MA 2611 Lab6

Jacob Nguyen

2022-10-06

1. A snack food company uses a machine to package bags of peanuts, weighing 454 ounces each. Assume the net weights of the bags are normally distributed. As part an in-line inspection, operators check the weights of 25 randomly selected bags: 456.1 454.9 463.4 454.4 439.9 439.4 433.6 454.4 441.2 451.7 451.1 454.1 449.7 450.1 449.6 449.8 448.2 451.5 447.9 449.2 455.1 454.5 459.2 453.7 456.5

```
peanuts = c(456.1, 454.9, 463.4, 454.4, 439.9, 439.4, 433.6, 454.4, 441.2, 451.7,
451.1, 454.1, 449.7, 450.1, 449.6, 449.8, 448.2, 451.5, 447.9, 449.2, 455.1, 454.5,
459.2, 453.7, 456.5)
```

   a. Calculate a 90% confidence interval for the mean weight of the snack bags.

```
t.test(peanuts,mu=454,alternative="two.sided",conf.level=0.90)
```

```
##
##  One Sample t-test
##
## data:  peanuts
## t = -2.45, df = 24, p-value = 0.02196
## alternative hypothesis: true mean is not equal to 454
## 90 percent confidence interval:
##  448.5111 453.0249
## sample estimates:
## mean of x
##   450.768
```

   b. Perform a hypothesis test to determine whether the machine is packaging the bags of peanuts correctly with a significance level of 0.05.

```
t.test(peanuts,mu=454,alternative="two.sided")
```

```
##
##  One Sample t-test
##
## data:  peanuts
## t = -2.45, df = 24, p-value = 0.02196
## alternative hypothesis: true mean is not equal to 454
## 95 percent confidence interval:
##  448.0454 453.4906
## sample estimates:
## mean of x
##   450.768
```

We reject the null hypothesis and can determine that the machine is not packaging the bags of peanuts correctly.

2. In a randomly selected group of 371 students, 45 students chose the number seven when asked to pick a random number between one and twenty.

a. Construct a 99% confidence interval for the proportion of students in favor of choosing the number seven.

```
prop.test(x=45,n=371,p=0.05,correct=FALSE,conf.level=0.99)
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  45 out of 371, null probability 0.05
## X-squared = 39.699, df = 1, p-value = 2.962e-10
## alternative hypothesis: true p is not equal to 0.05
## 99 percent confidence interval:
##  0.08416546 0.17172958
## sample estimates:
##         p
## 0.1212938
```

b. Does this provide statistical evidence of bias in favor of the number seven, in that the proportion of students picking seven is higher than $1/20 = 0.05$? Perform a hypothesis test to provide evidence.

We reject the null hypothesis and conclude that there is enough statistical evidence that the proportion of students picking seven is higher than 1/20 because our p-value is less than $p = 0.05$.

3. A disgruntled visitor at Yellowstone National Park claimed they waited "longer than expected" to see Old Faithful erupt. According to the NPS website, the average wait time between geyser eruptions is 92 minutes. Use the R built-in data set "faithful."

a. Perform a hypothesis test against the visitor's claim at a significance level of 0.05. What conclusions do you reach?

```
t.test(faithful,mu=92,alternative="greater")
```

```
##
##  One Sample t-test
##
## data:  faithful
## t = -36.435, df = 543, p-value = 1
## alternative hypothesis: true mean is greater than 92
## 95 percent confidence interval:
##  34.71388      Inf
## sample estimates:
## mean of x
##  37.19242
```

We fail to reject the null hypothesis and conclude that there is not enough statistical evidence that the average wait time between geyser eruptions is greater than 92 minutes.

b. The NPS website also says the wait time for an eruption is between 35 and 120 mins. How does this compare to the 95% confidence interval? Would a confidence interval be appropriate for providing the range of wait time for each eruption? Why or why not?

Compared to the 95% confidence interval the lower bound of the wait time for an eruption is the same, but the upper bound of 120 minutes is not the same, as the 95% confidence interval states that the upper bound of wait time is infinite. A confidence interval would not be appropriate for providing the range of wait time for each eruption because the range is too big to be confident in, since the eruption could not happen for an infinite amount of time.