Homework 5

Aji John

Winter Quarter 2018

Question 1

Perform a Poisson regression analysis to evaluate an association between 5 year all-cause mortality and creatinine by comparing the relative risk of death (or risk ratio of death) across groups defined by continuous serum creatinine level. (Only provide a formal report of inference when asked to.)

(a)

Provide an interpretation of the slope and the intercept in the Poisson regression model, and include the numerical values of the slope and intercept in your interpretation.

Intercept when exponentiated is the base risk. The numerical value is 0.057.

(b)

Give full inference for an association between 5 year all-cause mortality and serum creatinine levels from the Poisson regression model.

From Poisson regression analysis, we estimate that the relative risk of death in 5 years due to CRT increases by 156% for every 1 unit increase in CRT, and this increase is highly significant ($P < 0.0001$). A 95% CI suggests that this observation is not unusual if the true percent increaseof the more recent CRT is anywhere from 93.8% to 239.9% higher than the previous CRT level.

(c)

Compare the association results in part b that are based on risk ratios to using a logistic regression model where odds ratios of death within 5 years are used as the summary measure for an association with serum creatinine level (i.e., question 3 in homework 4). Briefly describe any similarities or differences in the association results.

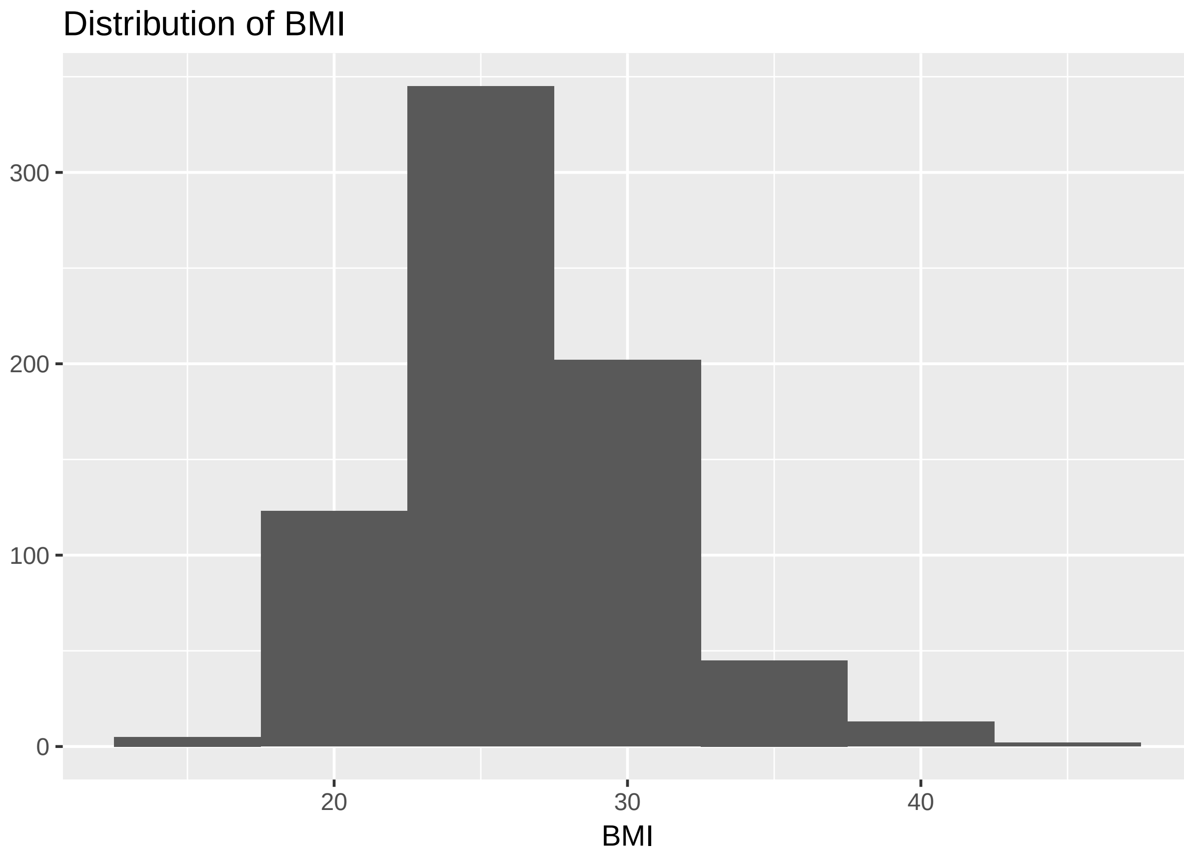
Odds radio is 5.986 (When exponentiated) in logistic regression, and when used in POisson, it is 2.564. We have the probability of dying in 5 years with low CRT is 0.13 and with high CRT is 0.26, as the probability of the event is nearly double, it is not rare.

