Pstat120C Homewrok 3

Yu Tian

2022-08-026

```
library(ggplot2)
library(tidyverse)
```

Practice

2.

```
A = c(9.16, 13.29, 12.07, 11.97, 13.31, 12.32, 11.78)
B = c(11.95, 15.15, 14.75, 14.79, 15.48, 13.47, 13.06)
C = c(11.47, 9.54, 11.26, 13.66, 11.18, 15.03, 14.86)
D = c(11.35, 8.73, 10.00, 9.75, 11.71, 12.45, 12.38)
data2 = data.frame(A,B,C,D)
data2
```

```
##       A     B     C     D
## 1  9.16 11.95 11.47 11.35
## 2 13.29 15.15  9.54  8.73
## 3 12.07 14.75 11.26 10.00
## 4 11.97 14.79 13.66  9.75
## 5 13.31 15.48 11.18 11.71
## 6 12.32 13.47 15.03 12.45
## 7 11.78 13.06 14.86 12.38
```

```
n = 7
k=4
Y_A = sum(data2$A)
Y_B = sum(data2$B)
Y_C = sum(data2$C)
Y_D = sum(data2$D)
mean_Y_A = Y_A/n
mean_Y_B = Y_B/n
mean_Y_C = Y_C/n
mean_Y_D = Y_D/n
sum_Y = Y_A+Y_B+Y_C+Y_D
mean_Y = sum_Y/(n*k)
CM = ((sum_Y)^2)/(n*k)
CM
```

a

```
## [1] 4273.595
```

```
Total_SS = sum((data2$A)^2)+sum((data2$B)^2)+sum((data2$C)^2)+sum((data2$D)^2)-CM
Total_SS
```

```
## [1] 97.03189
```

```
SST = (Y_A^2)/n+(Y_B^2)/n+(Y_C^2)/n+(Y_D^2)/n-CM
SST
```

```
## [1] 36.74969
```

```
SSE = Total_SS-SST
SSE
```

```
## [1] 60.2822
```

```
MST = SST/(k-1)
MST
```

```
## [1] 12.2499
```

```
MSE = SSE/(n*k-k)
MSE
```

```
## [1] 2.511758
```

```
F_value = MST/MSE
F_value
```

```
## [1] 4.87702
```

Test of Hypothesis:

$$H_0 = \mu_1 = \mu_2 = \mu_3 = \mu_4$$

$$H_a = \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$$

From the calculation above, we get the F-value for the F-test is 4.8770 with the numerator and denominator degrees of freedom $v_1 = k - 1 = 3, v_2 = n - k = 24$, respectively.

From the table of p-value of F-test, with the $\alpha = 0.05$, we get the p-value is 0.0089, which is smaller than $\alpha = 0.05$.

Thus, we reject the null hypothesis H_0, and conclude that there is sufficient evidence of a difference in mean wear among the four treatments.

b

```
# get t-value from table
t = 2.797
s_2 = SSE/(n*k-k)
s =  sqrt(s_2)
CI1 = (mean_Y_B-mean_Y_C) + t*s*sqrt(1/n+1/n)
CI1
```

```
## [1] 4.033735
```

```
CI2 = (mean_Y_B-mean_Y_C) - t*s*sqrt(1/n+1/n)
CI2
```

```
## [1] -0.7051633
```

Thus, from the calculation above, we get the confidence interval is (-0.7052, 4.0337).

c

```r
# get t-value from table
t1 = 1.711
CI_1 = mean_Y_A + t1*s/sqrt(n)
CI_1
```

```
## [1] 13.01063
```

```r
CI_2 = mean_Y_A - t1*s/sqrt(n)
CI_2
```

```
## [1] 10.96079
```

Thus, from the calculation above, we get the confidence interval is (10.9608, 13.0106).