SECI2143 PROBABILITY AND STATISTICAL DATA ANALYSIS
PROJECT II – STATISTICAL ANALYSIS

Relationship between Deaths Caused by Ischaemic Heart Disease and Gender from the year 2013 to 2018 in Malaysia

Lokessh A/L Pathmanatan
Section 10

Abstract - Ischemic heart disease, also called coronary heart disease (CHD) or coronary artery disease, is the term given to heart problems caused by narrowed heart (coronary) arteries that supply blood to the heart muscle. Although the narrowing can be caused by a blood clot or by constriction of the blood vessel, most often it is caused by build-up of plaque, called atherosclerosis. Thus, the relationship between deaths for males and females is being tested in this project. From the inferential statistics, the results clearly show that the number of deaths for males is clearly higher compared to the females throughout all the years.

Introduction

Ischemia is defined as inadequate blood supply (circulation) to a local area due to blockage of the blood vessels supplying the area. Ischemic means that an organ (e.g., the heart) is not getting enough blood and oxygen. When the blood flow to the heart muscle is completely blocked, the heart muscle cells die, which is termed a heart attack or myocardial infarction (MI). Most people with early (less than 50 percent narrowing) CHD do not experience symptoms or limitation of blood flow. They are most likely to occur during exercise or emotional stress, when the demand for the oxygen carried by the blood increases. This is a clinical syndrome characterized by discomfort in the chest, jaw, shoulder, back, or arms that is typically aggravated by exertion or emotional stress and relieved promptly with rest or by taking nitro-glycerine. Angina usually occurs in patients with CHD, but also can occur in individuals with valvular disease, hypertrophic cardiomyopathy, and uncontrolled hypertension. Infrequently, patients with normal coronary arteries may experience angina related to coronary spasm or endothelial dysfunction. The objective of this project is to determine the number of deaths between males and females from the year 2013-2018 in Malaysia.

What do I want to Study?

This report is about deaths between males and females in Malaysia from the year 2013-2018. I am doing a research on why the ischaemic heart disease continues to cause more and more deaths year after year. This one disease caused deaths has increased up to 60% in the past decade. My main aim is to determine the relationship between gender and the number of deaths caused by the ischaemic heart disease. I would also like to study on the reasons behind the causes of this disease and the necessary prevention that needs to be taken to overcome the whole situation and maybe even prevent it from happening in the first place.

Methodology

To carry out an inferential statistic, I am going to use the secondary data that I obtained from the Department of Statistics Malaysia. The target people are the males and females' citizens of Malaysia. Inferential statistic is carried out by using hypothesis testing of two samples, correlation, regression and chi square test of independence.

Results and Discussion

Hypothesis testing (2 samples)

Two sample tests on deaths caused by the Ischaemic Heart disease between each gender from 2013 – 2018.

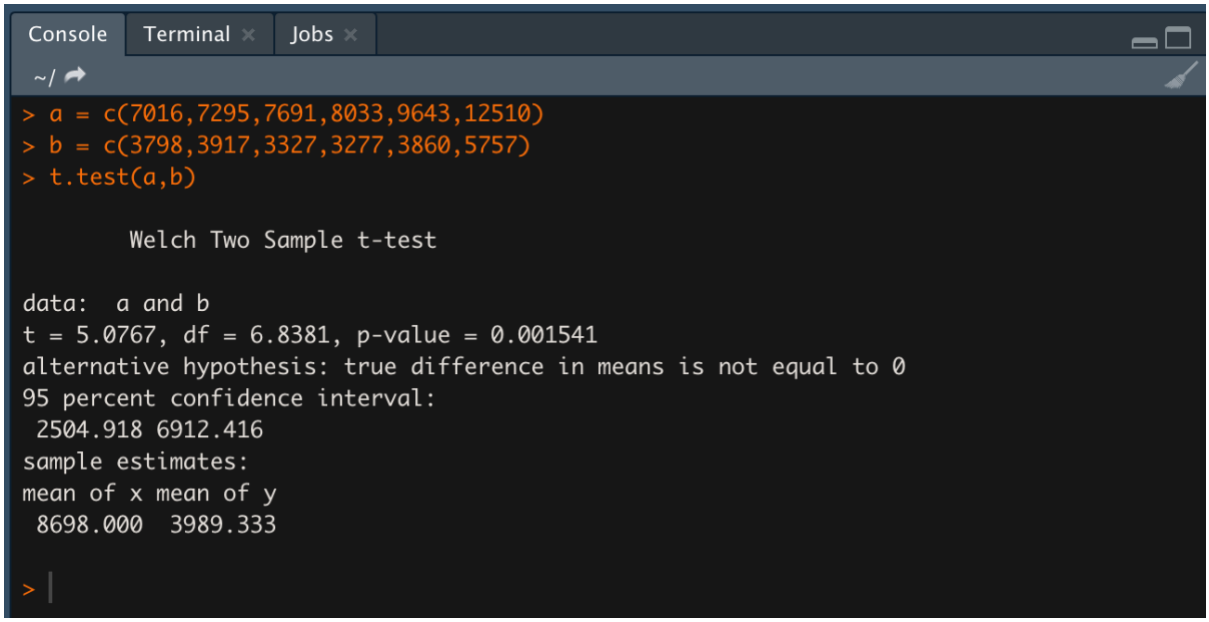
Let μ1 = sample mean of male of death cause by the heart disease from 2013 to 2018.