Question 2

Questions 3 and 4 below investigate associations between serum cholesterol level, age, sex, and body mass index (BMI). In this question we will obtain some summary statistics for these variables.

(a)

Create a variable for BMI using the height and weight measurements on the subjects. [Hint: Make sure that appropriate conversions of the weight and height measures are used in the calculation of BMI]. Provide a figure illustrating the distribution of BMI in the sample.



(b)

Provide suitable descriptive statistics for serum creatinine levels, age, sex, and BMI.

Variable	Male	Female	All subjects
Sample size	n = 121	n = 614	n = 735
CRTs[1]	1.2 (0.3); 0.7-4	0.9 (0.3); 0.5-3.2	1.1 (0.3); 0.5-4
Age (years)[1]	74.7 (5.6); 66-99	74.4 (5.3); 65-91	74.6 (5.5); 65-99
BMI (kg/mm)[1]	26.3 (3.8); 16-42	26.5 (4.9); 15-47	26.4 (4.3); 15-47

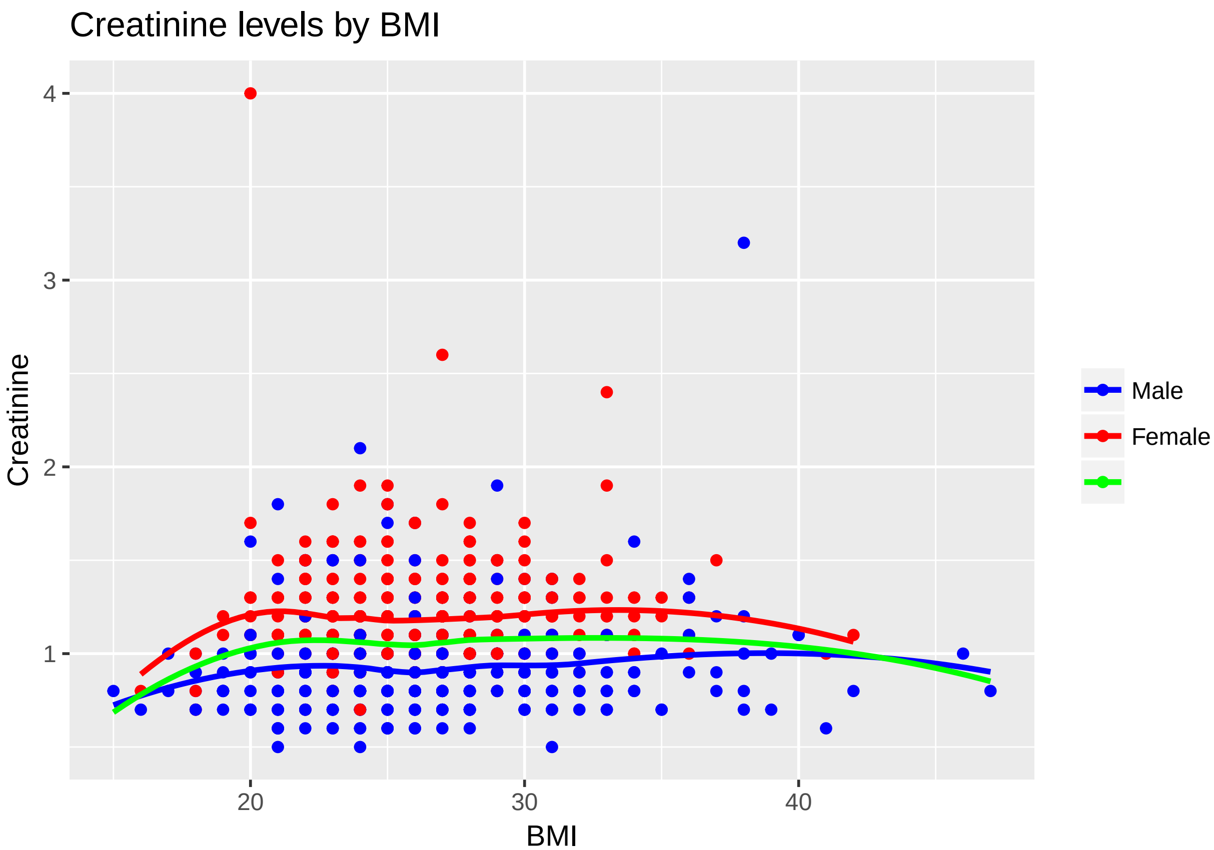
[1] mean (sd); min-max are reported

Question 3

We are interested in examining how mean serum creatinine levels vary by BMI and sex. In the questions below, you do not need to provide full statistical inference. Instead, just answer the following questions.

