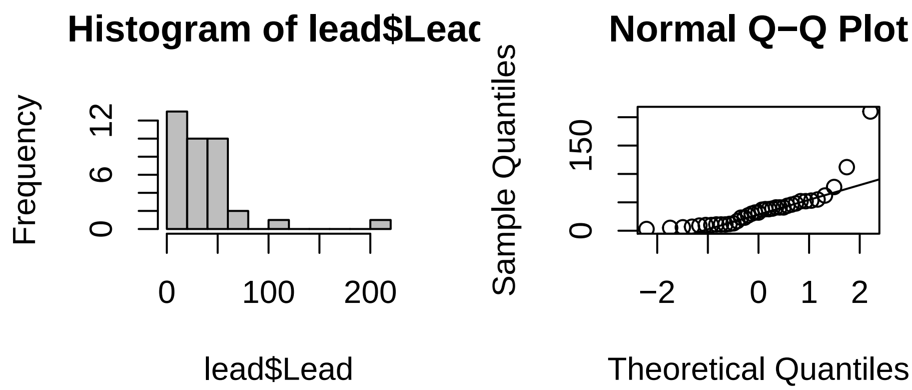
2F. n = 156 to achieve power of 0.80.

```
#A
power.t.test(n = 120, delta = 0.4, sd = 2, sig.level = 0.05,
             type = "one.sample", alternative = "one.sided")
#B
power.t.test(n = 120, delta = 0.4, sd = 1, sig.level = 0.05,
             type = "one.sample", alternative = "one.sided")
#C
power.t.test(n = 150, delta = 0.4, sd = 2, sig.level = 0.05,
             type = "one.sample", alternative = "one.sided")
#D
power.t.test(n = 120, delta = 0.4, sd = 2, sig.level = 0.01,
             type = "one.sample", alternative = "one.sided")
#E
power.t.test(n = 120, delta = 1, sd = 2, sig.level = 0.05,
             type = "one.sample", alternative = "one.sided")
#F
power.t.test(power = 0.8, delta = 0.4, sd = 2, sig.level = 0.05,
             type = "one.sample", alternative = "one.sided")
```

Q3 Lead (Inference for single mean/median)

3A. **(4 pts)** Lack of normality can be seen from histogram (skewed right), qqplot (shows curvature) and tests of normality (small p-value for SW test indicates evidence against normality). Diagnostics agree.

```
lead <- read.csv("C:/hess/STAT511_FA11/ASCII-comma/CH05/ex5-27.TXT", quote = " ' "  )
par(mfrow = c(1, 2))
hist(lead$Lead, breaks = 10, col = "grey")
qqnorm(lead$Lead); qqline(lead$Lead)
```




```
shapiro.test(lead$Lead)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  lead$Lead
## W = 0.69693, p-value = 1.928e-07
```

3B. mean = 37.24, median = 32

3C. Sign Test p-value = 0.6177
Fail to Reject H0, we cannot conclude that the population median is different from 30.

3D. 95% CI for the median: (17, 41)

```
#3C,D,E
library(BSDA)
SIGN.test(lead$Lead, md = 30)
ttest <- t.test(lead$Lead)
```

3E. 95% CI = ((24.87, 49.62))

3F. Results will vary.

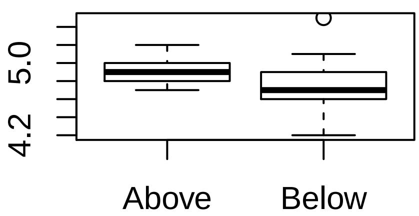
```
mean.fun <- function (d, i)
{ m <- mean(d[i])
n <- length(i)
v <- (n-1)*var(d[i])/n^2
c(m, v)
}
library(boot)
results <- boot(data = lead$Lead, mean.fun, R = 1000)
boot.ci(results, type = "all")
```

3G. Mean

Q4 Oxygen (2 sample t-test)

4A. Boxplots

```
Oxygen <- read.csv("C:/hess/STAT511_FA11/ASCII-comma/CH06/ex6-6.txt", quote = " ' ")
par(mfrow=c(1,1))
boxplot(Oxygen)
```



4B. Above mean = 4.92, sd = 0.157
Below mean = 4.74, sd = 0.32

4C. Since 0.320/0.157 > 2 (just barely!), the Welch-Satterthwaite test is preferred.

```
#4D,E
TestOut <- with(t.test(Above, Below, var.equal = FALSE), data = Oxygen)
```

4C. 95% CI = ((-0.012, 0.372))
Since the CI includes zero, we cannot conclude there is a difference between the means.

4E. p-value = 0.064. Fail to Reject H0. Cannot conclude there is a difference between the means.