Let μ2 = sample mean of female of death cause by the heart disease from 2013 to 2018.

$H_0 : \mu_1 - \mu_2 = 0$

$H_1 : \mu_1 \neq \mu_2$

Significance level, α = 0.10

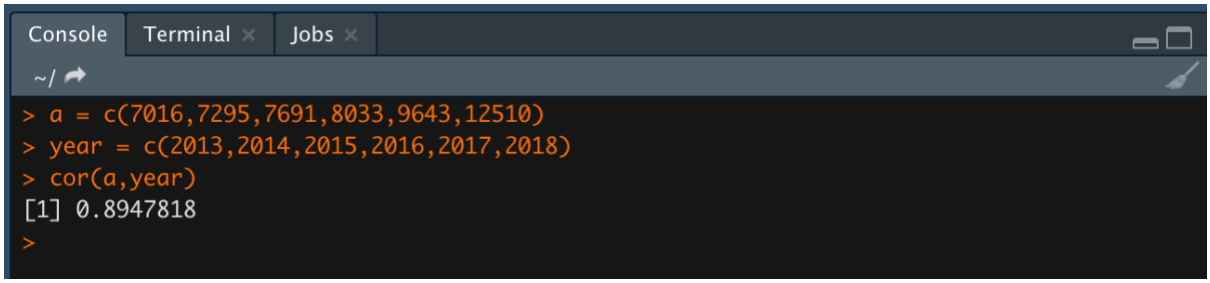


Variable a is the dataset for males and variable b is the dataset for females. Since the size of the sample is quite large, we can assume both samples are normally distributed. This is a two-tailed test, so there are two rejection area of P-value which is left and right of the test statistic, z = 5.0767. The P-value for test statistic is < 0.01. Since P-value less than α = 0.10, thus we reject the null hypotheses. We can conclude that the sample mean of males and females are not the same.

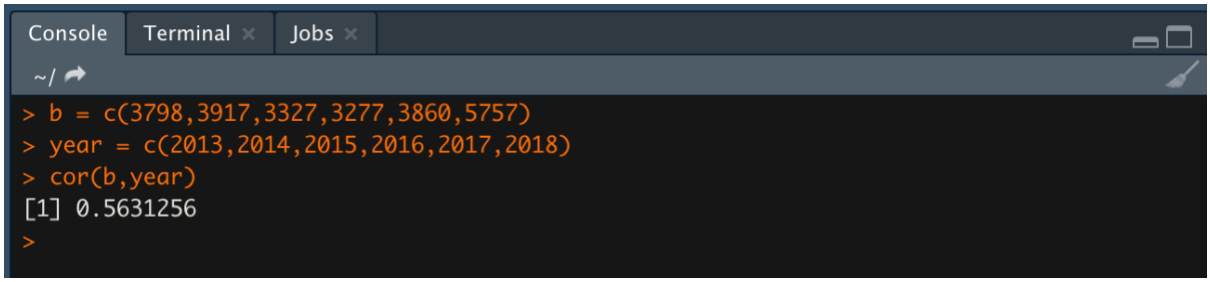
Table

Year	Male	Female
2013	7016	3798
2014	7295	3917
2015	7691	3327
2016	8033	3277
2017	9643	3860
2018	12510	5757

