ACADGILD

# SESSION 9: Statistical Inference

## Assignment 2

## PROBLEM STATEMENT

1. Calculate the p-value for the test in problem no 2.
2. How do you test the proportions and compare against hypothetical props? Test hypothesis: proportion of automatic cars is 40%

## SOLUTION

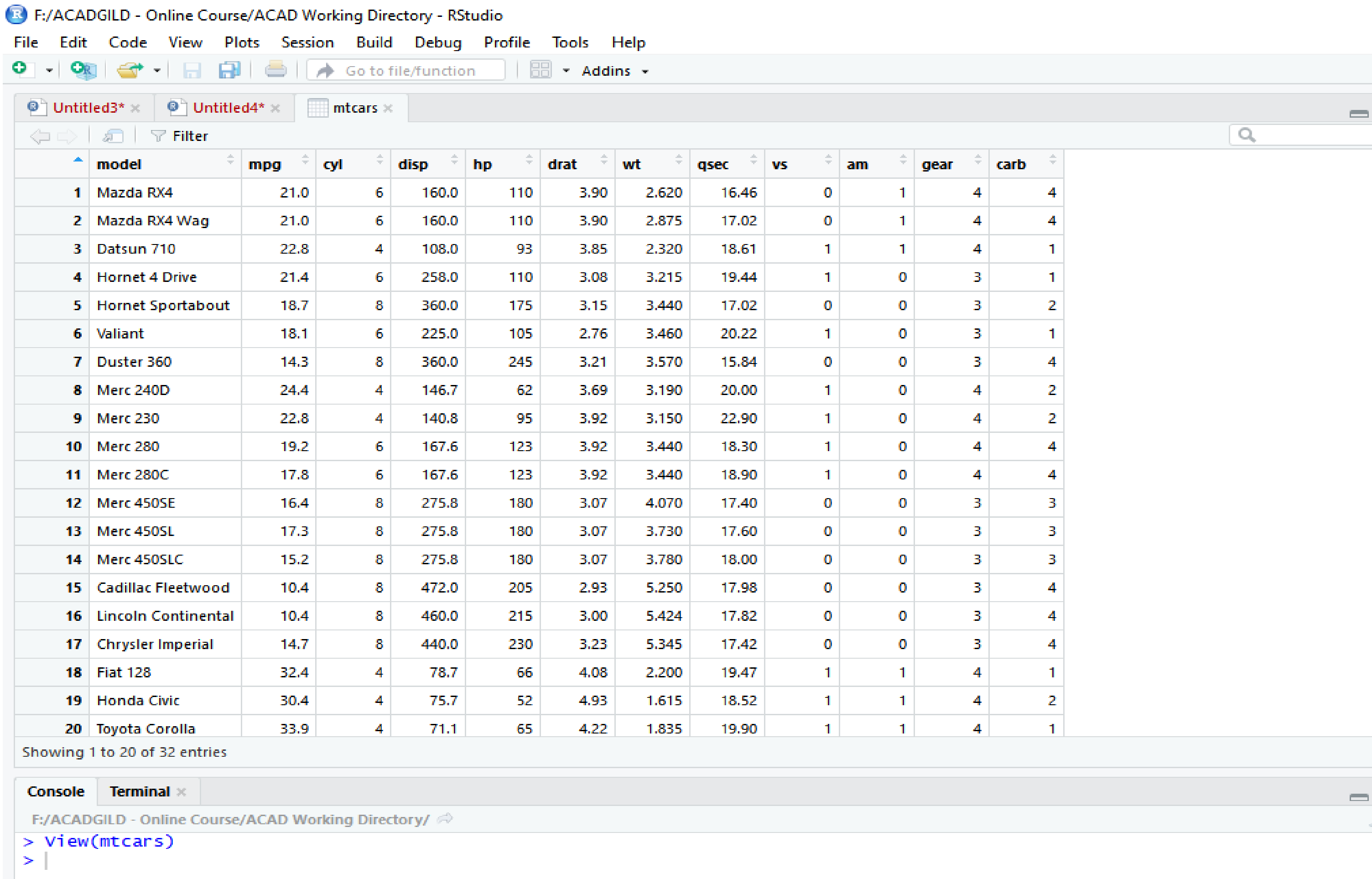
### 1. Calculate the p-value for the test in Problem no 2.

**The R-script for the given problem is as follows:**

```
library(readr)

library(psych)

mtcars <- read_csv("F:/ACADGILD - Online Course/1. DATA SETS/mtcars.csv")

View(mtcars)

mtcars

str(mtcars)

describe(mtcars$am)

table(mtcars$am)

# Calculate the P Value for the test in Problem 2.

t.test(mtcars$am,mu=10,conf.level = 0.95)

t.test(mpg~am,data = mtcars)

# OR

phat <- 13/(13 + 19)

(phat - 0.4)/sqrt(0.4 * 0.6/(13 + 19))

prop.test(13, 13 + 19, p = 0.4, alternative = "less",
          conf.level = 0.95, correct = FALSE)
```

