HW3 KEY

NOTE: These solutions are for personal use and should not be shared with other students.

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

Diet (Q1 - Q8)

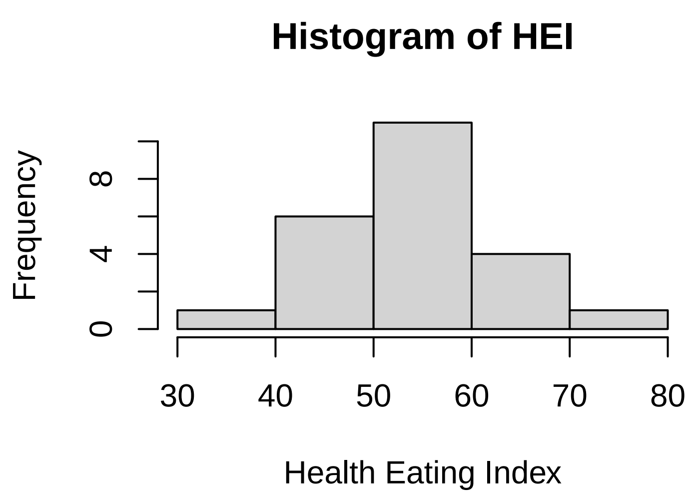
Q1

The population is all women receiving SNAP benefits (food stamps) and living in Denver, CO.

OK to claim a larger population *if* it is acknowledged that this requires assuming the Denver population are representative of a larger population. Also OK to narrow the population, e.g. “women receiving SNAP benefits, living in Denver CO, who are willing to answer surveys”.

Q2

```r
Diet <- read.csv("Diet.csv")
hist(Diet$HEI, main = "Histogram of HEI",xlab="Health Eating Index")
```



Q3

Mean:

```r
HEI_mean <- mean(Diet$HEI)
HEI_mean
```

```
## [1] 54.78261
```

Standard Deviation:

```r
HEI_sd <- sd(Diet$HEI)
HEI_sd
```

```
## [1] 8.612657
```

Q4

95% CI: (51.058, 58.507)

```r
tcrit <- qt(0.975,df=22)
HEI_lower <- HEI_mean - tcrit*HEI_sd/sqrt(23)
HEI_upper <- HEI_mean + tcrit*HEI_sd/sqrt(23)
c(HEI_lower,HEI_upper)
```

```
## [1] 51.05822 58.50700
```

```r
#Or, use t.test()
t.test(Diet$HEI)
```

```
##
##  One Sample t-test
##
## data:  Diet$HEI
## t = 30.505, df = 22, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  51.05822 58.50700
## sample estimates:
## mean of x
##  54.78261
```

Q5

Precise language may vary. Any amount of rounding is fine. Acceptable answers:

- We can be 95% confident that the true population mean HEI score (or μ_{HEI}) is between 51.058 and 58.507.
- The range of plausible values for μ_{HEI} is 51.058 to 58.507
- The range of values for μ_{HEI} that would be reasonably likely to produce data similar to ours is 51.058 to 58.507.
- μ_{HEI} should be somewhere between 51.058 to 58.507, because it was make using a method that would successfully cover μ_{HEI} 95% of the time, under repeated sampling.

Q6

The CI is valid because the data appear to be approximately normally distributed (even though the sample size is "moderate").

Q7 (4 pts)

Reject H_0 because 60 is not included in the confidence interval.

Can also say:

- We conclude the population mean is different from 60.
- We have evidence the population mean is different from 60.

For full points, the reason for this conclusion must be based on the confidence interval.

Q8 (4 pts)

p-value = 0.0082

Some acceptable interpretations:

- The probability of obtaining a test statistic at least as large our ours, if H_0 were true, is 0.0082.
- Results at least as extreme as ours would occur 0.82% of the time if H_0 were true.
- The probability of getting data that disagree with H_0 by at least as much as these do is less than $\alpha = 0.05$, so we reject H_0.
- P-value is less than 0.05 and rejects the null hypothesis, because there is a less than 5% chance of getting p<0.05 when the null hypothesis is true.

