Homework 4

October 14, 2019

Math A7800: Homework 4
Name: Jorge Monzon Diaz
Email: jorgemd503@gmail.com

0.1 Dependencies

[1]:
```
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.patches import Ellipse
import pandas as pd
```

1 Exercise 5.1

1.1 (a) Evaluate T^2, for testing $H_0 : \boldsymbol{\mu^T} = [7, 11]$, using the data

$$\boldsymbol{X} = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$$

$$T^2 = n(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0})^T \boldsymbol{S^{-1}}(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0})$$

$$S = \frac{1}{n-1}(\boldsymbol{X} - \boldsymbol{1}\boldsymbol{\bar{x}^T})^T(\boldsymbol{X} - \boldsymbol{1}\boldsymbol{\bar{x}^T})$$

$$\bar{X} = \begin{bmatrix} 6 \\ 10 \end{bmatrix}$$

$$S = \frac{1}{3}\left(\begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix} - \begin{bmatrix} 6 & 10 \\ 6 & 10 \\ 6 & 10 \\ 6 & 10 \end{bmatrix}\right)^T \left(\begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix} - \begin{bmatrix} 6 & 10 \\ 6 & 10 \\ 6 & 10 \\ 6 & 10 \end{bmatrix}\right) = \frac{1}{3}\begin{bmatrix} -4 & 2 & 0 & 2 \\ 2 & -1 & -1 & 0 \end{bmatrix}\begin{bmatrix} -4 & 2 \\ 2 & -1 \\ 0 & -1 \\ 2 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 8 & -10/3 \\ -10/3 & 2 \end{bmatrix}$$

$$S^{-1} = \frac{1}{(16 - \frac{100}{9})}\begin{bmatrix} 2 & 10/3 \\ 10/3 & 8 \end{bmatrix} = \begin{bmatrix} 9/22 & 15/22 \\ 15/22 & 18/11 \end{bmatrix}$$

$$T^2 = n(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0})^T \boldsymbol{S^{-1}}(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0}) = 4\begin{bmatrix} -1 & -1 \end{bmatrix}\begin{bmatrix} 9/22 & 15/22 \\ 15/22 & 18/11 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = 4\begin{bmatrix} -1.0909 & -2.3182 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$T^2 = 4(3.4091) = 13.6364$$

1.2 (b) Specify the distribution of T^2 for the situation in (a)

$$T^2 \sim \frac{(n-1)p}{(n-p)}F_{p,n-p} = \frac{(3)2}{(2)}F_{2,2} = 3F_{2,2}$$

1.3 (c) Using (a) and (b), test H_0 at the $\alpha = .05$ level. What conclusion do you reach?

$$3F_{2,2}(.05) = 3(19) = 57$$

We fail to reject our null hypothesis at $\alpha = .05$ as $13.6364 \not> 57$.

2 Excercise 5.5

The quantities $\bar{\boldsymbol{x}}$, $\boldsymbol{S}$, and $\boldsymbol{S}^{-1}$ are given in Example 5.3 for the transformed microwave-radiation data. Conduct a test of the null hypothesis $H_0 : \boldsymbol{\mu^T} = [.55, .60]$ at the $\alpha = 0.5$ level of significance. Is your result consistent with the 95% confidence ellipse for $\boldsymbol{\mu}$ pictured in Figure 5.1?

$$\bar{X} = \begin{bmatrix} .564 \\ .603 \end{bmatrix}$$

$$S^{-1} = \begin{bmatrix} 203.018 & -163.391 \\ -163.391 & 200.228 \end{bmatrix}$$

$$T^2 = n(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0})^T \boldsymbol{S^{-1}}(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0}) = 42 \begin{bmatrix} 0.014 & 0.003 \end{bmatrix} \begin{bmatrix} 203.018 & -163.391 \\ -163.391 & 200.228 \end{bmatrix} \begin{bmatrix} 0.014 \\ 0.003 \end{bmatrix}$$

$$T^2 = 42 \begin{bmatrix} 2.3521 & -1.6866 \end{bmatrix} \begin{bmatrix} 0.014 \\ 0.003 \end{bmatrix} = 42(0.0279)$$

$$T^2 = 1.1705$$

$$T^2 \sim \frac{(n-1)p}{(n-p)} F_{p,n-p} = \frac{(41)2}{(40)} F_{2,40}$$

$$\frac{(41)2}{(40)} F_{2,40}(.05) = \frac{82}{40}(3.23) = 6.6215$$

We fail to reject our null hypothesis at $\alpha = .05$ because $1.1705 \ngtr 6.6215$. This is consistent with the 95% confidence ellipse for $\boldsymbol{\mu}$ pictured in Figure 5.1 as $\boldsymbol{\mu^T} = [.55, .60]$ falls within the ellipse.

