hk_1_0918

Ruoyuan Qian

2019/9/18

Problem 1

a)

I found that there are two rows of variable names which may cause some errors when reading the data into R. As a result, I use `skip=` to skip first row and read others into R.

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    72 obs. of  20 variables:
##  $ Group     : num  1 1 1 1 1 1 1 1 1 1 ...
##  $ Age       : num  57 65 61 54 64 41 59 57 67 44 ...
##  $ Gender    : num  1 1 1 2 1 2 1 2 1 2 ...
##  $ Race      : num  1 1 1 2 1 1 1 2 1 1 ...
##  $ HTN       : num  1 1 1 1 1 0 1 0 1 0 ...
##  $ T2DM      : num  0 1 0 1 1 0 1 0 1 0 ...
##  $ Depression: num  0 0 0 0 0 1 0 0 0 0 ...
##  $ Smokes    : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ PRE...9   : num  160 126 120 140 148 116 142 123 160 106 ...
##  $ POST...10 : num  163 92 121 121 123 130 121 113 124 111 ...
##  $ PRE...11  : num  102 59 67 81 63 82 69 77 81 65 ...
##  $ POST...12 : num  107 57 68 65 58 87 68 65 80 65 ...
##  $ PRE...13  : num  33 25.7 26.1 41.6 29.2 31.7 27.6 25.7 36.6 32.3 ...
##  $ POST...14 : num  32.7 25.7 25.3 39.7 28.4 31.3 28.4 25.8 35.6 31.9 ...
##  $ PRE...15  : num  60 40 88 44 48 66 37 69 40 45 ...
##  $ POST...16 : num  62 43 67 24 52 56 44 73 35 49 ...
##  $ PRE...17  : num  110 133 114 112 63 62 89 117 77 126 ...
##  $ POST...18 : num  107 96 98 75 58 86 81 129 73 131 ...
##  $ PRE...19  : num  96 106 92 401 96 75 66 96 113 91 ...
##  $ POST...20 : num  105 132 95 162 216 92 72 71 101 92 ...
```

```
## [1] FALSE
```

```
##
##  1  2  3
## 51 19  2
```

```
##
##
## |                    | Control (N=36)  | Intervention (N=36) |
## |:-------------------|:---------------:|:-------------------:|
## |Age                 |                 |                     |
## |-  Mean (SD)        | 51.500 (10.809) |   53.583 (9.581)    |
## |-  Median           |     51.000      |       55.500        |
## |Gender              |                 |                     |
## |-  female           |   20 (55.6%)    |     20 (55.6%)      |
## |-  male             |   16 (44.4%)    |     16 (44.4%)      |
## |Race                |                 |                     |
## |-  African American |   22 (61.1%)    |     31 (86.1%)      |
## |-  Hispanic         |   14 (38.9%)    |      5 (13.9%)      |
## |Depression          |                 |                     |
```