Parts of the above interpretations can be combined. We are looking for an answer that contains a definition of the p-value in it. For full credit, answer must contain some reference to "assuming the null hypothesis is true", or "when the null hypothesis is true" or "using Type I error rate of 0.05".

```r
HEI_test_stat <- (HEI_mean-60)/(HEI_sd/sqrt(23))
HEI_pvalue <- pt(abs(HEI_test_stat),df=22,lower.tail=F)*2
HEI_pvalue
```

```
## [1] 0.008206317
```

```r
#Or, use t.test()
t.test(Diet$HEI, mu = 60)
```

```
##
##  One Sample t-test
##
## data:  Diet$HEI
## t = -2.9052, df = 22, p-value = 0.008206
## alternative hypothesis: true mean is not equal to 60
## 95 percent confidence interval:
##  51.05822 58.50700
## sample estimates:
## mean of x
##  54.78261
```

CI (Q9 - Q11)

Q9

Decrease

Q10

Increase

Q11

Decrease

- One way to explain Q9-Q11 is to use the formula of ME or the formula of the width which is $2 \times ME = 2 \times t_{\alpha/2,df} \times \frac{s}{\sqrt(n)}$. The changes in sample size, standard deviation and confidence level can be clearly seen in the formula. It is also fine to give an intuitive explanation.
- Q9: The estimation is more precise as the sample size increases and thus obtaining a narrower CI.
- Q10: A larger standard deviation gives us a wider CI.
- Q11: We have less confidence to believe that a narrow CI contains the true parameter. Or, to get a CI with a lower rate of success at covering the true parameter, we make it cover a narrower range of values.

Grading note: Full credit for some reasonable explanation.

Oxygen (Q12 - Q15)

Q12 (5 pts)

```
ybar = 4.62
s = 0.58
n = 10
mu0 = 5
RR = qt(0.975, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
```

```
##                name value
## 1 Rejection Region  2.26
## 2   Test Statistic -2.07
```

Q12A (2 pts)

Reject H_0 if |t| > RR = 2.26.

Q12B (2 pts)

t = -2.07.

Q12C (1 pt)

Fail to reject H_0.

Interpretation is not necessary for credit, but it is correct to also say:

- We cannot conclude the population mean is different from 5.
- We do not have sufficient evidence that the population mean is different from 5.
- Data like ours would not be unlikely if the population mean was 5.

Q13 (5 pts)

```
RR = qt(0.05, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
```

```
##                name value
## 1 Rejection Region -1.83
## 2   Test Statistic -2.07
```

Q13A (2 pts)

Reject H_0 if t < RR = -1.83.

Q13B (2 pts)

t = -2.07.

Q13C (1 pt)

Reject H_0.

Interpretation is not necessary for credit, but it is correct to also say:

- We conclude the population mean is less than 5.
- We have evidence that the population mean is less than 5.
- Data like ours would be unlikely to occur if the population mean was greater than or equal to 15.

Q14 (5 pts)

```
n = 40
RR = qt(0.975, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
```

```
##                name value
## 1 Rejection Region  2.02
## 2   Test Statistic -4.14
```

Q14A (2 pts)

Reject H_0 if $|t| > RR = 2.02$.

Q14B (2 pts)

$t = -4.14$.

Q14C (1 pt)

Reject H_0.

Q15

Many acceptable answers, such as:

- Increased sample size makes it easier/more likely to reject H_0.
- Increased sample size is associated with increased power.
- Increased sample size tends to make test statistics larger and p-values smaller.

Technical note: these three statements are true *when H_0 is false.* In this problem, we increased sample size and assumed our sample mean was unchanged. But, if H_0 was true, increasing sample size would tend to move our sample mean closer to the null value.