3 Exercise 5.10(c)

Refer to the bear growth data in Example 1.10 (see Table 1.4). Restrict your attention to the measurements of length.

Obtain the 95% T^2 confidence ellipse for the mean increase in length from 2 to 3 years and the mean increase in length from 4 to 5 years.

[2]:
```
x = pd.read_excel (r'hw4.xlsx', sheet_name='data 510')
print(x)

#find difference in lengths
lengths = pd.DataFrame({'Length 2-3':x['Length 3']-x['Length 2'],'Length 4-5':
↪x['Length 5']-x['Length 4']})
print(lengths)

#print means
print("\nx-mean:")
print(lengths.mean(axis = 0))
```

```
   Length 2  Length 3  Length 4  Length 5
0       141       157       168       183
1       140       168       174       170
2       145       162       172       177
3       146       159       176       171
4       150       158       168       175
5       142       140       178       189
6       139       171       176       175
   Length 2-3  Length 4-5
0          16          15
1          28          -4
2          17           5
3          13          -5
4           8           7
5          -2          11
6          32          -1

x-mean:
Length 2-3    16.0
Length 4-5     4.0
dtype: float64
```

$$S = \frac{1}{n-1}(\boldsymbol{X} - \mathbf{1}\bar{\boldsymbol{x}}^{\boldsymbol{T}})^T(\boldsymbol{X} - \mathbf{1}\bar{\boldsymbol{x}}^{\boldsymbol{T}})$$

$$S = \frac{1}{6}\left(\begin{bmatrix} 16 & 15 \\ 28 & -4 \\ 17 & 5 \\ 13 & -5 \\ 8 & 7 \\ -2 & 11 \\ 32 & -1 \end{bmatrix} - \begin{bmatrix} 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \end{bmatrix}\right)^T \left(\begin{bmatrix} 16 & 15 \\ 28 & -4 \\ 17 & 5 \\ 13 & -5 \\ 8 & 7 \\ -2 & 11 \\ 32 & -1 \end{bmatrix} - \begin{bmatrix} 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \\ 16 & 4 \end{bmatrix}\right)$$

$$S = \frac{1}{6}\begin{bmatrix} 0 & 12 & 1 & -3 & -8 & -18 & 16 \\ 11 & -8 & 1 & -9 & 3 & 7 & -5 \end{bmatrix}\begin{bmatrix} 0 & 11 \\ 12 & -8 \\ 1 & 1 \\ -3 & -9 \\ -8 & 3 \\ -18 & 7 \\ 16 & -5 \end{bmatrix}$$

$$S = \frac{1}{6}\begin{bmatrix} 798 & -298 \\ -298 & 350 \end{bmatrix} = \begin{bmatrix} 133 & -298/6 \\ -298/6 & 350/6 \end{bmatrix}$$

$$S^{-1} = \frac{9}{47624}\begin{bmatrix} 350/6 & 298/6 \\ 298/6 & 133 \end{bmatrix} = \begin{bmatrix} 0.011024 & 0.009386 \\ 0.009386 & 0.025135 \end{bmatrix}$$

The 95% confidence region is given by:

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu_0})^T \boldsymbol{S^{-1}}(\bar{\mathbf{X}} - \boldsymbol{\mu_0}) \leq \frac{(n-1)p}{(n-p)} F_{p,n-p}(0.05)$$

$$\frac{(6)4}{(3)} F_{4,3}(0.05) = 8 * 9.12 = 72.96$$

$$7\begin{bmatrix} 16 - \mu_{2-3} & 4 - \mu_{4-5} \end{bmatrix}\begin{bmatrix} 0.011024 & 0.009386 \\ 0.009386 & 0.025135 \end{bmatrix}\begin{bmatrix} 16 - \mu_{2-3} \\ 4 - \mu_{4-5} \end{bmatrix} \leq 72.96$$

$$\begin{bmatrix} 16 - \mu_{2-3} & 4 - \mu_{4-5} \end{bmatrix}\begin{bmatrix} 0.011024 & 0.009386 \\ 0.009386 & 0.025135 \end{bmatrix}\begin{bmatrix} 16 - \mu_{2-3} \\ 4 - \mu_{4-5} \end{bmatrix} \leq 10.42$$