(a)

Create a scatterplot of serum creatinine levels versus BMI. Use different symbols and/or colors for each sex group, and include LOWESS (or LOESS) curves for each sex group.



(b)

What observations do you make from the scatterplot in part a regarding the association between serum creatinine levels and BMI?

- Some outliers between 25 and 35 range of bmi
- Sparse observations between 40 and 50
- Variance not the same across different BMI levels
- Linear relationship looks suspicious, need to validate it

(c)

Is there evidence from descriptive statistics (question 2) and the scatterplot in part a that sex modifies the association between serum creatinine level and BMI? Explain your reasoning.

We have SD for mean CRT between males and females is 0.3, and the mean is 1.2 and 0.9 for males/females respectively. We see that the estimated mean CRT for female is within one SD of male subjects, so, it is very likely that the means are same across both the age groups. Also, by looking at the scatter plot the lines do not cross, so it is not an effect modifier.

(d)

Is there evidence from descriptive statistics (question 1) and the scatterplot in part a that sex confounds the association between serum creatinine level and BMI? Explain your reasoning.

No, there is no evidence of sex confounding association between CRT and BMI. From the scatter plot it is not marginally different when conditionally plotted on sex(see above figure). Additionally,lines nearly follow the same trend, and are not parallel for each bmi group.

(e)

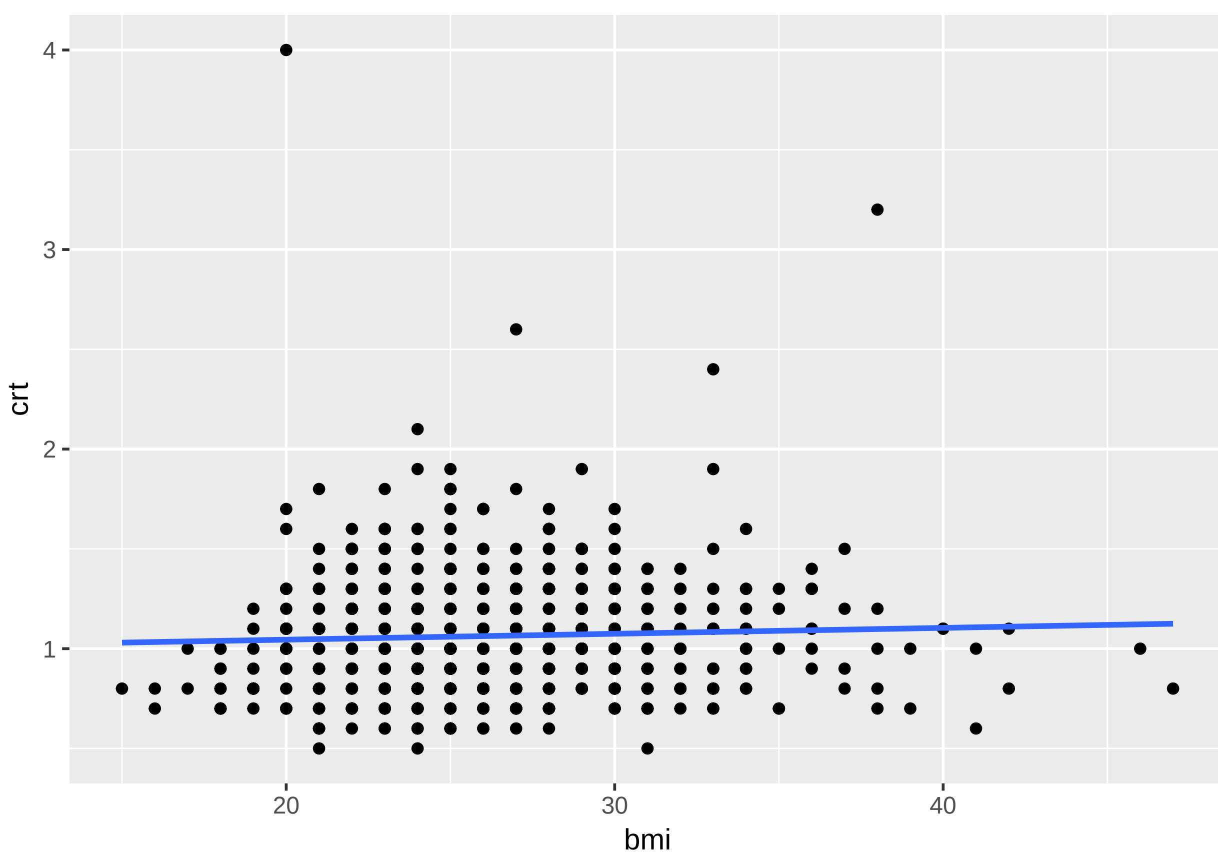
Perform an analysis to determine whether mean serum creatinine levels differ across sex groups. Briefly describe the analysis that you performed and clearly state the basis of your conclusion regarding an association.