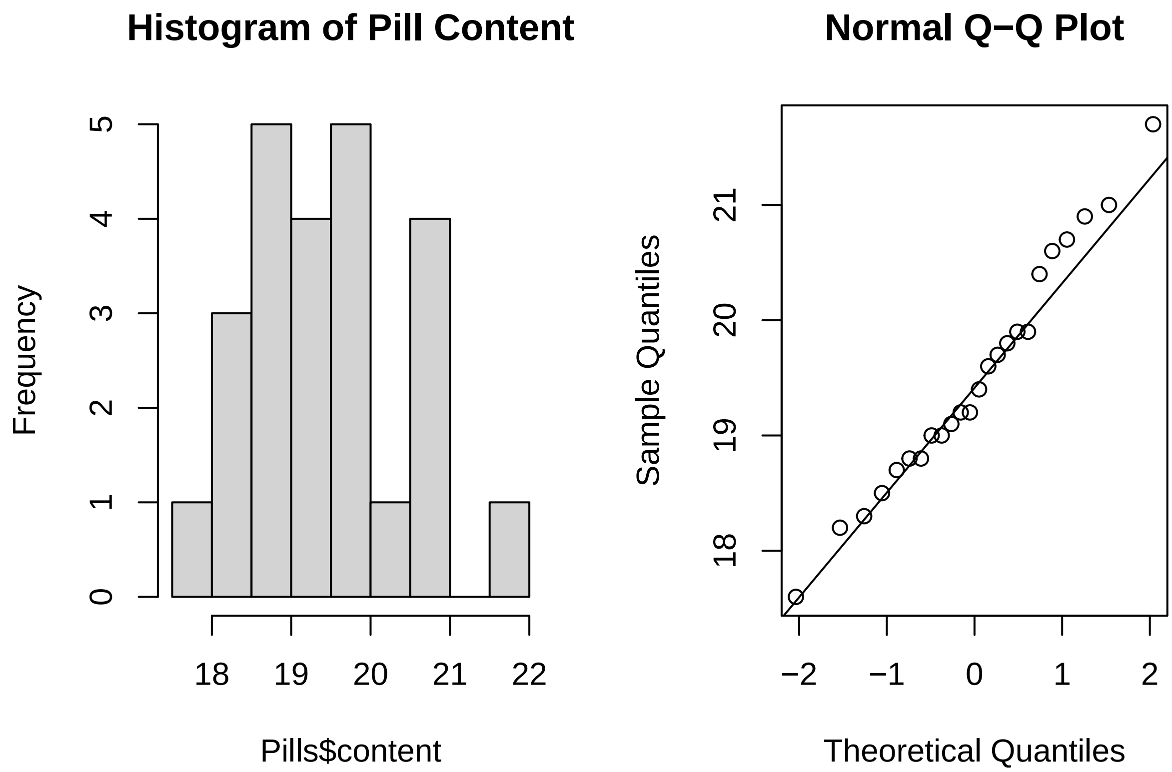
In other words, when H_0 is true, the probability of rejecting H_0 is α (typically 0.05), *regardless of the sample size.*

Pills (Q16 - Q19)

Q16

The histogram is (approximately) bell shaped and the qqplot is (approximately) linear, which support that the data came from a normal distribution, or a distribution that is not substantially non-normal. It is OK to note that these data show a little bit of right-skew, but this should not be used to make a strong conclusion of "therefore the data did not come from a normal distribution". With a sample size so small, it is not reasonable to expect data to form a perfect normal distribution; what we are looking for are clear signs of non-normality.

```
Pills <- read.csv("Pills.csv")
par(mfrow = c(1, 2))
hist(Pills$content, main = "Histogram of Pill Content")
qqnorm(Pills$content);qqline(Pills$content)
```



Q17

```
Q17Out <- t.test(Pills$content)
Pills_mean <- as.numeric(Q17Out$estimate)
Pills_CI <- Q17Out$conf.int
Pills_lower <- Pills_CI[1]
Pills_upper <- Pills_CI[2]
```

mean = 19.5
95% CI = (19.08, 19.92)

```
Q17Out
```

```
##
##  One Sample t-test
##
## data:  Pills$content
## t = 95.158, df = 23, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  19.07609 19.92391
## sample estimates:
## mean of x
##      19.5
```

Q18, Q19 Grading notes:
- Hypotheses (2 pts), test statistic (1 pt), p-value (1 pt), conclusion in context (2 pts).
- When stating hypotheses, need to specify population parameter (in this case mu = μ = population mean) for full credit.

