Sepehr Akbari

April 12th, 2024

Hypothesis testing with t

Instructions

For each of the following problems,

(a) Write null and alternative hypotheses appropriate to this study.

(b) Compute the t-score of the sample mean.

(c) Compute the p-value of the sample mean.

(d) Are the results statistically significant at level $\alpha = 0.05$?

(e) What conclusions, if any, can be drawn from this study? Answer in ordinary human language.

(f) Verify your calculations using the t.test( ) function.

Problem 1

A fluorescent lamp manufacturer advertises that the mean life of their lamps is $10,000$ hours. You worry that it's less. Use the lamp data set, available on Moodle, to test this claim at significance level $\alpha = 0.05$.

Answer:

Loading the dataset:

```
library(readxl)
lamp <- read_excel("/Users/sepehrakbari/Documents/LFC/Semester 2/MATH 150/DSs/lamps.xlsx")
```

(a) Null & Alternative Hypotheses

Null Hypotheses: $H_0 = 10,000$

Alternative Hypotheses: $H_1 < 10,000$

(b) t-score

```r
populationMean <- 10000
mean <- mean(lamp$hours_of_use)
stndDev <- sd(lamp$hours_of_use)
count <- length(lamp$hours_of_use)

t_score <- (mean - populationMean) / (stndDev / sqrt(count))

cat("t-score:", t_score, "\n")
```

```
t-score: -1.376544
```

(c) p-value

```r
p_value <- pt(t_score, count - 1)

cat("p-value:", p_value, "\n")
```

```
p-value: 0.08925758
```

(d) statistical significancy test

```r
alpha <- 0.05

if (p_value < alpha) {
  cat("Statistically significant (H1).")
} else {
  cat("Not statistically significant (H0).")
}
```

```
Not statistically significant (H0).
```

(e) conclusion

We cannot confirm with this data that the average life of a lamp is less than 10,000 hours. Either the null hypotheses ($H_0 = 10000$) is true, or a larger sample size or a lower significance level might be needed for a clearer picture.

(f) verification

```r
t.test(lamp$hours_of_use, mu = 10000, alternative = "less")
```

```
    One Sample t-test

data:  lamp$hours_of_use
t = -1.3765, df = 31, p-value = 0.08926
alternative hypothesis: true mean is less than 10000
95 percent confidence interval:
     -Inf 10097.12
sample estimates:
mean of x
 9580.875
```

Problem 2

A guidebook says that the average time between eruptions of the Wyoming's Old Faithful geyser is 75 minutes. Use built-in R data set faithful to test this clasim at significance level $\alpha = 0.05$.

Answer:

(a) Null & Alternative Hypotheses

Null Hypotheses: $H_0 = 75$

Alternative Hypotheses: $H_1 \neq 75$

(b) t-score

```r
populationMean <- 75
mean <- mean(faithful$waiting)
stndDev <- sd(faithful$waiting)
count <- length(faithful$waiting)

t_score <- (mean - populationMean) / (stndDev / sqrt(count))

cat("t-score:", t_score, "\n")
```

```
t-score: -4.977387
```

(c) p-value

```
p_value <- 2 * pt(t_score, count - 1)

cat("p-value:", p_value, "\n")
```

```
p-value: 1.147569e-06
```

(d) statistical significancy test

```
alpha <- 0.05

if (p_value < alpha) {
  cat("Statistically significant (H1).")
} else {
  cat("Not statistically significant (H0).")
}
```

```
Statistically significant (H1).
```

(e) conclusion

The calculation shows that there is enough evidence supporting that the mean eruption in the Old Faithful geyser is indeed 75 minutes.

(f) verification

```
t.test(faithful$waiting, mu = 75, alternative = "two.sided")
```

```
    One Sample t-test

data:  faithful$waiting
t = -4.9774, df = 271, p-value = 1.148e-06
alternative hypothesis: true mean is not equal to 75
95 percent confidence interval:
 69.27418 72.51994
sample estimates:
mean of x
 70.89706
```