Homework 3

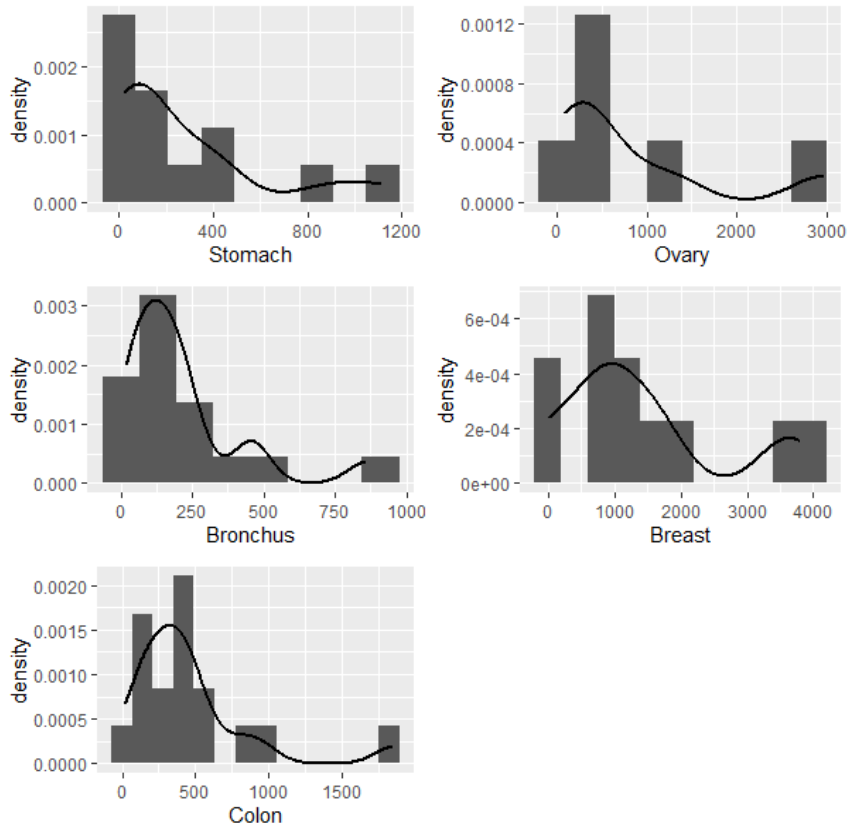
Ruoyuan Qian

2019/10/26

# Problem 2

## a)

According to Plot 1, the distributions of each group are heavily skewed, so median and IQR are more suitable and accurate to be descriptive statistics in the situation.



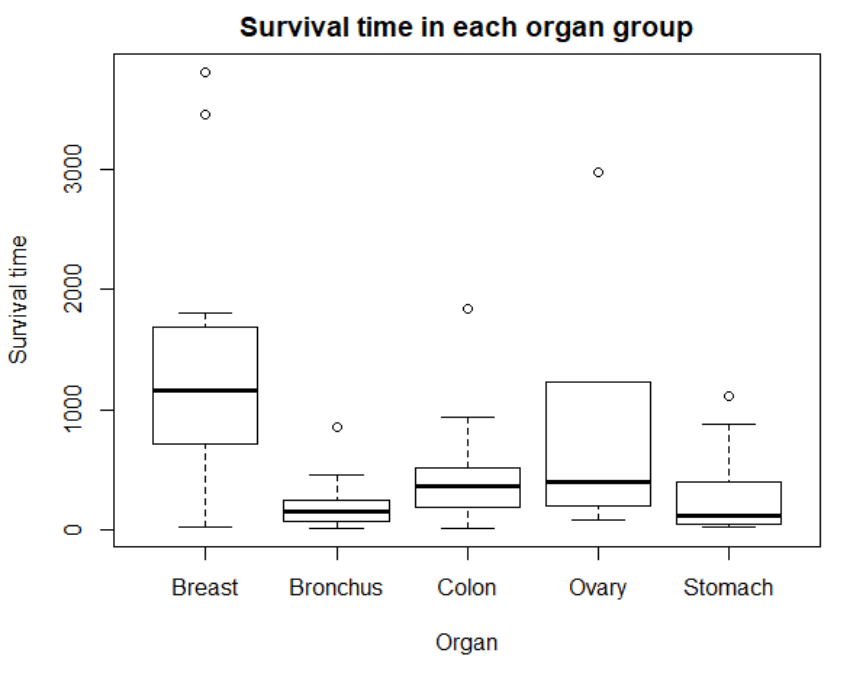
Plot 1 Distributions of organ group

According to Table 1, the median of survival times in Breast group is significantly different from others whereas the median of Stomach is similar to the median of Bronchus as well as the median of Colon is similar to the median of Ovary.