Q18 (6 pts)

```
Q18Out <- t.test(Pills$content, mu = 20)
```

Q18A (2 pts)

H_0: $\mu = 20$ vs H_A: $\mu \neq 20$

Q18B (2 pts)

TS: t = -2.44
p-value = 0.022811

Q18C (2 pts)

(Since p-value < 0.05, we reject H0.)

Conclusion in context:

- We conclude that the (population) mean amount is different from 20mg.

- We have evidence that the (population) mean amount is different from 20mg.
- Data like ours would be unlikely to occur if the (population) mean amount was equal to 20mg.

```
Q18Out
```

```
##
##  One Sample t-test
##
## data:  Pills$content
## t = -2.44, df = 23, p-value = 0.02281
## alternative hypothesis: true mean is not equal to 20
## 95 percent confidence interval:
##  19.07609 19.92391
## sample estimates:
## mean of x
##      19.5
```

Q19 (6 pts)

```
Q19Out <- t.test(Pills$content, mu = 20, alternative = "less")
```

Q19A (2 pts)

H_0: $\mu \geq 20$ vs H_A: $\mu < 20$

Q19B (2 pts)

TS: t = -2.44
p-value = 0.0114055

Q19C (2 pts)

(Since p-value < 0.05, we reject H0.)

Conclusion in context:

- We conclude that the (population) mean amount is less than 20mg.
- We have evidence that the (population) mean amount is less 20mg.
- Data like ours would be unlikely to occur if the (population) mean amount was greater than or equal to 20mg.

Side note: using this kind of statistical test to decide whether to destroy product would be unwise in practice, because if H_0 is even slightly false, it will be rejected given a large enough sample size. Since it is not reasonable to expect infinite precision in a manufacturing process, a realistic testing method would be performed using tolerance bounds, such as 20mg $\pm$ δ, where δ is some acceptable deviation. We could then reject the batch if the mean was significantly less than 20mm - δ or significantly greater than 20mm + δ.

Or, if we wanted a stricter method on quality control, we would only accept the batch if the mean was significantly **greater than 20mm - δ and significantly less than 20mm + δ.** *This avoids treating a failure to reject the null as evidence that the null is true, a bad practice particularly when sample sizes are small and Type II errors are common.*

```
Q19Out
```

```
##
##  One Sample t-test
##
## data:  Pills$content
## t = -2.44, df = 23, p-value = 0.01141
## alternative hypothesis: true mean is less than 20
## 95 percent confidence interval:
##      -Inf 19.85121
## sample estimates:
## mean of x
##      19.5
```

CUE (Q20 - Q21)

Q20

From the Empirical Rule, about 99.7% of values in a normal distribution will fall within 3 standard deviations of the mean. This means there should be roughly 6 standard deviations from the left endpoint to the right endpoint of the interval that contains 99.7% of CUE values.

s = (0.90 - 0.36)/6 = 0.09

```
sd <- (0.90 - 0.36)/6
```

Q21

First, use $n = (\frac{t \times s}{ME})^2$ and start with $t = 2$, giving $n = (\frac{2 \times 0.09}{0.03})^2 = 36$. Then write R code that checks margin of error for values of n close to 36.

Answer: n = 38 gives ME = 0.0296.

```
alpha <- 0.05
n <- seq(32, 40, 1)
ME <- qt(1-(alpha/2), df = n-1)*sd/sqrt(n)
out <- data.frame(n, ME)
out
```

```
##    n         ME
## 1 32 0.03244846
## 2 33 0.03191261
## 3 34 0.03140248
## 4 35 0.03091608
## 5 36 0.03045162
## 6 37 0.03000749
## 7 38 0.02958226
## 8 39 0.02917463
## 9 40 0.02878340
```

Zinc (Q22 - Q27)

Q22 (4 pts)

Power = 0.573 with n = 85.

