Biostatistics 518 HW#2

1. Provide a suitable descriptive analysis for the association between CRP and fibrinogen both overall, and separately for groups having no prior history of diagnosed CVD or having prior diagnosed CVD.
	1. Methods: Scatter plot of Fibrinogen vs. CRP with lowess smooth lines for all patients, patients with history of cardiovascular disease (CVD) and those without history of CVD.



* 1. Results:
		1. Overall there is a pattern of increase in CRP with increasing fibrinogen. The pattern of increase becomes most apparent at levels of fibrinogen near 300 mg/dL and the change in CRP for each change in fibrinogen seems to stabilize at fibrinogen levels near 600mg/dL. The variability in CRP level seems to increase as fibrinogen level increases.
		2. The trend is similar among patients with and without history of CVD, except for at levels of fibrinogen between approximately 600-700mg/dL, where patients with no history of CVD have a very slightly higher CRP level at a given level of fibrinogen.
1. t-test analyses exploring an association between mean fibrinogen and prior history of CVD.
	1. Presuming that the standard deviation of fibrinogen is similar within each group defined by presence or absence of prior history of CVD.
		1. Methods: A t-test presuming equal variances was performed to look for a difference in mean fibrinogen between subjects with and without history of cardiovascular disease. 2-sided p-value and 95% CI was calculated for the difference in means.
		2. Results: Subjects with history of CVD had a mean serum fibrinogen level 14.9 mg/dL higher than those without history of CVD (334.5 vs. 319.6). With 95% confidence, we can say that this result would not be unusual if the true population difference in mean fibrinogen level is between 10.4 and 19.3 mg/dL higher in patients with history of CVD than in those without. With a 2-sided p-value of <0.0001, we can reject the null hypothesis that there is no difference in CRP levels in patients with and without history of CVD.
	2. Using linear regression.
		1. Methods: Simple linear regression was performed to look for association between mean serum fibrinogen levels and history of CVD. The analysis was done presuming equal variances in fibrinogen between the groups (robust standard errors were not calculated).
		2. Results and Comparison with t-test: Here, the intercept (391.6) provides the mean fibrinogen level in subjects without history of CVD. In other words, it gives the value of mean fibrinogen when X=0 (no history of cardiovascular disease). The slope of the line (14.9) is identical to the difference in means given by the t-test. Because the prevdis variable is binary, this shows that for every increase in X (from 0=no history of disease, to 1=history of disease), there will be a mean difference in Y(fibrinogen) of 14.9mg/dL. This is the same difference in means given by the t-test. The 2-sided p-value and 95% confidence interval for the slope are <0.0001 and (10.4-19.3), respectively. These are identical to the p-value and 95% CI for the t-test above.
	3. Allowing for differences in SD of fibrinogen across groups.
		1. Methods: A t-test not presuming equal variances was performed to look for a difference in mean fibrinogen between subjects with and without history of cardiovascular disease. 2-sided p-value and 95% CI was calculated for the difference in means.
		2. Results: Subjects with history of CVD had a mean serum fibrinogen level 14.9 mg/dL higher than those without history of CVD (334.5 vs. 319.6). With 95% confidence, we can say that this result would not be unusual if the true population difference in fibrinogen level is between 10.1 and 19.7 mg/dL higher in patients with history of CVD than those without. With a 2-sided p-value of <0.0001, we can reject the null hypothesis that there is no difference in CRP levels in patients with and without history of CVD.
	4. Using linear regression allowing for differences in SD of CRP among subjects with and without history of CVD.
		1. Methods: Simple linear regression was performed to look for association between serum fibrinogen levels and history of CVD. The analysis was done without presuming equal variances in fibrinogen between the groups (robust standard errors were calculated).
		2. Results and Comparison with t-test: Again, the intercept (391.6) provides the mean fibrinogen level in subjects without history of CVD. The slope of the line (14.9) is again identical to the difference in means given by the t-test. The 2-sided p-value and confidence interval for the slope are <0.0001 and (10.1-19.7), respectively. These are identical to the p-value and 95% CI for the t-test above. The 95% confidence interval for the test not assuming equal variances is slightly wider than for the test presuming equal variances.