Moreover, as for Plot 2, the distributions of Ovary and Stomach are extremely skewed, compared with others.

Table 1 Descriptive statistics of each group

|  |  |  |
| --- | --- | --- |
| **Organ** | **Median** | **IQR (25%-75%)** |
| Stomach | 124 | 46 - 396 |
| Bronchus | 155 | 72 - 245 |
| Colon | 372 | 189 - 519 |
| Ovary | 406 | 239.75 - 1039.5 |
| Breast | 1166 | 723 - 1692.5 |



Plot 2 Boxplot of survival times by group

## b)

(1) Population parameter of interest:

The true means of survival times of 5 organ groups(Stomach: , Bronchus: , Colon: , Ovary: , Breast: ).

(2) Hypothesis:

,

*at least two group means are different*

(3) Significant level:



(4) Assumption:

Normality assumed.

(5) Form of the test statistic and its null distribution:









 under the null hypothesis

Table 2 ANOVA table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Df** | **Sum of Square (SS)** | **Mean Sum of Square** | **F-Statistics** | **P-value** |
| Between | 4 | 11535761 | 2883940 | 6.433 | 0.000229 |
| Within | 59 | 26448144 | 448274 |  |  |
| Total | 63 | 37983905 |  |  |  |

(6) Decision rule:

Reject  if 

(7) Interpretation:

Since , we reject the null hypothesis with 0.01 significance level, and we conclude that there is at least two of group means are different.

## c)

Pair-wise comparisons:

1. Bonferroni adjustment:

Based on part b), here we have 10 comparisons(). Assuming that denotes the comparison between *i*th and *j*th group().

The adjusted significance level is:



We can implement the method above for all pair-wise comparisons and we can get the Table 3.

Table 3 P-value of Pair-wise comparisons under Bonferroni adjustment

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Breast | Bronchus | Colon | Ovary |
| Bronchus | 0.00025 | - | - | - |
| Colon | 0.00608 | 1.00000 | - | - |
| Ovary | 1.00000 | 0.38575 | 1.00000 | - |
| Stomach | 0.00153 | 1.00000 | 1.00000 | 0.75283 |

1. Tukey adjustment:

Assuming that denotes the comparison between *i*th and *j*th group().

The hypothesis is:



Significant level:



Form of the test statistic and its null distribution:



where,  is the overall number of group, is the total number of observation in all groups. *q* represents studentized range distribution in which parameters are k and n-k. We reject the null if . Here we have 10 pairwise comparisons which result in 10 P-values:

Table 4 P-value of Pair-wise comparisons under Tukey adjustment

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Breast | Bronchus | Colon | Ovary |
| Bronchus | 0.000239 | - | - | - |
| Colon | 0.005307 | 0.82084 | - | - |
| Ovary | 0.56309 | 0.227108 | 0.665912 | - |
| Stomach | 0.001396 | 0.998146 | 0.956829 | 0.377292 |

1. Dunnett adjustment:

Table 5 P-value of Pair-wise comparisons under Dunnett adjustment

|  |  |
| --- | --- |
| $Breast | p-value |
| Bronchus-Breast | 9.10x10-5 |
| Colon-Breast | 0.00221 |
| Ovary-Breast | 0.36692 |
| Stomach-Breast | 0.00056 |

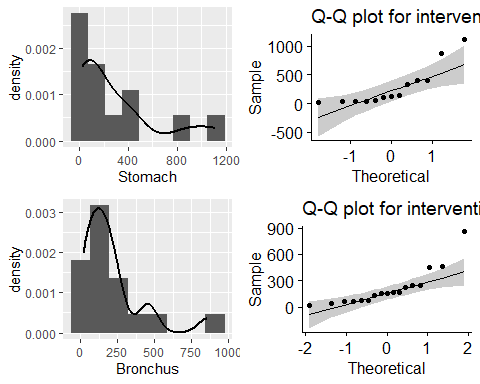
Conclusion:

