Biostat 518

Homework #2

Total Score: 159/195

1. Descriptive statistics of CRP and fibrinogen

Mean Fibrinogen levels by CRP categories

|  |  |
| --- | --- |
|  |  CRP mg/L |
|  | <1mg/L | 1-3mg/L | > 3mg/L |
| CVD history (n)1 |  |  |  |
| All (4899)2 | 294.6 (50.2, 109-540) | 314.3 (51.5, 138-482) | 357.9 (75.2, 132-872) |
| No prior CVD (3777)2 | 293.2 (49.2, 109-482) | 313.6 (51.2, 183-482) | 354.2 (72.4, 132-872) |
| Prior CVD (1122)2 | 300.8 (53.9, 171-540) | 316.7 (52.5, 138-470) | 367.7 (81.5, 175-695) |

1 Cardiovascular History (CVD) with (# participants)

2 Descriptive statistics presented represent mean (SD, min-max)



Methods: The relationship between serum crp levels and fibrinogen levels was explored using descriptive statistics. Initially a two-way scatter plot was created to look at general relationship between the two continuous variables. Subgroups of CRP were then created according to existing recommendations for risk of cardiovascular disease by CRP level: below 1mg/L (low risk of heart disease), 1-3 mg/L (average risk of heart disease), greater than 3mg/L (high risk of heart disease). Data is presented according to mean fibrinogen levels in each of these categories. A similar analysis was performed in the subgroup of participants with prior CVD and without prior CVD. Finally, mean fibrinogen and crp levels were examined after stratification by prior CVD.

Results: As seen in the above table, as serum CRP levels increase, mean fibrinogen levels also tend to increase. As illustrated in the scatter plot, this correlation seems to exponentially rise with the more extreme values of serum fibrinogen and crp. In addition, there is heteroscedacity with variance in data increasing with more extreme data. In the total 4899 participants, mean fibrinogen levels were 323.0 mg/dL. Stratified by prior CVD: in the 3777 participants without prior CVD, mean fibrinogen levels were 319.6mg/dL while in the 1122 participants with prior CVD, mean fibrinogen levels were 334.5mg/dL. In the total study participants, mean crp level was 3.61 mg/L. Stratified by cardiac history, those without prior CVD had mean CRP level of 3.38 mg/L and those with prior CVD had mean crp of 4.40 mg/L. Since prior CVD history is associated with both fibrinogen levels and crp levels, it is a possible confounding variable.

2. T-tests

a. Method: Mean serum fibrinogen levels were compared between participants with a prior history of CVD and participants without a prior history of CVD. The differences in the mean fibrinogen levels were compared using a t-test that assumed equal variance between the groups. Confidence intervals were obtained using similar methods.

Results: In the 3791 participants without a history of prior CVD, the mean fibrinogen level was 319.57mg/dL. In the 1124 participants with a history of prior CVD, the mean fibrinogen level was 334.46mg/dL. The difference in mean fibrinogen levels between the two groups was 14.89mg/dL. With 95% confidence, we can say that participants with a prior history of CVD had a mean fibrinogen level between 10.42 and 19.34mg/dL higher than participants without prior CVD history. The t-test gives a p value of < 0.001, meaning that we can reject the null hypothesis that the mean fibrinogen levels are the same between participants with and without prior cardiovascular disease.

b. The same analysis in part a can be done using classical linear regression which assumes equal variance between the 2 groups. Here, the intercept of 314.57 represents the point estimate for mean fibrinogen level in participants without prior CVD. The model gives us a slope of 14.86mg/dL; this is the same as the central tendency of the difference in mean fibrinogen levels between the participants with prior history of CVD and those without prior history of CVD. Similarly the 95% confidence intervals (10.42, 19.35) are the same as those obtained from the t-test that assumes equal variance because the two statistical analyses have the same standard error. The p-values in both analyses are very statistically significant at < 0.001.

c. Method: Mean serum fibrinogen levels were compared between participants with a prior history of CVD and participants without a prior history of CVD. The differences in the mean fibrinogen levels were compared using a t-test that did not assume equal variance between the groups. Confidence intervals were obtained using similar methods.