	5. In the t-test that presumes equal variances, if the group with the higher variance has the smaller sample size, then the t-test will be anti-conservative (more likely to report a smaller p-value and a tighter confidence interval). In this case, the subjects with history of CVD have a smaller sample size and a larger variance, meaning that the t-test that presumes equal variance will give us a tighter confidence interval than the t-test that does not presume equal variance (which is exactly what we see). If we could see all the digits of the p-value, the p-value for the test that presumes equal variance will be smaller than the p-value for the t-test that does not.
2. Methods: Linear regression with robust standard error was performed for the association between CRP and mean fibrinogen. A 2-sided p-value and 95% confidence interval were calculated for the slope.
	1. The estimated intercept is 304 mg/dL. This represents the estimated mean fibrinogen level when the CRP level is 0.
	2. The estimated slope is 5.25. This indicates that for every 1mg/L change in serum CRP level, the expected change in mean fibrinogen level is 5.25 mg/dL.
	3. The linear regression model tells us that for every 1mg/L change in serum CRP level, the expected change in mean fibrinogen level is an increase of 5.25 mg/dL. With 95% confidence we can say that the true population increase in mean fibrinogen level for every 1 mg/L increase in CRP is between 4.60 and 5.90 mg/dL higher. With a 2-sided p-value of <0.0001, we can reject the null hypothesis that there is no association between serum CRP level and mean fibrinogen level.
	4. See Table 1 at end of problem 6.
3. Methods: Linear regression with robust standard error was performed for the association between log-transformed CRP and mean fibrinogen. A p-value and 95% confidence interval was calculated for the slope.
	1. The estimated intercept is 295.57 mg/dL. This represents the estimated mean fibrinogen level when logCRP level is 0, or when actual CRP level is 1 mg/L.
	2. The estimated slope is 36.83. This indicates that for every 1-unit change in log-transformed serum CRP level, the expected change in mean fibrinogen level is 36.83mg/dL.
	3. The linear regression model tells us that for every 1 unit change in log-transformed serum CRP level, the expected change in mean fibrinogen level is an increase of 36.83mg/dL. With 95% confidence we can say that the true population increase in mean fibrinogen level for every 1 unit increase in logCRP is between 34.58 and 39.09 mg/dL higher. With a p-value of <0.0001, we can reject the null hypothesis that there is no association between serum CRP level and mean fibrinogen level.
	4. See Table 1 at end of problem 6.
4. Methods: Linear regression with robust standard error was performed for the association between CRP and geometric mean fibrinogen. A 2-sided p-value and 95% confidence interval were calculated for the slope.
	1. The intercept is 300.77 mg/dL. This represents the geometric mean fibrinogen level when CRP level is 0.
	2. The estimated slope is 1.014. This indicates that for every 1 mg/L change in serum CRP level, the expected percent-change geometric mean fibrinogen level is 1.4%.
	3. The linear regression model tells us that for every 1 unit change in serum CRP level, the expected percent-change in geometric mean fibrinogen level is an increase of 1.4%. With 95% confidence we can say that the true population increase in percent geometric mean fibrinogen level for every 1 mg/L increase in CRP is between 1.2 % and 1.6% higher. With a p-value of <0.0001, we can reject the null hypothesis that there is no association between serum CRP level and geometric mean fibrinogen level.
	4. See Table 1 at end of problem 6.
5. Methods: Linear regression with robust standard error was performed for the association between log-transformed CRP and geometric mean fibrinogen. A p-value and 95% confidence interval were calculated for the slope.
	1. The intercept is 292.54 mg/dL. This represents the estimated geometric mean fibrinogen level when logCRP level is 0, or when actual CRP level is 1.
	2. The estimated slope is 1.11. This indicates that for every 1-unit change in log-transformed serum CRP level, the expected percent change in geometric mean fibrinogen level is 11%.
	3. The linear regression model tells us that for every 1 unit change in log-transformed serum CRP level, the expected percent change in geometric mean fibrinogen level is an increase of 11%. With 95% confidence we can say that the true population increase in geometric mean fibrinogen level for every 1 unit increase in logCRP is between 10% and 12% higher. With a p-value of <0.0001, we can reject the null hypothesis that there is no association between log-transformed serum CRP level and geometric mean fibrinogen level.
	4. See Table 1 below.