$$|S - I\lambda| = \begin{vmatrix} 133 - \lambda & -298/6 \\ -298/6 & 350/6 - \lambda \end{vmatrix} = \lambda^2 - \frac{574\lambda}{3} + \frac{47624}{9} = 0$$

$$\lambda = 157.80, 33.533$$

$$S - \lambda_1 I = \begin{bmatrix} 75.2 & -49.6667 \\ -49.6667 & -99.4667 \end{bmatrix}, u_1 = \begin{pmatrix} -298.4 \\ 149 \end{pmatrix}$$

$$S - \lambda_2 I = \begin{bmatrix} 99.467 & -49.6667 \\ -49.6667 & 24.8 \end{bmatrix}, u_2 = \begin{pmatrix} 74.4 \\ 149 \end{pmatrix}$$

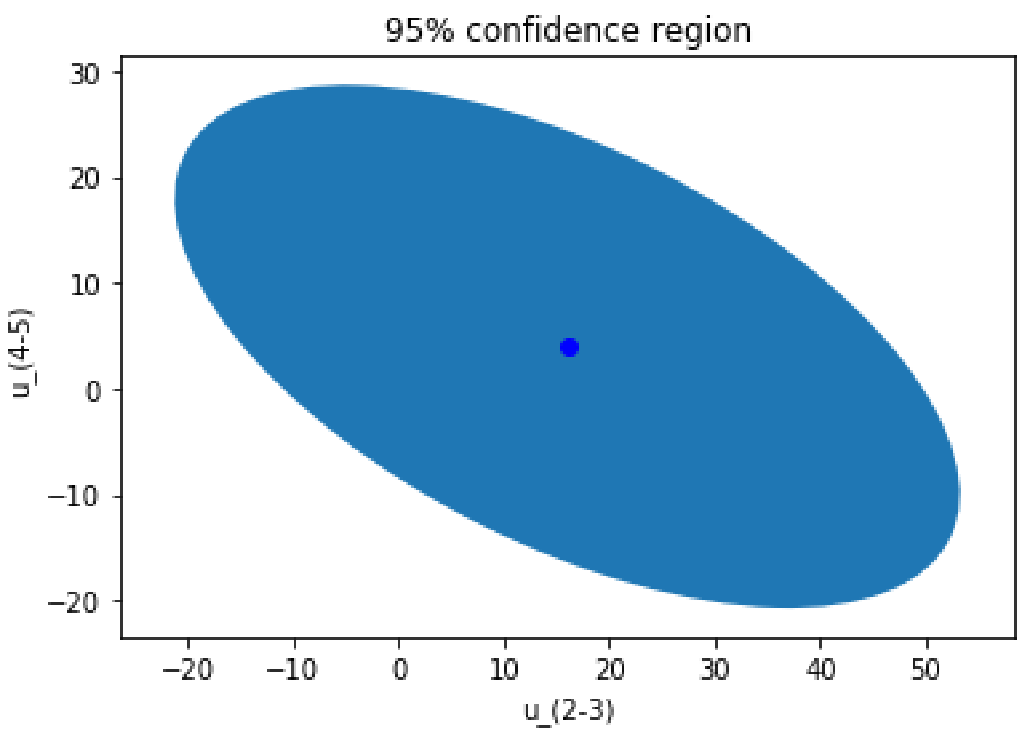
Half-axis length given by:

$$\sqrt{\frac{(n-1)p}{(n-p)n} F_{p,n-p}(0.05)\lambda_i} = \sqrt{10.4229\lambda_i}$$

Semi-major axis length = 40.5553
Semi-minor axis length = 18.6952

[3]:
```
#create ellipse
ell=Ellipse((16,4), 2*40.5553, 2*18.6952, np.degrees(np.arctan2(149, -298.4)))
#plot center
plt.plot(16, 4, marker='o', color="blue")

#create plot
plt.title("95% confidence region")
plt.xlabel("u_(2-3)")
plt.ylabel("u_(4-5)")
ax = plt.gca()
ax.add_patch(ell)
ax.autoscale()
plt.show()
print(ell)
```




```
Ellipse(xy=(16, 4), width=81.1106, height=37.3904, angle=153.4656786561316)
```