Correlation



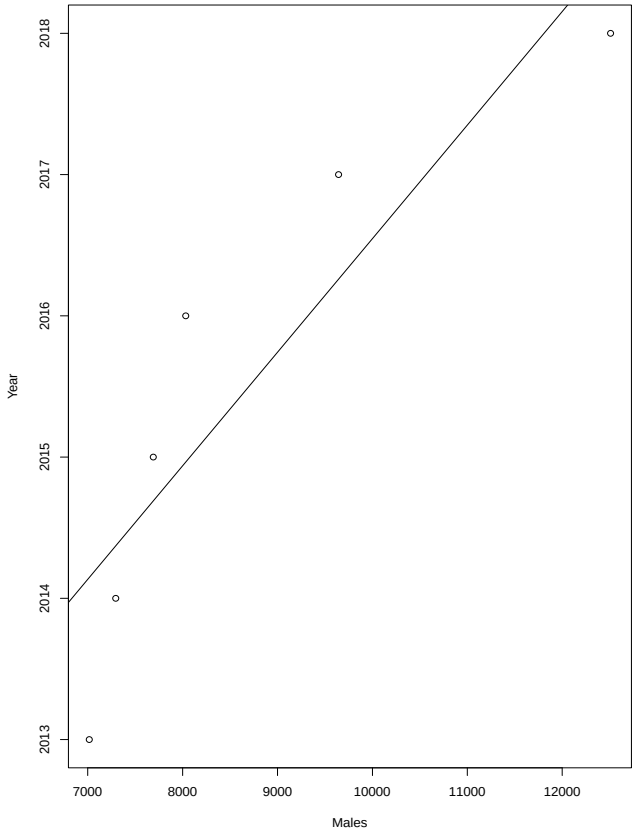
⊗ Based on the computed correlation value for the Ischaemic Heart Disease Deaths in Males across the years which is 0.8947818, it can be concluded that there is a strong positive relationship between the number of male deaths with the respective years.



⊗ Based on the computed correlation value for the Ischaemic Heart Disease Deaths in Females across the years which is 0.5631256, it can be concluded that there is a moderately strong relationship between the number of female deaths with the respective years.

By comparing these 2 values, there is sufficient evidence that the number of heart disease deaths for males has a much higher influence with the respective years compared to the females.

Regression

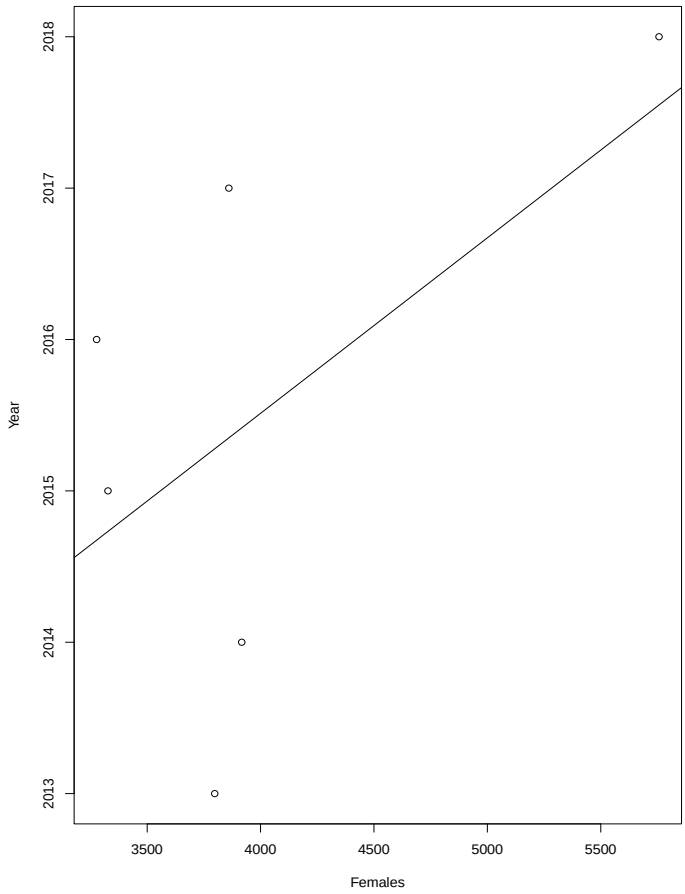



```
Console   Terminal ×   Jobs ×
~/
> plot(a,year, ylab="Year", xlab="Males")
> abline(lm(year~a))
> model = lm(year~a)
> model

Call:
lm(formula = year ~ a)

Coefficients:
(Intercept)            a
  2.009e+03    8.039e-04

>
```




```
Console   Terminal ×   Jobs ×
~/
> plot(b,year, ylab="Year", xlab="Females")
> abline(lm(year~b))
> model = lm(year~b)
> model

Call:
lm(formula = year ~ b)

Coefficients:
(Intercept)            b
  2.011e+03    1.159e-03

>
```

Based on the graphs that have been generated, a simple regression model is produced. The regression models are positive linear models, which have a straight-line relationship. Also through regression, we can predict the number of heart disease deaths in both genders are gradually increasing. The deaths in males are also increasing in a much quicker rate compared to the females.

y_{male} = 2.009e+03 + 8.039e-04x

⊗ From the equation, we can see that the estimated change in average number of deaths for males in increasing by 8.039e-04. There is also no zero deaths for all the years. 2.009e+03 is the portion of number of male deaths from 2013 - 2018.

y_{female} = 2.011e+03 + 1.159e-03x

⊗ From the equation, we can see that the estimated change in average number of deaths for females in increasing by 1.159e-03. There is also no zero deaths for all the years. 2.011e+03 is the portion of number of female deaths from 2013 - 2018.

