1) µ0 = 15 mg/day

Conjectured true population mean, µA = 17 mg/day

Hypothesis:

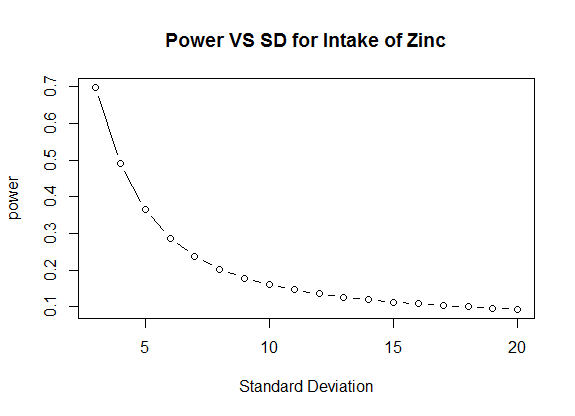
H0 : µ<= µ0 : µ<= 15

HA : µ> µ0 : µ> 15

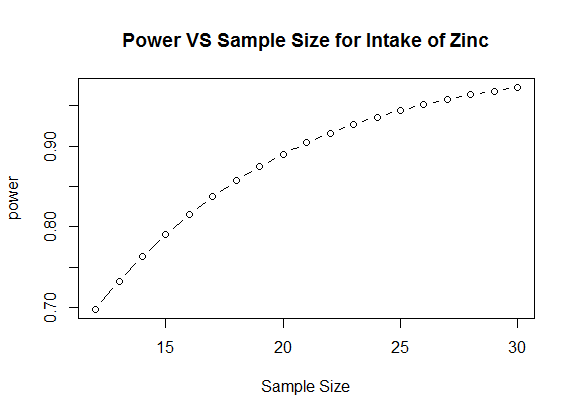
n = 12, α = 0.05, conjectured s.d σ = 3 mg/day

A) power = 0.6981908 [See appendix for R code and Calculation]

B) If the sample deviation was larger, the power would be lower than that of the power that was calculated in part A. Following graph shows the Power ~ SD relation for the data.



C) If the sample size was larger, power would increase. The following graph shows the Power ~ Sample size relation for Zinc data.

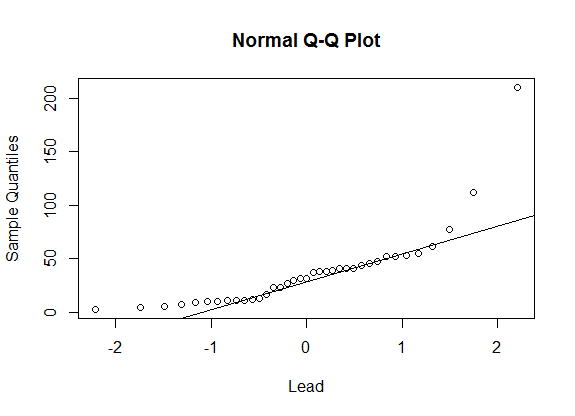
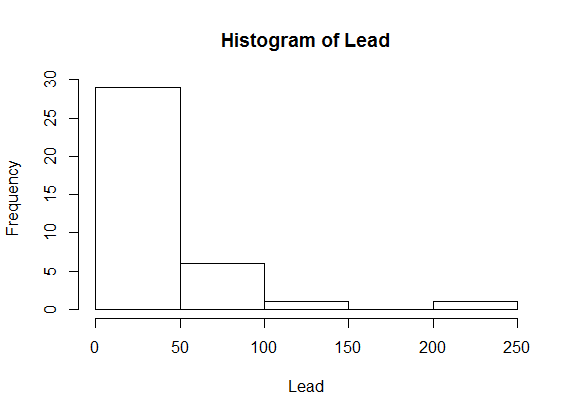


D) If α = 0.10, power = 0.8260609 [See Appendix for calculation], which is higher than the power calculated in part A.

E) For µA = 16 mg/day, power = 0.2874441 [See Appendix for calculation], which is lower than the power calculated in part A.

F) For Power = 0.9, sample size = 21 (rounded up) [See Appendix for calculation].

2) A)



Shapiro-Wilk normality test p-value = 1.928e-07 < 0.05. [See apendix]  
So, the null hypothesis (lead sample data is normally distributed) can be rejected.

The result of the shapiro-wilk test matches the histogram and normal Q-Q plot. From the Histrogram we can see that the lead sample data is skewed right, and from the normal Q-Q plot we see that the plot is not a straight line.  
  
B) Mean = 37.24324

Median = 32

C) H0 : M = 30

HA : M ≠ 30

From the sign test, we obtain,  
s (#of values > 30) = 20

p-value = 0.6177

As P value > α = 0.05, we cannot reject H0 .

We do not have enough evidence to reject the population median = 30 with 95% confidence.

D)

95% confidence interval (using Upper Archieved CI) is (17.0000,41)

E)

H0 : µ= µ0 : µ= 30

HA : µ≠ µ0 : µ≠ 30

p-value = 0.2431 > α = 0.05. So we cannot reject H0.

We do not have enough evidence to reject that the population mean = 30 with 95% confidence.