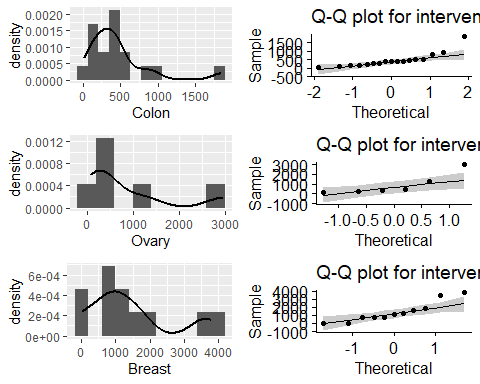
At 0.01 original significant level, all methods detect consistent combinations of comparison, which means all of 3 methods give the same result. For example, take ‘Breast’ group as reference, all methods detect significant difference in Bronchus-Breast comparison, Colon-Breast comparison and Stomach-Breast comparison.

The difference of 3 methods is P-value. Although they get the same result, their p-value is different. Moreover, according to Table 3, 4, 5, Bonferroni adjustment is the most conservative than others with the largest relative P-values. Dunnett adjustment is the most sensitive one with the smallest relative P-values.

## d)

According to Plot 3, all distributions are extremely skewed, and points in all the Q-Q plots are not linear except Colon group. And then Shapiro-Wilk test is performed to test normality as well.





Plot 3 Distributions and Q-Q plots for each Organ group

Shapiro-Wilk test

Hypothesis:

The data are sampled from a population having a normal distribution,

The data are sampled from a population not having a normal distribution

Table 6 P-value of Shapiro-Wilk test for each group

|  |  |
| --- | --- |
| **Data** | **p-value** |
| Stomach | 0.002075 |
| Bronchus | 0.0007186 |
| Colon | 0.0006134 |
| Ovary | 0.029 |
| Breast | 0.07431 |

Conclusion:

From the Table 6, at 0.05 significance level, we reject the null at conclude that there are sufficient evidence that all the distributions except Breast group are not normal. As a result, the underlying assumption of ANOVA is questionable.

### i)

Transformation can be done to transform the data into normality.

Kruskal-Wallis (KW) test can be used when the normality is questionable.

### ii)

Kruskal-Wallis (KW) test:

Kruskal-Wallis公式

## Kruskal-Wallis chi-squared = 14.954, df = 4, p-value = 0.004798

Since, we reject the null hypothesis with 0.01 significance level, and we conclude that there is at least two of group medians are different.

ANOVA analysis:

The ANOVA analysis can be referred from the Problem 2 b).

Table 2 ANOVA table

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Source** | **Df** | **Sum of Square (SS)** | **Mean Sum of Square** | **F-Statistics** | **P-value** |
| Between | 4 | 11535761 | 2883940 | 6.433 | 0.000229 |
| Within | 59 | 26448144 | 448274 |  |  |
| Total | 63 | 37983905 |  |  |  |

Both of tests reject the null and detect the difference of 5 groups. However, the ANOVA implements means while Kruskal-Wallis (KW) test implements medians. Furthermore, the P-value of Kruskal-Wallis (KW) test is larder than ANOVA’s, which means the ANOVA analysis may inflate the difference due to its questionable underlying assumption.

# Problem 3

## a)

Assume that placebo and iron groups are denoted as Group 1 (non-zinc group), zinc and zinc+ iron groups are denoted as Group 2 (non-zinc group). And let a, b, c, d denote the original groups placebo, iron, zinc, zinc+iron respectively.

Let denotes the average of Group 1, denotes the average of Group 2.

,

,

,



## b)

Since, , then we can calculate the standard deviation for each original group:

,

,

,



F test can be conducted to detect if the *sd* in Group placebo (a) and iron (b) are equal.

Hypothesis:



With the significant level pre-specified with 0.05, compute the test statistics:

,

 ,

Since , at 0.05 significant level, we fail to reject the null and conclude that the variances of placebo and iron groups are equal.

As the test above, the variances of Group zinc (c) and Group zinc+iron (d) can be tested similarly.

Hypothesis:



With the significant level pre-specified with 0.05, compute the test statistics:

,

 ,

Since , at 0.05 significant level, we fail to reject the null and conclude that the variances of zinc and zinc+iron groups are equal.





## c)

### i)

We have known that ‘true’ effect size is , the ‘true’ standard deviations are , , , , and , in the study two-sided significance test is used：





We need to enroll 162 subjects for each group.

### ii)

We have known that ‘true’ effect size is , the ‘true’ standard deviations are , , , , and , in the study two-sided significance test is used:





We need to enroll 170 subjects for non-zinc group and 86 subjects for zinc group.