Results: In the 3791 participants without a history of prior CVD, the mean fibrinogen level was 319.57mg/dL. In the 1124 participants with a history of prior CVD, the mean fibrinogen level was 334.46mg/dL. The difference in mean fibrinogen levels between the two groups was14.89mg/dL. With 95% confidence, we can say that participants with a prior history of CVD had a mean fibrinogen level between 10.09 and 19.68mg/dL higher than participants without CVD. The t-test gives a p value of < 0.001, meaning that we can reject the null hypothesis that the mean fibrinogen levels are the same between participants with and without prior cardiovascular disease.

d. Similarly, we can do the same analysis in part c with a linear regression that does not assume equal variances to calculate the standard error. The model gives us the exact same point estimates for the intercept and the slope as the t-test gives for the mean fibrinogen level in participants with no prior CVD history and central tendency of mean difference in fibrinogen levels between the groups. The robust SE is 2.44 which is also the same as the t-test SE, resulting in the same 95% confidence interval for the difference in mean fibrinogen levels between the groups (10.42 and 19.68). The p-values in both analyses are very statistically significant at < 0.001.

e. The t-test that assumes equal variances (part a) and the t-test that does not assume equal variances (part c) will give the same point estimates for mean fibrinogen levels in the two groups and difference in mean fibrinogen between the two groups. However, the t-test that assumes equal variances is based on 4913 degrees of freedom while the t-test that does not assume equal variances is based on only 1664.57 degrees of freedom. This results in a larger p-value and wider confidence intervals in the t-test that does not assume equal variances. Of note, because the sample size is very large, the difference in the CI between the 2 analyses is relatively small.

3. Fibrinogen and CRP

Methods: The relationship between serum fibrinogen and untransformed crp levels was analyzed using a simple linear regression. Standard errors and confidence intervals were obtained using robust estimates.

(see part 3 for results of statistical analysis)

a. Classical linear regression given an intercept of 304.015. This is the estimated value of mean fibrinogen level when serum crp is 0mg/L.

b. Classical linear regression given an estimated slope of 5.251 mg/dL. This is the estimated change in mean fibrinogen level for every 1mg/L increase in serum crp levels.

c. From linear regression analysis, we estimate that for each 1mg/L increase in mean serum crp, there is a 5.251mg/dL increase in mean serum fibrinogen level. A 95% CI suggests that this observation would not be unusual if the true difference in mean fibrinogen level per 1mg/L increase in crp were between 4.604mg/dL and 5.898mg/dL. Because the two sided P value is < 0.001 we reject the null hypothesis that there is no linear trend in average fibrinogen level and crp.

d.

|  |  |
| --- | --- |
| CRP level | Fitted mean Fibrinogen (mg/dL) |
| 1mg/L | 309.266 |
| 2mg/L | 314.517 |
| 3mg/L | 319.768 |
| 4mg/L | 325.019 |
| 6mg/L | 335.520 |
| 8mg/L | 346.022 |
| 9mg/L | 351.273 |
| 12mg/L | 367.025 |

4.

Methods: Geometric mean crp levels were generated by log transforming serum crp. Of note, to avoid the error of generating inaccurately missing data when calculating geometric means, participants whose serum CRP levels were recorded as 0 were replaced by 0.5mg/L. This was based on the assumption that participants did not actually have a CRP level of 0mg/L, but in reality a smaller value then was clinically assayable by laboratory techniques. The relationship between serum fibrinogen and log transformed crp levels was analyzed using a simple linear regression with robust estimates of SE and 95% CI.

Results: From linear regression analysis, we estimate that for each 1mg/L increase in log transformed serum crp, there is a 36.833 mg/dL increase in mean serum fibrinogen level. A 95% CI suggests that this observation would not be unusual if the true difference in mean fibrinogen level per 1mg/L increase in log transformed mean crp levels were between 34.577mg/dL and 39.089mg/dL. Because the two sided P value is < 0.001 we reject the null hypothesis that there is no linear trend in average fibrinogen level and log transformed crp levels.

