Module 3 | R Practice: mtcars dataset

Trang Tran

CPS, Northeastern University

ALY6010 | Probability Theory and Introductory Statistics

Patrick McQuillan

Mar 12, 2023

Introduction

The mtcars data present fuel consumption and 10 aspects of automobile design and performance for 32 car models. Here below are the data dictionary and the head rows of this dataset.

A data frame with 32 observations on 11 (numeric) variables.

[, 1]	mpg	Miles/(US) gallon
[, 2]	cyl	Number of cylinders
[, 3]	disp	Displacement (cu.in.)
[, 4]	hp	Gross horsepower
[, 5]	drat	Rear axle ratio
[, 6]	wt	Weight (1000 lbs)
[, 7]	qsec	1/4 mile time
[, 8]	vs	Engine (0 = V-shaped, 1 = straight)
[, 9]	am	Transmission (0 = automatic, 1 = manual)
[,10]	gear	Number of forward gears

Table 1: Data Dictionary

```
> head(mtcars)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Hypothesis Testing

#Right-tailed t-test for the mean weight (1000 lbs.) of the cars

```
        One Sample t-test

data:  mtcars$wt
t = 1.256, df = 31, p-value = 0.1092
alternative hypothesis: true mean is greater than 3
95 percent confidence interval:
 2.923979      Inf
sample estimates:
mean of x
  3.21725
```

H0: Mean weight (wt) = 3.0

H1: Mean weight (wt) > 3.0

CI: 95%, $\alpha = 0.05$

The right-tailed t-test yields a p-value of 0.1092, which is greater than the significance level (α) of 0.05. Thus, we do not have enough evidence to reject the null hypothesis, meaning we cannot conclude that the mean weight of the cars is greater than 3 (*1000 lbs). The sample mean weight is 3.21725, and the 95% confidence interval for the mean is (2.923979, Inf).

#Two-tailed t-test for mean qsec (Fastest time to travel 1/4 mile from standstill (in seconds))

```
        One Sample t-test

data:  mtcars$qsec
t = -3.6445, df = 31, p-value = 0.0009706
alternative hypothesis: true mean is not equal to 19
99 percent confidence interval:
 16.98193 18.71557
sample estimates:
mean of x
 17.84875
```

H0: Mean qsec = 19

H1: Mean qsec $\neq$ 19

CI: 99%, α = 0.01

The two-tailed t-test yields a p-value of 0.0009706, which is considerably less than the significance level (α) of 0.01. Hence, we reject the null hypothesis and conclude that the mean ¼ mile time of the cars is different from 19. As we can see, the sample mean qsec is estimated at 17.84875, and the 99% confidence interval for the mean is (16.98193, 18.71557).

#Left-tailed t-test of mean disp (engine displacement in cubic inches)

```
        One Sample t-test

data:  mtcars$disp
t = -0.8799, df = 31, p-value = 0.1928
alternative hypothesis: true mean is less than 250
95 percent confidence interval:
     -Inf 267.8698
sample estimates:
mean of x
 230.7219
```

H0: Mean disp = 250

H1: Mean disp $<$ 250

CI: 95%, α = 0.05

With a p-value of 0.1928, we fail to reject the null hypothesis at the 0.05 significance level (0.1928 > 0.05). Therefore, we do not have enough evidence to conclude that the population mean of the displacement variable is less than 250 cubic inches. The mean of the sample is 230.7219 cubic inches, and the 95% confidence interval for the population mean is (-Inf, 267.8698). We would need more data or a different test to draw a stronger conclusion.

Conclusion

Based on the results of the above t-tests, we can summarize that there is insufficient evidence to support the claims that the average weight of cars is greater than 3000 lbs. and that the mean engine displacement is less than 250 cubic inches, with a confidence level of 95%. However, we can conclude with 99% confidence that the mean of the fastest time to travel a quarter mile from a standstill is not equal to 19.