4 Exercise 5.11(a)

A physical anthropologist performed a mineral analysis of nince ancient Peruvian hairs. The results for the chromium (x_1) and strontium (x_2) levels, in parts per million (ppm), were as follows:

[4]:
```
x = pd.read_excel (r'hw4.xlsx', sheet_name='data 511')
print(x)

#print means
print("\nx-mean:")
print(x.mean(axis = 0))
```

```
   x1(Cr)  x2(St)
0    0.48   12.57
1   40.53   73.68
2    2.19   11.13
3    0.55   20.03
4    0.74   20.29
5    0.66    0.78
6    0.93    4.64
7    0.37    0.43
8    0.22    1.08

x-mean:
x1(Cr)     5.185556
x2(St)    16.070000
dtype: float64
```

It is known that low levels (less than or equal to .100 ppm) of chromium suggest the presence of diabetes, while strontium is an indication of animal protein intake.

Construct and plot a 90% joint confidence ellipse for the population mean vector $\boldsymbol{\mu}^T = [\mu_1, \mu_2]$, assuming that these nine Peruvian hairs represent a random sample from individuals belonging to a particular ancient Peruvian culture.

$$S = \frac{1}{n-1}(\boldsymbol{X} - \boldsymbol{1}\bar{\boldsymbol{x}}^{\boldsymbol{T}})^T(\boldsymbol{X} - \boldsymbol{1}\bar{\boldsymbol{x}}^{\boldsymbol{T}})$$

$$S = \frac{1}{8}\left(\begin{bmatrix} .48 & 12.57 \\ 40.53 & 73.68 \\ 2.19 & 11.13 \\ .55 & 20.03 \\ .74 & 20.29 \\ .66 & .78 \\ .93 & 4.64 \\ .37 & .43 \\ .22 & 1.08 \end{bmatrix} - \begin{bmatrix} 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \end{bmatrix}\right)^T \left(\begin{bmatrix} .48 & 12.57 \\ 40.53 & 73.68 \\ 2.19 & 11.13 \\ .55 & 20.03 \\ .74 & 20.29 \\ .66 & .78 \\ .93 & 4.64 \\ .37 & .43 \\ .22 & 1.08 \end{bmatrix} - \begin{bmatrix} 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \\ 5.185556 & 16.07 \end{bmatrix}\right)$$

$$S = \begin{bmatrix} 176.0042 & 287.2412 \\ 287.2412 & 527.8493 \end{bmatrix}$$

$$S^{-1} = \frac{1}{176.0042(527.8493) - 287.2412(287.2412)} \begin{bmatrix} 527.8493 & -287.2412 \\ -287.2412 & 176.0042 \end{bmatrix} = \begin{bmatrix} .0508 & -.0276 \\ -.0276 & .0169 \end{bmatrix}$$

The 90% confidence region is given by:

$$n(\bar{\boldsymbol{X}} - \boldsymbol{\mu_0})^T \boldsymbol{S^{-1}} (\bar{\boldsymbol{X}} - \boldsymbol{\mu_0}) \leq \frac{(n-1)p}{(n-p)} F_{p,n-p}(0.1)$$

$$\frac{(8)2}{(7)} F_{2,7}(0.1) = \frac{16}{7} * 3.26 = 7.45$$

$$9 \begin{bmatrix} 5.185556 - \mu_1 & 16.07 - \mu_2 \end{bmatrix} \begin{bmatrix} .0508 & -.0276 \\ -.0276 & .0169 \end{bmatrix} \begin{bmatrix} 5.185556 - \mu_1 \\ 16.07 - \mu_{42} \end{bmatrix} \leq 7.45$$

$$\begin{bmatrix} 5.185556 - \mu_1 & 16.07 - \mu_2 \end{bmatrix} \begin{bmatrix} .0508 & -.0276 \\ -.0276 & .0169 \end{bmatrix} \begin{bmatrix} 5.185556 - \mu_1 \\ 16.07 - \mu_{42} \end{bmatrix} \leq .8278$$

$$|S - I\lambda| = \begin{vmatrix} 176.0042 - \lambda & 287.2412 \\ 287.2412 & 527.8493 - \lambda \end{vmatrix} = \lambda^2 - 703.853\lambda + 10396.2 = 0$$

$$\lambda = 688.759, 15.0941$$

$$S - \lambda_1 I = \begin{bmatrix} -512.7548 & 287.2412 \\ 287.2412 & -160.9097 \end{bmatrix}, u_1 = \begin{pmatrix} .5602 \\ 1 \end{pmatrix}$$

$$S - \lambda_2 I = \begin{bmatrix} 160.9101 & 287.2412 \\ 287.2412 & 512.7552 \end{bmatrix}, u_2 = \begin{pmatrix} -1.7851 \\ 1 \end{pmatrix}$$