```
## |-  no              |   23 (63.9%)    |     26 (72.2%)      |
## |-  yes             |   13 (36.1%)    |     10 (27.8%)      |
## |Smokes             |                 |                     |
## |-  no              |   31 (86.1%)    |     31 (86.1%)      |
## |-  yes             |    5 (13.9%)    |      5 (13.9%)      |
## |HTN                |                 |                     |
## |-  no              |   16 (44.4%)    |     14 (38.9%)      |
## |-  yes             |   20 (55.6%)    |     22 (61.1%)      |
## |T2DM               |                 |                     |
## |-  no              |   17 (47.2%)    |     23 (63.9%)      |
## |-  yes             |   19 (52.8%)    |     13 (36.1%)      |
```

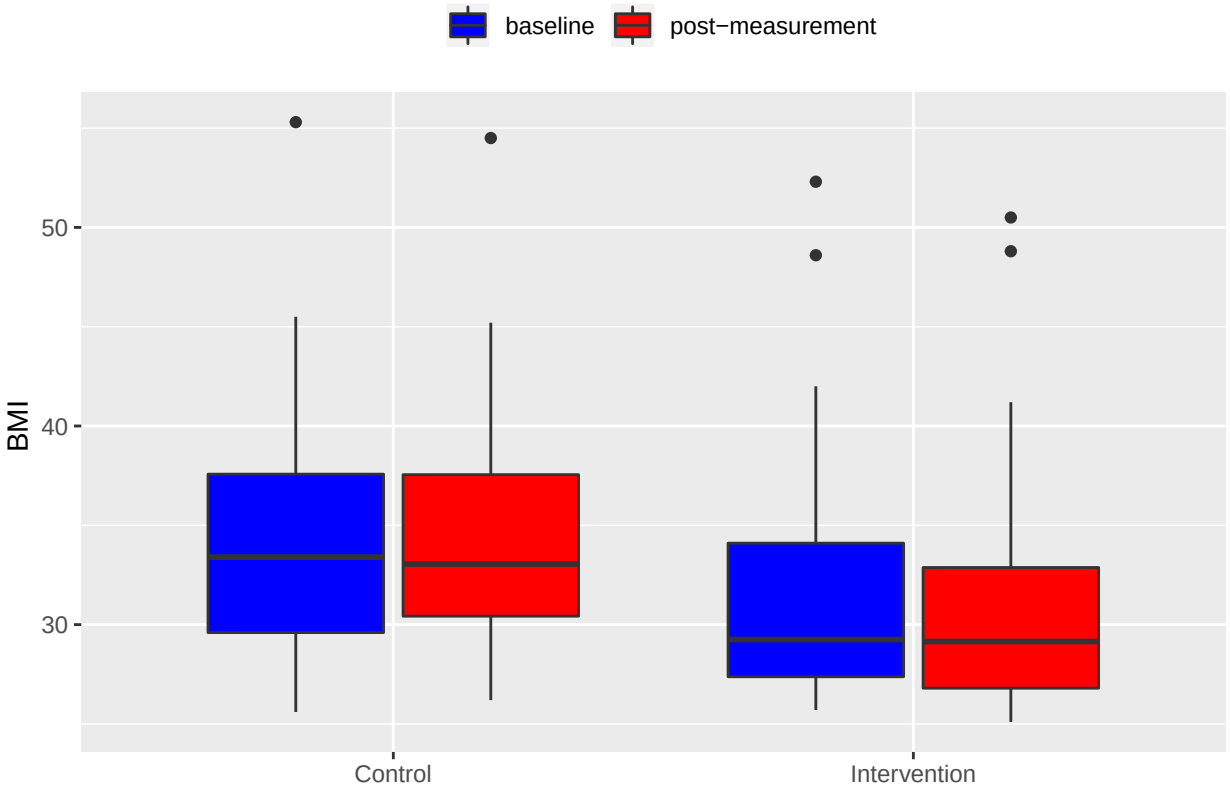
b)

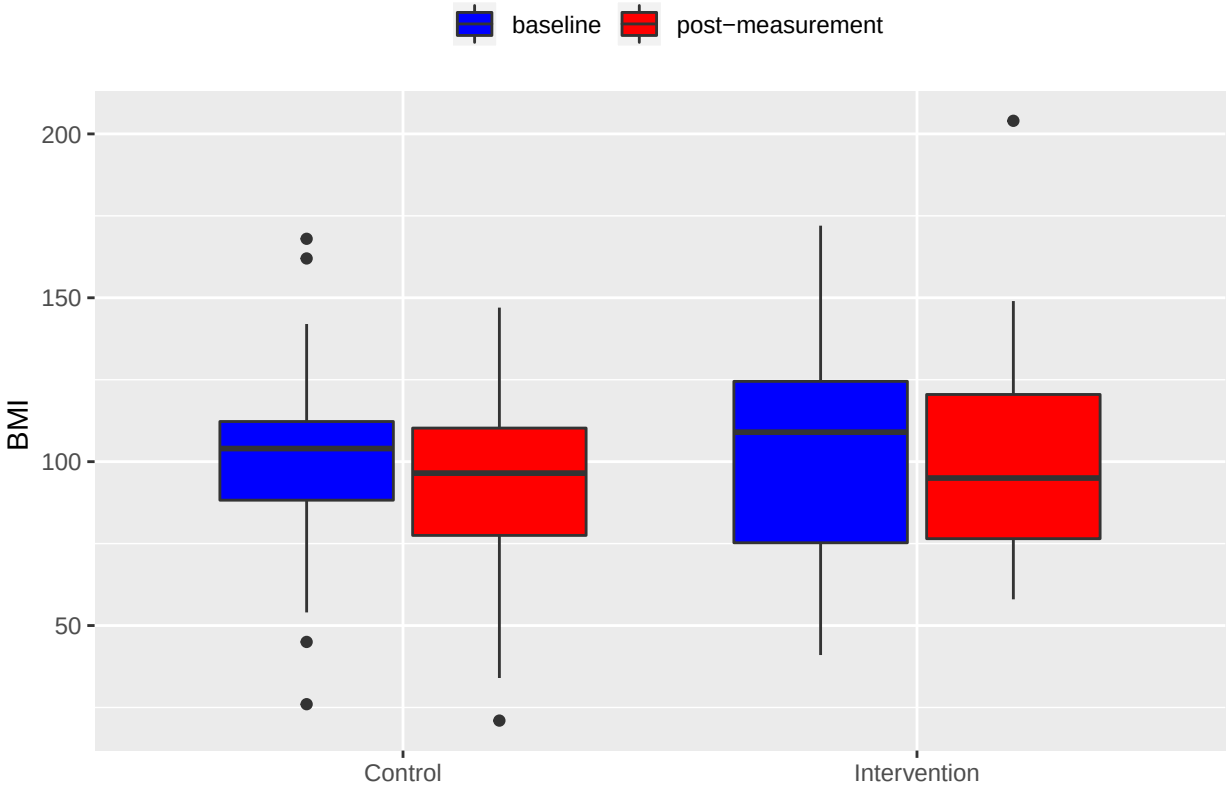
b) i

