Group 2: Kush Parmar, Jo-Anne Rivera, Simerjeet Mudhar
I pledge my Honor that I have abided by the Stevens Honor System. -Kush, Jo-Anne, Simerjeet

Problem 1.

i. Null Hypothesis: The population mean is equal to 42.

Alternative Hypotheses:

- The population mean is not equal to 42.
- The population mean is greater than 42.
- The population mean is less than 42.

ii. Test 1: $H_0 : \mu = 42 \quad H_a: \mu < 42$

- $\overline{X} = 43.17$
- $z = \frac{\sqrt{20}(43.17-42)}{3.5} = 1.494971$
- $P_{H_0}(Z < 1.494971) = pnorm(1.494971) = 0.932539$
- $\alpha = 0.05 < 0.932539$, we do not reject H_0.

Test 2: $H_0 : \mu = 42 \quad H_a: \mu > 42$

- $\overline{X} = 43.17$
- $z = \frac{\sqrt{20}(43.17-42)}{3.5} = 1.494971$
- $P_{H_0}(Z > 1.494971) = 1 - pnorm(1.494971) = 0.06746101$
- $\alpha = 0.05 < 0.06746101$, we do not reject H_0.

Test 3: $H_0 : \mu = 42 \quad H_a: \mu \neq 42$

- $\overline{X} = 43.17$
- $z = \frac{\sqrt{20}(43.17-42)}{3.5} = 1.494971$
- $z = Z_\alpha(0.05) = 1.96$
- $-1.96 < 1.494971 < 1.96$ we do not reject H_0

iii. Test 1: $H_0 : \mu = 42 \quad H_a: \mu < 42$

-
```
x=c(41.5,50.7,36.6,37.3,34.2,45.0,48.0,
    43.2,47.7,42.2,43.2,44.6,48.4,46.4,
    46.8,39.2,37.3,43.5,44.3,43.3)
```

```
> t.test(x,alternative = c("less"), mu = 42)

        One Sample t-test

data:  x
t = 1.1852, df = 19, p-value = 0.8747
alternative hypothesis: true mean is less than 42
95 percent confidence interval:
     -Inf 44.87702
sample estimates:
mean of x
    43.17
```

- we do not reject H_0

Test 2: $H_0 : \mu = 42 \quad H_a : \mu > 42$

```
x=c(41.5,50.7,36.6,37.3,34.2,45.0,48.0,
    43.2,47.7,42.2,43.2,44.6,48.4,46.4,
    46.8,39.2,37.3,43.5,44.3,43.3)
```

```
> t.test(x,alternative = c("greater"), mu = 42)

        One Sample t-test

data:  x
t = 1.1852, df = 19, p-value = 0.1253
alternative hypothesis: true mean is greater than 42
95 percent confidence interval:
 41.46298      Inf
sample estimates:
mean of x
    43.17
```

- we do not reject H_0

Test 3: $H_0 : \mu = 42 \quad H_a : \mu \neq 42$

```
x=c(41.5,50.7,36.6,37.3,34.2,45.0,48.0,
    43.2,47.7,42.2,43.2,44.6,48.4,46.4,
    46.8,39.2,37.3,43.5,44.3,43.3)
```

```
> t.test(x,alternative = c("two.side"), mu = 42)

        One Sample t-test

data:  x
t = 1.1852, df = 19, p-value = 0.2506
alternative hypothesis: true mean is not equal to 42
95 percent confidence interval:
 41.10375 45.23625
sample estimates:
mean of x
    43.17
```

- we do not reject H_0

iv. The difference between the z-test statistic and the t-test statistic stems from the fact that the z-test statistic is used to determine if the two samples' mean calculated is different from the population standard deviation. While the t-test statistic is trying to determine how different averages found in the data set are different when the standard deviation is unknown

