City University of Hong Kong

2019-2020 Sem B

EE3211 Modelling Techniques

Exercise project

Topic 1

April 18,2020

(Deadline: End of Week 13, i.e. April 24, 2020)

Background

Diabetes mellitus (DM), also known as diabetes. According to international diabetes Atlas (2019), diabetes caused 4.2 million deaths total and at least USD 760 billion dollars in health expenditure in a year.

Yet, various factor may associate with diabetes, example of blood pressure, drinking habits may affected to infection of diabetes. Diabetes was also known as a bad life habit disease. People who have a poor life habit may easier to get diabetes. In fact, a hypothesis was made to recognize this issue.

Objective

This report is aims to identify several body-factors (Ages, weight, cholesterol level etc.) and life habit would or not associated with diabetes. In following parts, data will be collected from The National Health and Nutrition Examination Survey (NHANES) Diabetes Questionnaire Data, Body Measures Data, Blood Pressure & Cholesterol Questionnaire Data, Alcohol Use Questionnaire Data (2015-2016).

Several methods will be done for testing the relationship between diabetes and variables. sample ID," have diabetes", Weight, Height, and BMI, "had high blood pressure", "had high cholesterol level" "Had at least 12 alcohol drinks/1 year", and "alcoholic drinks/day - past 12 months" "Overweight" were collected to analysis.

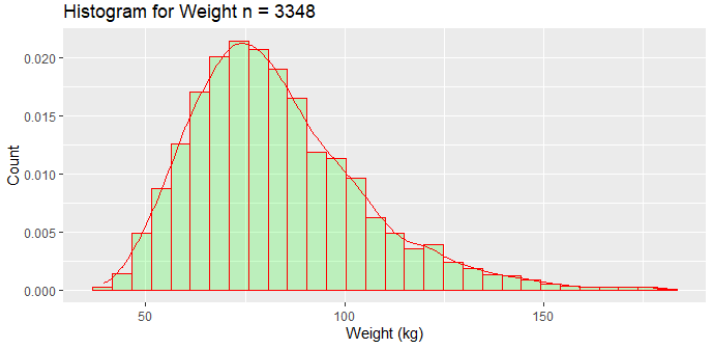
Methods

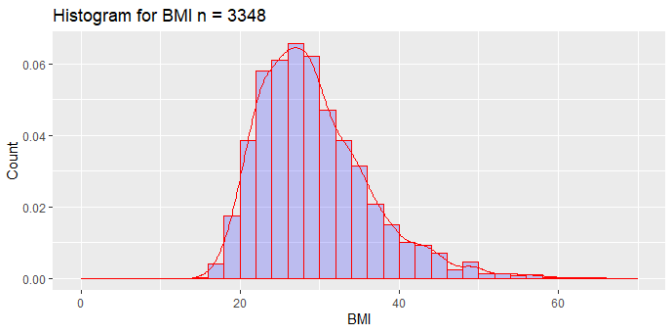
Data:

3348 valid data were collected with 9 variables, all missing, NA, refuse to answer were excluded from dataset. A summary and used code were established in appendix 1,2. A variable "Overweight" is according to subjects BMI, 1 represent >30BMI and 0 represent <30 BMI. Any correlation will be tested within these data.

Test method:

Chi-Square Test, relative risk and logistic regression method will be used for identifying the relationship between having diabetes and other variables. Chi-Square Test, relative risk will be used for overweight and diabetes. Logistic regression method used for identity all the variables associated with diabetes.





Result

A null hypothesis was assumed that there is no association between diabetes and overweight. Chi-Square Test is used to figure out the hypothesis in table 1. [X-squared = 72.089, df = 1, p-value = 2.2e-16 < 0.05]. The null hypothesis is rejected, and we accept the alternative hypothesis. There is an association between overweight and diabetes.

Chi-Square Test	diabetes		
overweight	Yes	No	total
Yes	227	162	289
No	1063	1896	2959
total	1290	2058	3348

Table 1. Chi-Square table for diabetes and overweight.

In additions, a relative risk test was also calculated with overweight and diabetes. A result of 2.186 was calculated by (227/289) / (1063/2959). It reflected people who overweight had 2.186 times the risk of diabetes compared to subjects who did not overweight.

A logistic regression method is used for identifying the relationship between diabetes and other variables. A null hypothesis was assumed that there is no association between diabetes and a variable. In table 2, p-value with a value small than 0.05 reflected a rejection of null hypothesis and accept the alternative hypothesis, these is an association between diabetes. >0.05 reflected a accept of null hypothesis. As a result, overweight, high blood pressure and high cholesterol level are associated with diabetes.

diabetes	y = c + mx / signal = βo + β1×conc βo = theoretical y-intercept β1 = theoretical slope		
Risk factor	βo	β1	p-value
Height	1.9454016	-0.0002191	0.7497
overweight	1.94023	-0.08209	**1.188e-09**
blood pressure	1.6823	0.1346	**< 2.2e-16**
cholesterol level	1.83219	0.04459	**3.207e-06**
alcohol taking. (every 12 / 1 yr)	1.92271	-0.01264	0.5382
alcohol taking. (avg pre day)	1.895343	0.004844	0.08549

Table 2. logistic regression method

Conclusion

Among all the data and tests, a conclusion that overweight is one of the main factors associated with diabetes. In all cases, overweighed can be a main reason for infected diabetes. In additions, overweighed people are more likely to have diabetes compare to people who are not overweighed. This can lead to a result that overweighed could associated with diabetes.

Apart from the above result, high blood pressure and high cholesterol level are also associated with diabetes. A special mention about alcohol taking per day, although the p-value is bigger than 0.05 but close to 0.05, no strong evident to justify an association between alcohol taking and diabetics. A study was found an association between over drinking alcohol and overweight (Traversy, G., & Chaput, J.-P ,2015). Overweighed, high blood pressure and high cholesterol level or alcohol taking are mostly from a poor life habit or genetic disease. In former, a good eating habit and health lifestyle can reduce the possibility to have overweight. A good rest cycle, habit can reduce chances of high blood pressure and high cholesterol level. Drink less or none alcohol may consider a good life habit also.

To prevent having diabetes, it is important to reduce the chance of having high blood pressure, high cholesterol level and overweight. It is suggested to have a better eating habit and lifestyle to reduce those three factors. In especially tackle with overweighed. Future study should address the relationship between diabetes and these 3 factors or more.

References:

IDF Diabetes Atlas 9th edition 2019. (n.d.). Retrieved from https://www.diabetesatlas.org/en/
Appendix 1: summary of dataset

Traversy, G., & Chaput, J.-P. (2015, March). Alcohol Consumption and Obesity: An Update. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4338356/