CODICEPERSONA_PROBLEMA

Marco Scarpelli

DATA

Print dataframe

```
##   accesses men women
## 1      180  29    22
## 2      235  21    22
## 3      208  12    31
## 4      220  16    36
## 5      190  25    27
## 6      208  39    28
```

```
## [1] 24  3
```

Point a

Assumptions

- Independence (we are told to assume this in the problem's text)
- Multivariate normality since we have a low amount of observations

Let us test multivariate normality:

```
##            Test        HZ    p value MVN
## 1 Henze-Zirkler 0.8135282 0.07245885 YES
```

We can see that we have a p-value of 7%, which is above 5% but below 10%. We continue with our analysis.

```
##                inf    center       sup
## accesses 195.64510 211.12500 226.60490
## men       14.88768  20.79167  26.69566
## women     22.38644  27.04167  31.69689
```

Axes direction:

```
##            [,1]        [,2]       [,3]
## [1,] 0.98529663  0.05642585  0.1612658
## [2,] 0.09957569 -0.95665950 -0.2736554
## [3,] 0.13883524  0.28568993 -0.9482120
```

Other info:

```
##                    accesses       men     women
## mean             211.125000 20.791667 27.041667
## radius             3.177302  3.177302  3.177302
## semi.axes.length  15.693925  5.842251  3.965390
```

The expression for this confidence interval is:

$$\left\{ m \in R^2 \middle| n \left(\overline{X} - m \right)^T \mathcal{S}^{-1} \left(\overline{X} - m \right) < F^* \right\}$$

with F^* equal to:

$$\frac{(n-1)p}{n-p} F_{1-\alpha}(p, n-p)$$

Point b

The expression for the simultaneous confidence intervals:

$$\text{SimCI}_{1-\alpha} \left(A\mu \right) = \left\{ \mathbf{a}^T \overline{\mathbf{X}} \pm \sqrt{\frac{(n-1)p}{n-p} F_{1-\alpha}(p, n-p)} \sqrt{\frac{\mathbf{a}^T \mathcal{S} \mathbf{a}}{n}} \right\}$$

```
##         inf    center       sup
## 1 195.64510 211.12500 226.60490
## 2  14.88768  20.79167  26.69566
## 3  22.38644  27.04167  31.69689
## 4  40.56411  47.83333  55.10255
```

Point b

```
##
##  One Sample t-test
##
## data:  linear_comb
## t = 2.7816, df = 23, p-value = 0.005305
## alternative hypothesis: true mean is greater than 0
## 95 percent confidence interval:
##  2.15273     Inf
## sample estimates:
## mean of x
##  5.608333
```