F)

95 percent confidence interval: (24.86550,49.62099)

G) Studentized confidence interval mean= (27.46,57.19)

H) Assuming the cumulative lead exposure is of interest, the mean would be of more interest.

**APPENDIX**

**#QUESTION 1**

> #1A

> power.t.test(n=12,delta=2,sd=3,sig.level = 0.05,type = "one.sample", alternative = "one.sided")

One-sample t test power calculation

n = 12

delta = 2

sd = 3

sig.level = 0.05

power = 0.6981908

alternative = one.sided

>

> #1B

> testSD<-seq(3,20,1)

> powerVal1B\_DiffSD <- power.t.test(n=12,delta=2,sd=testSD,sig.level = 0.05,type = "one.sample", alternative = "one.sided")

> plot(powerVal1B\_DiffSD$power ~ testSD, type = "b", xlab = "Standard Deviation", ylab = "power", main = "Power VS SD for Intake of Zinc")

>

> #1C

> testSampleSize <- seq(12,30,1)

> powerVal1C\_Diffn <- power.t.test(n=testSampleSize,delta=2,sd=3,sig.level = 0.05,type = "one.sample", alternative = "one.sided")

> plot(powerVal1C\_Diffn$power ~ testSampleSize, type = "b", xlab = "Sample Size", ylab = "power", main = "Power VS Sample Size for Intake of Zinc")

>

> #1D

> power.t.test(n=12,delta=2,sd=3,sig.level = 0.10,type = "one.sample", alternative = "one.sided")

One-sample t test power calculation

n = 12

delta = 2

sd = 3

sig.level = 0.1

power = 0.8260609

alternative = one.sided

>

> #1E

> power.t.test(n=12,delta=1,sd=3,sig.level = 0.05,type = "one.sample", alternative = "one.sided")

One-sample t test power calculation

n = 12

delta = 1

sd = 3

sig.level = 0.05

power = 0.2874441

alternative = one.sided

>

> #1F

> power.t.test(delta=2,sd=3,p=0.9,sig.level = 0.05,type = "one.sample", alternative = "one.sided")

One-sample t test power calculation

n = 20.69914

delta = 2

sd = 3

sig.level = 0.05

power = 0.9

alternative = one.sided

#QUESTION2

> hist(DataHW4\_2$X.Lead.,xlab = "Lead", main = "Histogram of Lead")

> qqnorm(DataHW4\_2$X.Lead.,xlab = "Lead")

> qqline(DataHW4\_2$X.Lead., )

> shapiro.test(DataHW4\_2$X.Lead.)

Shapiro-Wilk normality test

data: DataHW4\_2$X.Lead.

W = 0.69693, p-value = 1.928e-07

|  |
| --- |
| > #B  > mean(DataHW4\_2$X.Lead.)  [1] 37.24324  > median(DataHW4\_2$X.Lead.)  [1] 32  >  > #C  > hist(DataHW4\_2$X.Lead.,xlab = "Lead", main = "Histogram of Lead")  > summary(DataHW4\_2$X.Lead.)  Min. 1st Qu. Median Mean 3rd Qu. Max.  3.00 11.00 32.00 37.24 46.00 210.00  > sort(DataHW4\_2$X.Lead.)  [1] 3 5 6 7 9 10 10 11 11 11 12 13 17 23 23 27 30 32 32  [20] 37 38 38 39 41 41 41 44 46 48 52 52 53 55 62 77 112 210  > library(BSDA)  > SIGN.test(DataHW4\_2$X.Lead., md=30)  One-sample Sign-Test  data: DataHW4\_2$X.Lead.  s = 20, p-value = 0.6177  alternative hypothesis: true median is not equal to 30  95 percent confidence interval:  17.34363 41.00000  sample estimates:  median of x  32  Conf.Level L.E.pt U.E.pt  Lower Achieved CI 0.9011 23.0000 41  Interpolated CI 0.9500 17.3436 41  Upper Achieved CI 0.9530 17.0000 41 |
|  |
| |  | | --- | | > t.test(DataHW4\_2$X.Lead. , mu=30)  One Sample t-test  data: DataHW4\_2$X.Lead.  t = 1.1868, df = 36, p-value = 0.2431  alternative hypothesis: true mean is not equal to 30  95 percent confidence interval:  24.86550 49.62099  sample estimates:  mean of x  37.24324 | |

**2G**

> mean.fun <- function(d,i)

+ {

+ m <- mean(d[i])

+ n <- length(i)

+ v <- (n-1)\*var(d[i])/n^2

+ c(m,v)

+ }

> set.seed(7255)

> resultsHW4\_2F <- boot(data=DataHW4\_2$X.Lead.,mean.fun, R=1000)

> boot.ci(resultsHW4\_2F,type="all")

BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS

Based on 1000 bootstrap replicates

CALL :

boot.ci(boot.out = resultsHW4\_2F, type = "all")

Intervals :

Level Normal Basic Studentized

95% (25.31, 48.78 ) (23.66, 47.92 ) (27.46, 57.19 )

Level Percentile BCa

95% (26.57, 50.83 ) (28.02, 53.82 )

Calculations and Intervals on Original Scale

Some BCa intervals may be unstable