To get the results for the table below, log transformation of CRP was placed into the model: E(Fibrinogen| crp) = βo + (log crp)β1

|  |  |
| --- | --- |
| CRP level | Fitted mean Fibrinogen (mg/dL) |
| 1mg/L | 295.566 |
| 2mg/L | 321.097 |
| 3mg/L | 336.032 |
| 4mg/L | 346.628 |
| 6mg/L | 361.562 |
| 8mg/L | 372.159 |
| 9mg/L | 376.497 |
| 12mg/L | 387.093 |

5.

Methods: The relationship between log-transformed mean serum fibrinogen and crp levels was analyzed using a simple linear regression with robust estimates of SE and 95% CI. Estimates of the slope were exponentiated to determine the percent change in geometric mean fibrinogen.

Results: From linear regression analysis, we estimate that for each 1mg/L increase in serum crp, there is a 1.40% increase in geometric mean serum fibrinogen level. 95% CI suggests that this observation would not be unusual if the true percent increase in geometric mean fibrinogen level per 1mg/L increase in mean crp levels were between 1.22% and 1.58%. Because the two sided P value is < 0.001 we reject the null hypothesis that there is no linear trend in geometric mean fibrinogen level and crp levels.

|  |  |
| --- | --- |
| CRP level | Fitted geometric mean Fibrinogen (mg/dL) |
| 1mg/L | 300.896 |
| 2mg/L | 309.513 |
| 3mg/L | 313.877 |
| 4mg/L | 317.984 |
| 6mg/L | 327.013 |
| 8mg/L | 336.299 |
| 9mg/L | 341.040 |
| 12mg/L | 355.669 |

6.

Methods: The relationship between log transformed serum fibrinogen and log transformed serum crp levels was analyzed using linear regression with robust estimates of SE and 95% CI. Of note, to avoid the error of generating inaccurately missing data when calculating geometric means, participants whose serum CRP levels were recorded as 0 were replaced by 0.5mg/L. This was based on the assumption that participants did not actually have a CRP level of 0mg/L, but in reality a smaller value then was clinically assayable by laboratory techniques. Estimates were exponentiated to gain inference on the geometric mean fibrinogen found in the below table.

Results: From linear regression analysis, we estimate that for each 1mg/L increase in log-transformed mean serum crp, there is a 11.1% increase in geometric mean serum fibrinogen level. A 95% CI suggests that this observation would not be unusual if the true percent increase in geometric mean fibrinogen level per 1mg/L increase in log-transformed mean crp levels were between 10.5% and 11.8%. Because the two sided P value is < 0.001 we reject the null hypothesis that there is no linear trend in geometric mean fibrinogen level and log-transformed crp levels.

|  |  |
| --- | --- |
| CRP level | Fitted geometric mean Fibrinogen (mg/dL) |
| 1mg/L | 292.536 |
| 2mg/L | 314.706 |
| 3mg/L | 328.446 |
| 4mg/L | 338.556 |
| 6mg/L | 353.337 |
| 8mg/L | 364.214 |
| 9mg/L | 368.764 |
| 12mg/L | 380.116 |

7.