Using t.test with unequal variance setting we find that mean CRT is not equal i.e. they differ across the sex groups. The p-value we find is < 0.05, and is found to be significant.

(f)

Perform an analysis to determine whether there is a linear trend in mean serum creatinine levels by BMI. Briefly describe the analysis that you performed and clearly state the basis of your conclusion regarding an association.

Question 2:



We find that slope is 3.18 * 10^-3 which at p-value of .3113 is not significant.

(g)

Perform an analysis to determine whether mean serum creatinine levels differ across sex groups after adjustment for BMI. Briefly describe the analysis that you performed and clearly state the basis of your conclusion regarding an association.

We find that after adjusting for sex, the slope is found to be still not significant (p-value 0.1928)

(h)

Perform an analysis to determine whether there is a linear trend in mean serum creatinine levels by BMI after adjustment for sex. Briefly describe the analysis that you performed and clearly state the basis of your conclusion regarding an association.

Linear trend is there as slope is .003813, but is found to be not significant(p-value 0.192).

(i)

Perform an analysis to determine if sex modifies the association between mean serum creatinine levels and BMI. Briefly describe the analysis that you performed and clearly state the basis of your conclusion regarding an association.

We used the below model (crt ~ bmi*male) to explore the effect modification. The interaction (bmi:male) is not found to be present as p-value is 0.5261.

(j)

How would you summarize the association between serum creatinine levels and BMI and sex? Provide a summary of your findings that is suitable for inclusion in a manuscript.

For each unit(kg/m^2) difference in bmi between two groups that have the same sex, the difference in mean CRT is 0.00523(95% CI:-0.00206 to 0.01254, with the group with higher bmi having a miniscule change in CRT. The results are highly atypical of what we might expect if there was no true difference in mean CRT levels between two bmi groups that have the same sex. (P-value is .159 which is found to be not significant at alpha = 0.05).

Between two sex groups that have the same bmi, the difference in mean CRT is 0.3703(95% CI:0.0532 to 0.6873, with the males having a higher CRT. The results are highly atypical of what we might expect if there was no true difference in mean CRT levels between two sex groups that have the same bmi (P-value is 0.02 which is found to be significant at alpha = 0.05).

Additionally, we find that there is no evidence of gender having an effect on the association between CRT and bmi.

Question 4

Now consider a multivariate linear regression analysis with serum creatinine level as the response and the variables age, sex, and BMI as predictors.

(a)

Provide an interpretation of the intercept in the regression model. Is the slope estimate scientifically useful?

The intercept is the estimate of the mean CRT level for new born female who has 0 bmi. It is not scientifically relevant, and is not possible to have it in real life.

(b)

Give full inference for the age slope in the regression model.

The age slope is the estimated change in mean CRT for one unit change (1 yr) keeping sex and bmi constant.

For each year difference in age between two groups that have the same sex and bmi, the difference in mean CRT is .005.823(95% CI:.001401 to 0.010), with the older group having a higher CRT. The results are highly atypical of what we might expect if there was no true difference in mean CRT levels between age groups that have the same sex and bmi(P-value is 0.0099)

(c)

Give full inference for the sex slope in the regression model.

The male slope is the estimated change in mean CRT between two sex groups by keeping age and bmi constant.

Between two sex groups that have the same age and bmi, the difference in mean CRT is 0.268(95% CI:0.229 to 0.306), with the males having a higher CRT. The results are highly atypical of what we might expect if there was no true difference in mean CRT levels between two sex groups that have the same age and bmi(P-value is < 0.00005)

(d)

Give full inference for the BMI slope in the regression model.

The bmi slope is the estimated change in mean CRT level for one unit change in bmi(mg/dl) by keeping male and age constant.

For each unit(kg/m^2) difference in bmi between two groups that have the same sex and age, the difference in mean CRT is .005351(95% CI:-.00006427 to 0.01077, with the group with higher bmi having a very miniscule change in CRT. The results are highly atypical of what we might expect if there was no true difference in mean CRT levels between two bmi groups that have the same sex and age. (P-value is 0.0528 which is found to be not significant at alpha = 0.05)