**The output of the R-Script (from Console window) is given as follows:**

```
> library(readr)

> library(psych)

> mtcars <- read_csv("F:/ACADGILD - Online Course/1. DATA
SETS/mtcars.csv")

Parsed with column
specification: cols(

  model =
  col_character(), mpg =
  col_double(),
 cyl = col_double(),

 disp = col_double(),

 hp = col_double(),

 drat = col_double(),

 wt = col_double(),

 qsec = col_double(),

 vs = col_double(),

 am = col_double(),

 gear = col_double(),

 carb = col_double()

)

> View(mtcars)
```




| | model | mpg | cyl | disp | hp | drat | wt | qsec | vs | am | gear | carb |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Mazda RX4 | 21.0 | 6 | 160.0 | 110 | 3.90 | 2.620 | 16.46 | 0 | 1 | 4 | 4 |
| 2 | Mazda RX4 Wag | 21.0 | 6 | 160.0 | 110 | 3.90 | 2.875 | 17.02 | 0 | 1 | 4 | 4 |
| 3 | Datsun 710 | 22.8 | 4 | 108.0 | 93 | 3.85 | 2.320 | 18.61 | 1 | 1 | 4 | 1 |
| 4 | Hornet 4 Drive | 21.4 | 6 | 258.0 | 110 | 3.08 | 3.215 | 19.44 | 1 | 0 | 3 | 1 |
| 5 | Hornet Sportabout | 18.7 | 8 | 360.0 | 175 | 3.15 | 3.440 | 17.02 | 0 | 0 | 3 | 2 |
| 6 | Valiant | 18.1 | 6 | 225.0 | 105 | 2.76 | 3.460 | 20.22 | 1 | 0 | 3 | 1 |
| 7 | Duster 360 | 14.3 | 8 | 360.0 | 245 | 3.21 | 3.570 | 15.84 | 0 | 0 | 3 | 4 |
| 8 | Merc 240D | 24.4 | 4 | 146.7 | 62 | 3.69 | 3.190 | 20.00 | 1 | 0 | 4 | 2 |
| 9 | Merc 230 | 22.8 | 4 | 140.8 | 95 | 3.92 | 3.150 | 22.90 | 1 | 0 | 4 | 2 |
| 10 | Merc 280 | 19.2 | 6 | 167.6 | 123 | 3.92 | 3.440 | 18.30 | 1 | 0 | 4 | 4 |
| 11 | Merc 280C | 17.8 | 6 | 167.6 | 123 | 3.92 | 3.440 | 18.90 | 1 | 0 | 4 | 4 |
| 12 | Merc 450SE | 16.4 | 8 | 275.8 | 180 | 3.07 | 4.070 | 17.40 | 0 | 0 | 3 | 3 |
| 13 | Merc 450SL | 17.3 | 8 | 275.8 | 180 | 3.07 | 3.730 | 17.60 | 0 | 0 | 3 | 3 |
| 14 | Merc 450SLC | 15.2 | 8 | 275.8 | 180 | 3.07 | 3.780 | 18.00 | 0 | 0 | 3 | 3 |
| 15 | Cadillac Fleetwood | 10.4 | 8 | 472.0 | 205 | 2.93 | 5.250 | 17.98 | 0 | 0 | 3 | 4 |
| 16 | Lincoln Continental | 10.4 | 8 | 460.0 | 215 | 3.00 | 5.424 | 17.82 | 0 | 0 | 3 | 4 |
| 17 | Chrysler Imperial | 14.7 | 8 | 440.0 | 230 | 3.23 | 5.345 | 17.42 | 0 | 0 | 3 | 4 |
| 18 | Fiat 128 | 32.4 | 4 | 78.7 | 66 | 4.08 | 2.200 | 19.47 | 1 | 1 | 4 | 1 |
| 19 | Honda Civic | 30.4 | 4 | 75.7 | 52 | 4.93 | 1.615 | 18.52 | 1 | 1 | 4 | 2 |
| 20 | Toyota Corolla | 33.9 | 4 | 71.1 | 65 | 4.22 | 1.835 | 19.90 | 1 | 1 | 4 | 1 |