By comparing these 2 values of coefficient of determination, we have enough evidence to conclude that the number of heart disease deaths for females in nearer to the mean throughout all the years compared to the deaths for males.

Chi square test of independence

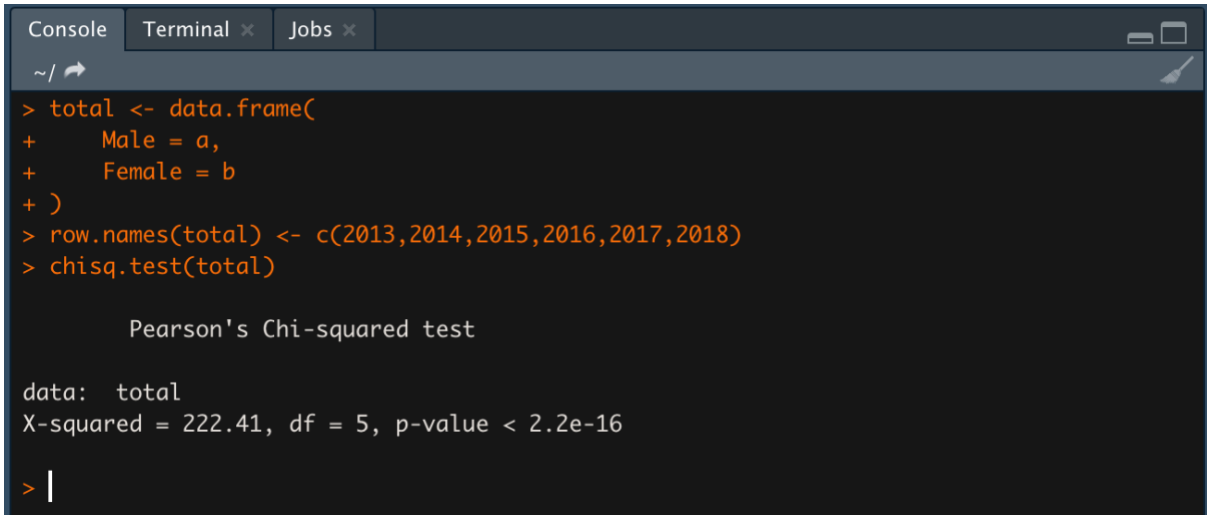
Book3

Filter

	Year	Male	Female
1	2013	7016	3798
2	2014	7295	3917
3	2015	7691	3327
4	2016	8033	3277
5	2017	9643	3860
6	2018	12510	5757

H_0 : The number of heart disease deaths for each gender is independent of the year.

H_1 : The number of heart disease deaths for each gender is dependent of the year.



Significance level, α = 0.10; $\mathcal{X}^2_{0.1,5}$ = 9.236

⊗ Since $\mathcal{X}^2$ = 222.41 > 9.236, we reject H_0. We have sufficient evidence that the number of heart disease deaths for each gender is dependent based on the respective year at significance level of α = 0.10.

To kind of wrap it all up, the results clearly show that there is a close relationship between the genders and the rate of these deaths. I believe that all the deaths caused by this disease should really not be taken lightly. The lack of appropriate actions are causing the deaths to have an endless raise. Progressive and effective measures must be taken immediately to ensure we can tackle these problems at the root cause. These changes will not be seen immediately, but this is more of a long-term problem, that needs step-by-step prevention. We also need to start educating the public about the early signs of these deadly diseases and what they can actually do, instead of panicking and worsening the situation.

Conclusion

Based on all the results of the report study, it is clear that we have come to certain conclusions. First of all, we can clearly see that the Ischaemic Heart Disease affects both genders in a different way, with males often being the majority and females the minority. Besides that, we also managed to conclude that the heart disease for both genders is also dependent based on the respective years. The mean deaths of females is also closer throughout all the years, unlike the males. This also implies that the number of heart disease deaths for females can easily alter depending on the year as there are rises and falls of deaths. Finally, we can see that the rate of heart disease deaths in total is not showing any signs of slowing down, which means that this upwards trend can also continue for the upcoming years. And we can clearly see that the rate of death in males is rapidly increasing compared to the death in females.