Note: By default (unless there is some compelling reason), a two-sided test should be used. But in this case the problem description ("greater than") motivates a one-sided alternative.

```
power.t.test(n = 85, delta = 0.3, sd = 1.5,
sig.level = 0.05, type = "one.sample",
alternative = "one.sided")
```

```
##
##      One-sample t test power calculation
##
##               n = 85
##           delta = 0.3
##              sd = 1.5
##       sig.level = 0.05
##           power = 0.5730619
##     alternative = one.sided
```

Q23 - Q26 Grading notes:

- full credits if students justify by trying different values of these factors that affect the power.
- full credit also for using the non-centrality parameter: $ncp = \frac{\mu_A - \mu_0}{\sigma/\sqrt(n)}$
- full credit for correctly stated intuitive answers, e.g. "increasing sample size makes it easier to reject H_0 because we have more information and a more precise estimate of the true parameter value"

Q23

Higher

Q24

Lower

Q25

Higher

Q26

Lower

Q27 (4 pts)

n = 156 to achieve power of 0.80.

```
power.t.test(delta = 0.3, sd = 1.5,
sig.level = 0.05, type = "one.sample",
alternative = "one.sided",power=0.8)
```

```
##
##      One-sample t test power calculation
##
##               n = 155.9257
##           delta = 0.3
##              sd = 1.5
##       sig.level = 0.05
##           power = 0.8
##     alternative = one.sided
```

Appendix

```r
#Retain (and do not edit) this code chunk!!!
library(knitr)
knitr::opts_chunk$set(echo = FALSE)
knitr::opts_chunk$set(message = FALSE)
Diet <- read.csv("Diet.csv")
hist(Diet$HEI, main = "Histogram of HEI",xlab="Health Eating Index")
HEI_mean <- mean(Diet$HEI)
HEI_mean
HEI_sd <- sd(Diet$HEI)
HEI_sd
tcrit <- qt(0.975,df=22)
HEI_lower <- HEI_mean - tcrit*HEI_sd/sqrt(23)
HEI_upper <- HEI_mean + tcrit*HEI_sd/sqrt(23)
c(HEI_lower,HEI_upper)
#Or, use t.test()
t.test(Diet$HEI)
HEI_test_stat <- (HEI_mean-60)/(HEI_sd/sqrt(23))
HEI_pvalue <- pt(abs(HEI_test_stat),df=22,lower.tail=F)*2
HEI_pvalue
#Or, use t.test()
t.test(Diet$HEI, mu = 60)
ybar = 4.62
s = 0.58
n = 10
mu0 = 5
RR = qt(0.975, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
RR = qt(0.05, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
n = 40
RR = qt(0.975, df = n-1)
TS = (ybar - mu0)/(s/sqrt(n))
print(data.frame(name=c("Rejection Region","Test Statistic"),value=round(c(RR,TS),2)))
Pills <- read.csv("Pills.csv")
par(mfrow = c(1, 2))
hist(Pills$content, main = "Histogram of Pill Content")
qqnorm(Pills$content);qqline(Pills$content)
Q17Out <- t.test(Pills$content)
Pills_mean <- as.numeric(Q17Out$estimate)
Pills_CI <- Q17Out$conf.int
Pills_lower <- Pills_CI[1]
Pills_upper <- Pills_CI[2]
Q17Out
Q18Out <- t.test(Pills$content, mu = 20)
Q18Out
Q19Out <- t.test(Pills$content, mu = 20, alternative = "less")
Q19Out
sd <- (0.90 - 0.36)/6
alpha <- 0.05
n <- seq(32, 40, 1)
ME <- qt(1-(alpha/2), df = n-1)*sd/sqrt(n)
```

```
out <- data.frame(n, ME)
out
power.t.test(n = 85, delta = 0.3, sd = 1.5,
sig.level = 0.05, type = "one.sample",
alternative = "one.sided")
power.t.test(delta = 0.3, sd = 1.5,
sig.level = 0.05, type = "one.sample",
alternative = "one.sided",power=0.8)
```