|  |  |
| --- | --- |
|  | **Fitted Values for Fibrinogen (mg/dL)** |
| **Comparisons across CRP level** | **Problem 3: (mean serum fibrinogen)** | **Problem 4: (mean serum fibrinogen)** | **Problem 5: (geometric mean fibrinogen)** | **Problem 6: (geometric mean fibrinogen)** |
| ***Differences*** |
| **2 mg/L – 1 mg/L** | 5.251 | 25.531 | 8.617 | 22.17 |
| **3 mg/L – 2 mg/L** | 5.251 | 14.935 | 4.36 | 13.74 |
| **4 mg/L – 1 mg/L** | 15.753 | 51.062 | 17.088 | 46.02 |
| **4 mg/L – 2 mg/L** | 10.502 | 25.531 | 8.471 | 23.85 |
| **6 mg/L – 3 mg/L** | 15.752 | 25.530 | 13.136 | 24.891 |
| **8 mg/L – 4 mg/L** | 21.003 | 25.531 | 18.315 | 25.658 |
| **9 mg/L – 6 mg/L** | 15.752 | 14.935 | 14.027 | 15.427 |
| **9 mg/L – 8 mg/L** | 15.753 | 14.935 | 14.027 | 15.427 |
| **12 mg/L – 6 mg/L** | 31.505 | 25.531 | 28.656 | 26.779 |
| ***Ratios*** |
| **2 mg/L / 1 mg/L** | 1.017 | 1.086 | 1.029 | 1.076 |
| **3 mg/L / 2 mg/L** | 1.017 | 1.047 | 1.014 | 1.044 |
| **4 mg/L / 1 mg/L** | 1.051 | 1.173 | 1.057 | 1.157 |
| **4 mg/L / 2 mg/L** | 1.033 | 1.080 | 1.027 | 1.076 |
| **6 mg/L / 3 mg/L** | 1.049 | 1.076 | 1.041 | 1.076 |
| **8 mg/L / 4 mg/L** | 1.065 | 1.074 | 1.058 | 1.044 |
| **9 mg/L / 6 mg/L** | 1.047 | 1.04 | 1.043 | 1.012 |
| **9 mg/L / 8 mg/L** | 1.015 | 1.012 | 1.014 | 1.076 |
| **12 mg/L / 6 mg/L** | 1.094 | 1.071 | 1.088 | 1.076 |

8. Results from problem 7:

a. In problems 3 and 5, the analysis looking at the linear relationship between fibrinogen and crp and the analysis looking at the linear relationship between geometric mean fibrinogen and crp, the analyses gave constant differences between fitted values that differed in an absolute increase in c. In particular, the differences were the same for a c increase of 1mg/L crp between 2-1 and 3-2. Similarly the differences were the same for a c increase of 3mg/dL crp between 4-1,6-3 and 9-6. This constant relationship seemed to break down towards the upper limits of crp when fibrinogen exponentially increased.

b. The analyses that gave constant ratios for the fitted values when comparing two groups that differed by an absolute increase of c, were also problems 3 and 5. For example, the fitted values of the linear regression of fibrinogen and crp gave a mean fibrinogen ratio of approximately 1.015 when crp levels increased by 1mg/L. Similarly, in problem 5, the geometric mean fibrinogen ratio was 1.014 for crp levels that increased by 1mg/L.

c. The analyses that gave constant differences in fitted values when comparing two groups that differed by a relative c-fold increase in crp levels were analyses 4 and 6 that looked at the relationships between the log-transformed crp value and fibrinogen (in problem 6, the geometric mean fibrinogen). For example, in problem 4 for crp values that increased 2-fold (i.e. 2-1, 4-2 6-3, 8-4, 12-6), the mean fibrinogen levels increased by about 25.53. Similarly, in problem 6, for the crp values that increased by 2-fold (i.e. 2-1, 6-3, 8-4, 12-6), the geometric mean fibrinogen levels increased by about 22/24.

d. The analysis that gave constant ratios in the fitted values when comparing two groups that differed by a relative *c*-fold increase in CRP levels were also analyses 4 and 6 which looked at the relationship between log-transformed crp and fibrinogen/ geometric mean fibrinogen. For example, in problem 4, when crp values increased 2 fold (2/1, 4/2, 6/3, 12/6), the ratio of mean fibrinogen levels was about 1.08. When crp values were 1.5 fold higher (3/2, 9/6), the ratio of mean fibrinogen levels was about 1.04. Similarly, in problem 6, when crp values increased 2 fold (2/1, 4/2, 6/3, 12/6), the ratio of geometric mean fibrinogen levels was also about 1.08. When crp values were 1.5 fold higher (3/2, 9/6), the ratio of geometric mean fibrinogen levels was about 1.04.

9. In choosing between the different analyses, it is important to look at the original data from our descriptive statistics. The relationship between crp and fibrinogen displays heteroscedacity with increased variance in fibrinogen levels with increasing crp levels. In addition, the relationship between fibrinogen and crp is appears nonlinear. Therefore we may gain more precision by log transforming our dependent variable. The statistical inference on log transformed fibrinogen levels may also be more interpretable as a percent change in geometric mean fibrinogen level – i.e. a 1mg/L increase in CRP leads to a 1.4% increase in fibrinogen levels.