Showing 1 to 20 of 32 entries

Console Terminal

F:/ACADGILD - Online Course/ACAD Working Directory/

```
> View(mtcars)
>
```

```
> mtcars
# A tibble: 32 x
12
   model              mpg   cyl  disp    hp  drat    wt  qsec    vs    am
gear carb
                    <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
   <chr>                                                <dbl> <dbl>
<dbl
>     <dbl>
 1 Mazda RX4        21       6  160    110  3.9   2.62  16.5     0     1
4     4
 2 Mazda RX4 Wag    21       6  160    110  3.9   2.88  17.0     0     1
4     4
 3 Datsun 710       22.8     4  108     93  3.85  2.32  18.6     1     1
4     1
 4 Hornet 4 Drive   21.4     6  258    110  3.08  3.22  19.4     1     0
3     1
   Hornet
 5 Sportabout       18.7     8  360    175  3.15  3.44  17.0     0     0
3     2
 6 Valiant          18.1     6  225    105  2.76  3.46  20.2     1     0
3     1
 7 Duster 360       14.3     8  360    245  3.21  3.57  15.8     0     0
3     4
```

```
   Mer
 8 c    240D         24.4    4 147.   62 3.69 3.19 20      1    0
4    2
   Mer
 9 c    230          22.8    4 141.   95 3.92 3.15 22.9    1    0
4    2
   Mer
10 c    280          19.2    6 168.  123 3.92 3.44 18.3    1    0
4    4

# ... with 22 more
rows > str(mtcars)

Classes ‘spec_tbl_df’, ‘tbl_df’, ‘tbl’ and 'data.frame': 32 obs.
of 12 variables:

 $ model: chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4
 Drive" ...

 $     :     21        22.8 18.7 18.1 14.3 24.4 22.8 19.2
 mpg   num   21        21.4 ...
 $     :            8
 cyl   num   6 6 4 6 6   8 4 4 6 ...
 $ disp :    16     10
 num         0   160 8   258 360 ...
       :     11
 $ hp  num   0   110 93 110 175 105 245 62 95 123 ...
 $ drat :    3.     3.85     3.15 2.76 3.21 3.69 3.92
 num         9   3.9 3.08     3.92 ...
       :     2.62       2.32 3.21 3.44
 $ wt  num   2.88                ...
 $ qsec :    16.5   18.6
 num         17     19.4     17 ...
       :            0
 $ vs  num   0 0 1 1 1   0 1 1 1 ...
       :            0
 $ am  num   1 1 1 0 0   0 0 0 0 ...
 $ gear :           3
 num         4 4 4 3 3   3 4 4 4 ...
 $ carb :           2
 num         4 4 1 1 1   4 2 2 4 ...

 - attr(*, "spec")=

  .. cols(

  ..   model = col_character(),

  ..   mpg = col_double(),

  ..   cyl = col_double(),
```

```
..   disp = col_double(),
..   hp = col_double(),
..   drat = col_double(),
..   wt = col_double(),
..   qsec = col_double(),
..   vs = col_double(),
..   am = col_double(),
..   gear = col_double(),
..   carb = col_double()
.. )
> #summary(mtcars$am)
> describe(mtcars$am)
   vars n mean sd median trimmed mad min max range skew
kurtosis se X1 1 32 0.41 0.5 0 0.38 0 0 1 1 0.36 -1.92 0.09 >
table(mtcars$am)

 0  1
19 13
> t.test(mtcars$am,mu=10,conf.level = 0.95)

      One Sample t-test

data: mtcars$am
t = -108.76, df = 31, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 10
95 percent confidence interval:
 0.2263446 0.5861554
```

```
sample estimates:

mean of x

  0.40625

> t.test(mpg~am,data = mtcars)

      Welch Two Sample t-test

data: mpg by am

t = -3.7671, df = 18.332, p-value = 0.001374

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

 -11.280194    -3.209684

sample estimates:

mean in group 0 mean in group 1

      17.14737     24.39231

>

> # OR

>

> phat <- 13/(13 + 19)

> (phat - 0.4)/sqrt(0.4 * 0.6/(13 +
19)) [1] 0.07216878

>

>

> prop.test(13, 13 + 19, p = 0.4, alternative = "less",

+         conf.level = 0.95, correct = FALSE)

      1-sample proportions test without continuity correction

data: 13 out of 13 + 19, null probability
0.4 X-squared = 0.0052083, df = 1, p-
value = 0.5288 alternative hypothesis:
```

```
true p is less than 0.4 95 percent
confidence interval:

 0.0000000 0.5508812

sample estimates:

       p

 0.40625
```

2. **How do you test the proportions and compare against hypothetical props?**

**Test hypothesis: proportion of automatic cars is 40%**

**The R-script for the given problem is as follows:**

```
prop.test(13, 32, p = 0.4, alternative = "less",

    conf.level = 0.95, correct = FALSE)

#OR
prop.test(table(mtcars$am)[2],nrow(mtcars),p=0.4,alternative = "less",conf.level =
0.95,correct=FALSE)
```

**The output of the R-Script (from Console window) is given as follows:**

```
> prop.test(13, 32, p = 0.4, alternative = "less",

+         conf.level = 0.95, correct = FALSE)

      1-sample proportions test without continuity correction

data: 13 out of 32, null probability 0.4
X-squared = 0.0052083, df = 1, p-value =
0.5288 alternative hypothesis: true p is
less than 0.4 95 percent confidence
interval:

 0.0000000 0.5508812
```

```
sample estimates:

     p

 0.40625

> #OR

>

> prop.test(table(mtcars$am)[2],nrow(mtcars),p=0.4,alter
native = "less",conf.level = 0.95,correct=FALSE)

      1-sample proportions test without continuity correction

data: table(mtcars$am)[2] out of nrow(mtcars), null
probability 0.4 X-squared = 0.0052083, df = 1, p-value =
0.5288

alternative hypothesis: true p is less
than 0.4 95 percent confidence interval:

 0.0000000 0.5508812

sample estimates:

        p

 0.40625
```

## Conclusion/Interpretation:

- Test Hypothesis: proportion of automatic cars is 40%.

- At confidence level of 0.95, since p- value is greater than alpha, we fail to reject the null hypothesis