29874408

Statistical Computing Written Report

Section 1: R programming I

Question 1.

A)

268 Patients that have diabetes

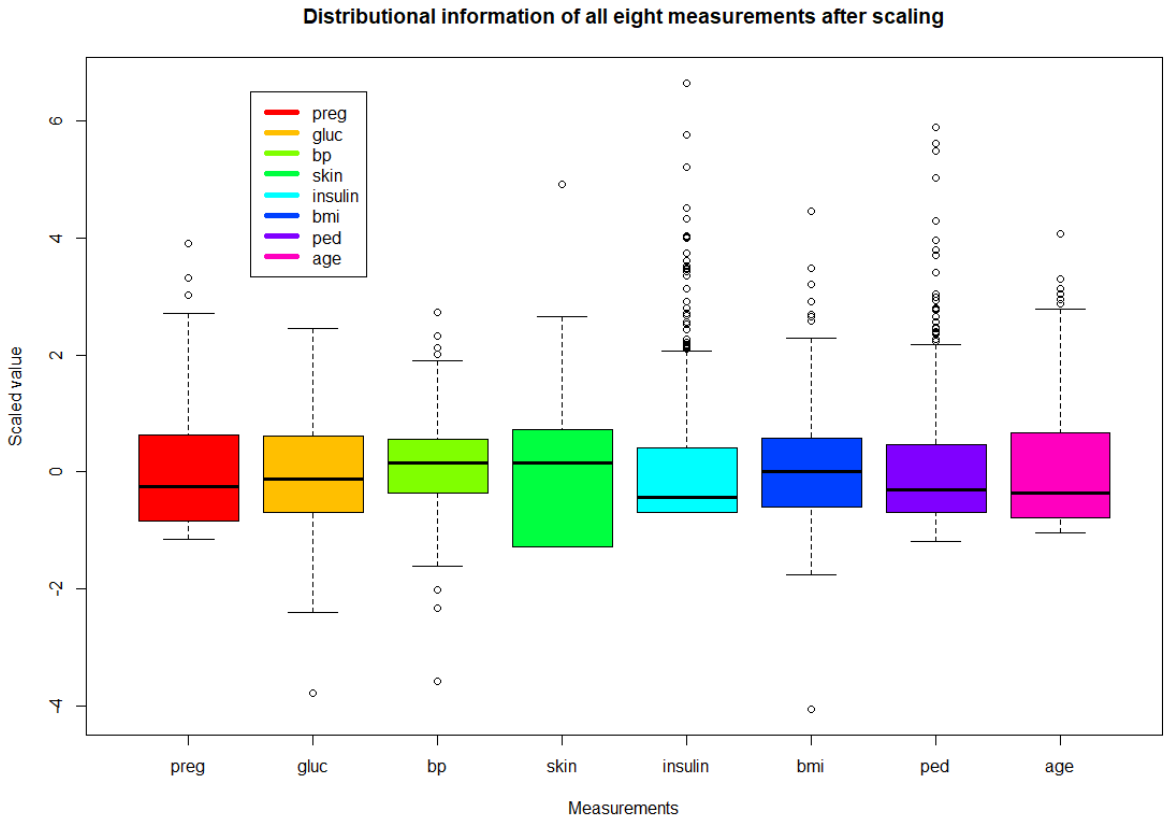
500 patients that DO NOT have diabetes, making for 768 total patients

Median patient age is 29 years old

Mean BMI for patients with diabetes is 35.142564

Mean BMI for patients without diabetes is 30.3042

B)



29874408

C)