Half-axis length given by:

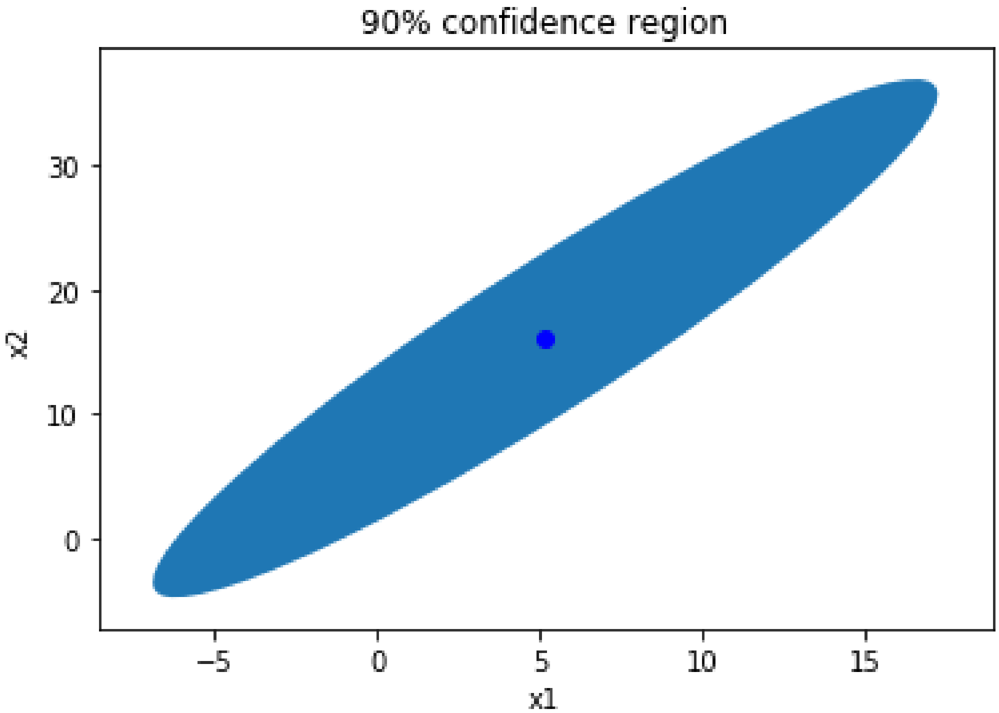
$$\sqrt{\frac{(n-1)p}{(n-p)n} F_{p,n-p}(0.1)\lambda_i} = \sqrt{.8278\lambda_i}$$

Semi-major axis length = 23.8776
Semi-minor axis length = 3.548

[5]:
```
#create ellipse
ell=Ellipse((5.185556,16.07), 2*23.8776, 2*3.548, np.degrees(np.arctan2(1, .
 ↪5602)))
#plot center
plt.plot(5.185556, 16.07, marker='o', color="blue")

#create plot
plt.title("90% confidence region")
plt.xlabel("x1")
plt.ylabel("x2")
ax = plt.gca()
ax.add_patch(ell)
ax.autoscale()
print(ell)
```

```
Ellipse(xy=(5.185556, 16.07), width=47.7552, height=7.096,
angle=60.74245093288949)
```



5 Exercise 5.13

Determine the approximate distribution of $-n\ln(|\hat{\mathbf{\Sigma}}|/|\hat{\mathbf{\Sigma}}_{\mathbf{0}}|)$ for the sweat data in Table 5.1

$$-n\ln\left(\frac{|\hat{\mathbf{\Sigma}}|}{|\hat{\mathbf{\Sigma}}_{\mathbf{0}}|}\right) \sim \chi^2_{v-v_0}$$

$$v = p + \frac{p(p+1)}{2} = 3 + 6 = 9, v_0 = \frac{p(p+1)}{2} = 6$$

$$-n\ln\left(\frac{|\hat{\mathbf{\Sigma}}|}{|\hat{\mathbf{\Sigma}}_{\mathbf{0}}|}\right) \sim \chi^2_3$$