Problem 2

```
> z=c(rnorm(20,1,4))
> z
 [1]  1.4390667 -2.7971079 -0.7484169  0.3354255  4.5807698  0.9125910
 [7]  0.8533899 -4.3572778  3.8028736  6.5665338 -5.5242370  3.4849257
[13] -3.1352735  1.0418153 -2.9862651  2.7660605 10.1889431  4.0841377
[19]  2.4182009  0.4875509
```

```
> y=(x+z)
```

```
> y
 [1] 42.93907 47.90289 35.85158 37.63543 38.78077 45.91259 48.85339 38.84272
 [9] 51.50287 48.76653 37.67576 48.08493 45.26473 47.44182 43.81373 41.96606
[17] 47.48894 47.58414 46.71820 43.78755
```

```
> mean(y)
[1] 44.34069
```

i. Null Hypothesis: The expected value for X is equal to the expected value for Y.
Alternative Hypotheses:

- The expected value for X is not equal to 44.34069.
- The expected value for X is greater than 44.34069.
- The expected value for X is less than 44.34069.

ii.Test 1: H_a:expected value for x < 44.34069.

```
> t.test(y,x,alternative = ("less"),paired=TRUE,conf.level = 0.99)

        Paired t-test

data:  y and x
t = 1.3663, df = 19, p-value = 0.9061
alternative hypothesis: true mean difference is less than 0
99 percent confidence interval:
     -Inf 3.34665
sample estimates:
mean difference
       1.170685
```

-
- we do not reject H_0

Test 2: H_a:expected value for x > 44.34069.

```
> t.test(y,x,alternative = ("greater"),paired=TRUE,conf.level = 0.99)

        Paired t-test

data:  y and x
t = 1.3663, df = 19, p-value = 0.09391
alternative hypothesis: true mean difference is greater than 0
99 percent confidence interval:
 -1.00528      Inf
sample estimates:
mean difference
       1.170685
```

-

- we do not reject H_0

Test 3: H_a:expected value for x ≠ 44.34069.

```
> t.test(y,x,alternative = ("two.side"),paired=TRUE,conf.level = 0.99)

	Paired t-test

data:  y and x
t = 1.3663, df = 19, p-value = 0.1878
alternative hypothesis: true mean difference is not equal to 0
99 percent confidence interval:
 -1.280716  3.622087
sample estimates:
mean difference
       1.170685
```

-
- we do not reject H_0

Problem 3

Binary Data: 0 1 0 0 0 0 1 0 1 0 0 0 1 1 1 0 0 0 0 0

a. Null Hypothesis: The probability is equal to 0.4.
   Alternative Hypotheses:
   - The probability is greater than 0.4.
   - The probability is less than 0.4.
   - The probability is not equal to 0.4.

b.

```
> prop.test(sum(y), length(y), p = 0.4, alternative = c("less"))

	1-sample proportions test with continuity correction

data:  sum(y) out of length(y), null probability 0.4
X-squared = 0.46875, df = 1, p-value = 0.2468
alternative hypothesis: true p is less than 0.4
95 percent confidence interval:
 0.0000000 0.5088713
sample estimates:
  p
0.3

> prop.test(sum(y), length(y), p = 0.4, alternative = c("greater"))

	1-sample proportions test with continuity correction

data:  sum(y) out of length(y), null probability 0.4
X-squared = 0.46875, df = 1, p-value = 0.7532
alternative hypothesis: true p is greater than 0.4
95 percent confidence interval:
 0.1453628 1.0000000
sample estimates:
  p
0.3
```

```
> prop.test(sum(y), length(y), p = 0.4, alternative = c("two.side"))

        1-sample proportions test with continuity correction

data:  sum(y) out of length(y), null probability 0.4
X-squared = 0.46875, df = 1, p-value = 0.4936
alternative hypothesis: true p is not equal to 0.4
95 percent confidence interval:
 0.1283909 0.5433071
sample estimates:
  p
0.3
```

c. When we look at the degree of freedom from the result of this calculation, we can see that it is only 1 while the population is 20 different MPGs. Hence, it has less power to reject the null hypothesis, which is why it is less reliable.