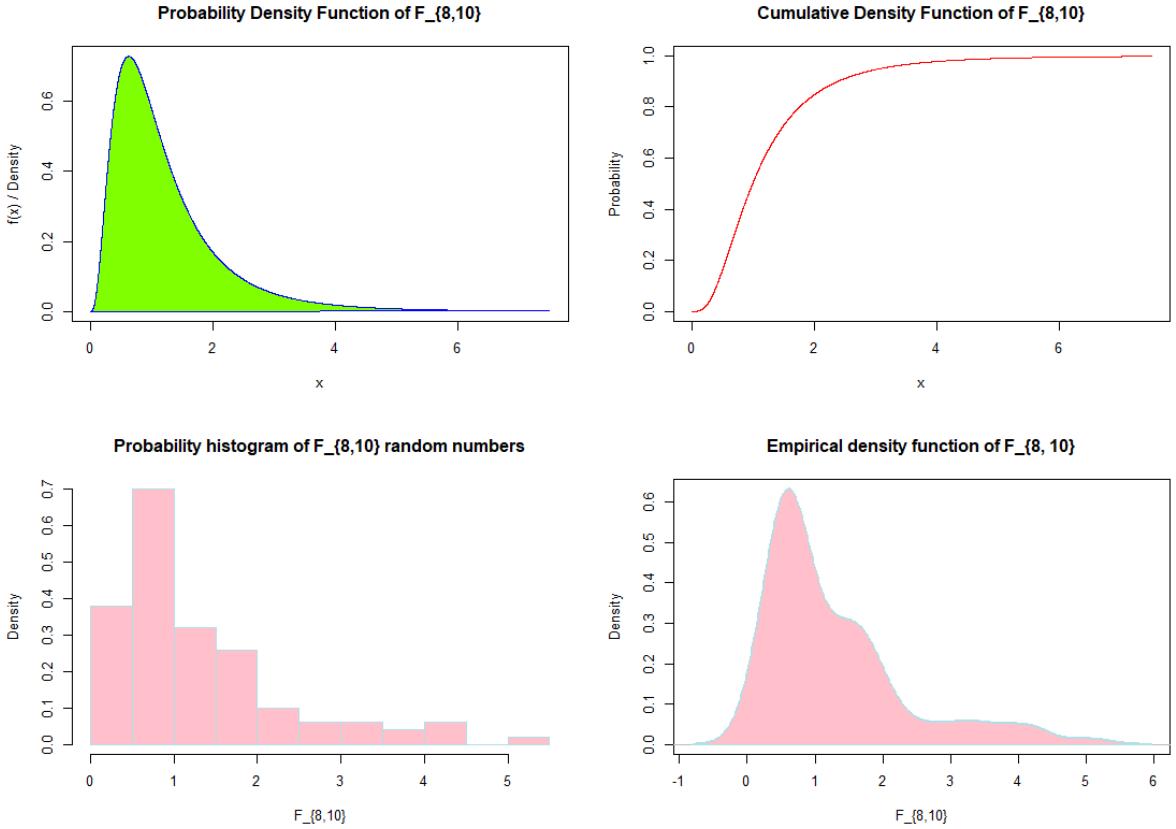
Null Hypothesis H_0: There is no difference between the means of the two groups

Alternative Hypothesis H_1: There is a difference between the means of the two groups

p-value for the t-test is 1.229807e-16

So we REJECT the null hypothesis as $p < 0.05$

D)



E)

Hotelling T^2 statistic is 333.3949

F)

Null Hypothesis H_0: For patients with and without diabetes the mean values of all measurements are equal

Alternative Hypothesis H_1: Mean values of all measurements are not equal

29874408

p-value is given by 7.355568e-55

So we REJECT the null hypothesis as p < 0.05

Question 2.

A)

Beta hat matrix given by

```
                  y1            y2
inter    1.8538942665 -0.8538942665
preg    -0.0205918715  0.0205918715
gluc    -0.0059202729  0.0059202729
bp       0.0023318790 -0.0023318790
skin    -0.0001545198  0.0001545198
insulin  0.0001805345 -0.0001805345
bmi     -0.0132440315  0.0132440315
ped     -0.1472374386  0.1472374386
age     -0.0026213938  0.0026213938
```

B)

```
> predictedy1
             y1        y2
[1,] 0.7103077 0.2896923
> predictedy2
             y1       y2
[1,] -0.2010554 1.201055
> predictedy3
             y1        y2
[1,] 0.3127774 0.6872226
```

We have seen that patient 1 is predicted to be a non-diabetic patient as y1 > y2 whereas patients 2 and 3 are predicted to be diabetic as y2 > y1.

Section 2: R programming II

Question 1.

A) Function

B) Function

C) Function

29874408

Question 2.

A) Function

B) Function

C) Function

Question 3.

A) Function

B) Function

C) Function

End of coursework