```
##    X Variable       Intervention.N.36      Intervention.N.36.
## 1  1                          baseline              six month
## 2  2      SYS          133.64 ± 15.11          125.06 ± 15.44
## 3  3              134 ( 121.5 - 144 )    124 ( 116.75 - 135 )
## 4  4    delta           -8.58 ± 17.17
## 5  5      DIA            75.44 ± 9.10           74.58 ± 12.37
## 6  6                74.5 ( 69 - 81 )        74 ( 65 - 80.5 )
## 7  7    delta            -0.86 ± 8.30
## 8  8      BMI            31.97 ± 6.58            31.21 ± 6.13
## 9  9          29.25 ( 27.375 - 34.1 ) 29.15 ( 26.8 - 32.875 )
## 10 10   delta            -0.76 ± 1.44
## 11 11     HDL           50.17 ± 11.85           50.17 ± 13.07
## 12 12                47.5 ( 40 - 60 )     48.5 ( 43 - 60.25 )
## 13 13   delta             0.00 ± 8.09
## 14 14     LDL          102.94 ± 33.84          100.50 ± 30.39
## 15 15           109 ( 75.25 - 124.5 )     95 ( 76.5 - 120.5 )
## 16 16   delta           -2.44 ± 21.27
## 17 17     GLU          116.64 ± 74.91          107.14 ± 38.65
## 18 18            94 ( 83.75 - 116.5 )    95.5 ( 85.25 - 129 )
## 19 19   delta           -9.50 ± 57.36
##              Control.N.36           Control.N.36.
## 1                baseline               six month
## 2          133.47 ± 15.94          130.14 ± 14.35
## 3   131 ( 122.5 - 143.5 )      127.5 ( 120 - 140 )
## 4           -3.33 ± 14.81
## 5            77.14 ± 9.66            75.69 ± 7.54
## 6       76 ( 68.75 - 85 )         76.5 ( 69 - 82 )
## 7           -1.44 ± 10.11
## 8            34.23 ± 6.16            34.51 ± 5.97
## 9  33.4 ( 29.6 - 37.575 ) 33.05 ( 30.425 - 37.55 )
## 10            0.28 ± 0.97
## 11          48.33 ± 13.70           45.19 ± 10.78
## 12    43.5 ( 39 - 54.25 )         43.5 ( 38 - 52 )
## 13           -3.14 ± 6.91
## 14          99.83 ± 29.06           93.61 ± 27.47
## 15 104 ( 88.25 - 112.25 )   96.5 ( 77.5 - 110.25 )
## 16          -6.22 ± 23.12
## 17         128.97 ± 73.86          126.61 ± 63.96
## 18     98 ( 81.75 - 139 )    106.5 ( 85 - 145.75 )
```

```
## 19          -2.36 ± 51.22
```

b) ii


baseline
post−measurement
BMI
50
40
30
Control
Intervention


baseline
post−measurement
BMI
200
150
100
50
Control
Intervention

b) iii

According to table of Pre/Post changes in metabolic parameters, structured exercise program has a greater impact on the`Systolic`, `Glu` variables in the intervention group than on the corresponding variables in the control group. So it means that the program may affect `Systolic` , `Glu` and `BMI`more apparently compared with the impact on the other variables.

Moreover, there are some variables that changed more severely in control than in intervention group, such as `diastolic`, `HDL` and `LDL`. In other words, the program is not benefial to the decline of `diastolic`, `HDL` and `LDL`.

As for the BMI in the study, in general, the values in intervention group are higher than the control group's while there is little difference of values within each group.

As for the LDL, for both groups, there is a trend that the values in post measurement are lower than the baseline's, and it is more obvious in control. Besides, in intervention , the values are more concentrated whereas values in control are more dispersed.

c)

It is an interventional clinical trial, and from the perspective of demographics, the two groups are basically balanced, which is favorable for the study.

However, there is some potential issues as well. For instance, the medians of baselines of BMI in two groups are of great difference, which may make the two groups incomparable. Besides, as for the LDL, the variances of two groups are extremely different, which may cause the similar problem like BMI.

Besides, the study dose not implement binding, so there may contain a lot of bias from doctors, subjects and so on. What is worse,the control is non-participating, as a result, the new program in the study is just compared with blank control instead of standard control. Due to that, the study can only tell the difference btween new program and non-participant but fails to test the difference btween new program and established one.

To sum up, the study contains plenty of bias and the result is not convincing enough.

Problem 2

The probability is The probability is`(0.6×0.001)/(0.6×0.001+0.05×0.999)=0.012`,

Problem 3

aaa