**Table 1**: Summary of fitted values for mean fibrinogen by CRP level by 4 different analytical methods.

|  |  |
| --- | --- |
|  | **Fitted Values for Fibrinogen (mg/dL)** |
| **CRP level** | **Problem 3** **Mean fibrinogen**  | **Problem 4****Mean fibrinogen** **(CRP log-transformed)** | **Problem 5****Geometric mean fibrinogen** | **Problem 6****Geometric mean fibrinogen** **(CRP log-transformed)** |
| **1 mg/L** | 309.27 | 295.57 | 304.97 | 293.64 |
| **2 mg/L** | 314.52 | 321.10 | 309.23 | 314.71 |
| **3 mg/L** | 319.77 | 336.04 | 313.55 | 328.45 |
| **4 mg/L** | 325.02 | 346.63 | 317.93 | 338.56 |
| **6 mg/L** | 335.52 | 361.57 | 326.89 | 353.34 |
| **8 mg/L** | 346.02 | 372.16 | 336.08 | 364.21 |
| **9 mg/L** | 351.27 | 376.50 | 340.77 | 368.76 |
| **12 mg/L** | 367.03 | 387.10 | 355.25 | 380.12 |

**Table 2**: Comparison of fitted values for mean fibrinogen by CRP level by 4 different analytical methods.

|  |  |
| --- | --- |
|  | **Fitted Values for Fibrinogen (mg/dL)** |
| **CRP level** | **Problem 3** **Mean fibrinogen**  | **Problem 4****Mean fibrinogen** **(CRP log-transformed)** | **Problem 5****Geometric mean fibrinogen** | **Problem 6****Geometric mean fibrinogen** **(CRP log-transformed)** |
| ***Differences*** |
| **2 mg/L - 1 mg/L** | 5.25 | 25.53 | 4.26 | 21.07 |
| **3 mg/L - 2 mg/L** | 5.25 | 14.91 | 4.32 | 13.74 |
| **4 mg/L - 1 mg/L** | 5.25 | 51.06 | 12.96 | 44.92 |
| **4 mg/L - 2 mg/L** | 10.50 | 25.53 | 8.70 | 23.85 |
| **6 mg/L - 3 mg/L** | 15.75 | 25.53 | 13.34 | 24.89 |
| **8 mg/L - 4 mg/L** | 21.00 | 25.53 | 18.15 | 25.65 |
| **9 mg/L - 6 mg/L** | 10.50 | 14.91 | 13.88 | 15.42 |
| **9 mg/L - 8 mg/L** | 5.25 | 4.34 | 4.69 | 4.55 |
| **12 mg/L - 6 mg/L** | 31.5 | 25.53 | 28.36 | 26.78 |
| ***Ratios*** |
| **2 mg/L / 1 mg/L** | 1.0169 | 1.0864 | 1.0140 | 1.0718 |
| **3 mg/L / 2 mg/L** | 1.0167 | 1.0465 | 1.0140 | 1.0437 |
| **4 mg/L / 1 mg/L** | 1.0509 | 1.1728 | 1.0425 | 1.1530 |
| **4 mg/L / 2 mg/L** | 1.0334 | 1.0795 | 1.0281 | 1.0758 |
| **6 mg/L / 3 mg/L** | 1.0493 | 1.0760 | 1.0425 | 1.0758 |
| **8 mg/L / 4 mg/L** | 1.0646 | 1.0737 | 1.0570 | 1.0758 |
| **9 mg/L / 6 mg/L** | 1.0469 | 1.0412 | 1.0425 | 1.0436 |
| **9 mg/L / 8 mg/L** | 1.0152 | 1.0117 | 1.0140 | 1.0125 |
| **12 mg/L / 6 mg/L** | 1.0939 | 1.0706 | 1.0868 | 1.0758 |

1. With respect to Table 2 above:
	1. The simple linear regression for association between CRP and mean fibrinogen with untransformed variables gave constant differences in the fitted values when comparing two groups that differed by an absolute increase of *c* units of CRP. For every increase of 1 in CRP, there was an increase of 5.25 in mean fibrinogen, for every increase in 2 in CRP, there was an increase of 10.5 in mean fibrinogen, etc.
	2. The linear regression for association between CRP and geometric mean fibrinogen gave constant ratios of fitted values when comparing groups that differed by an absolute difference of *c* in mean fibrinogen. For example, when comparing the value of ratios between a CRP of 3/2 or for 8/9 (values that differ by 1), we get a value of 1.0140 (or a 1.4% difference). When we compare values that differ by 3 (e.g 6/3 and 9/6) we get a ratio of 1.0425 (or 4.25% difference).
	3. The linear regression for association between logCRP and mean fibrinogen gave constant differences in the fitted values between two groups that differed by a relative *c*-fold increase in CRP levels. For example, the difference in fibrinogen levels for CRP levels that are 2-fold higher (4 and 2, 6 and 3, 8 and 4, 12 and 6), are all 25.53. The difference in fibrinogen levels for CRP levels that are 1.5 times higher (3 and 2, 9 and 6) are identical at 14.91.
	4. The linear regression for association between logCRP and geometric mean CRP gave constant ratios between groups that differed by a *c*-fold increase in CRP. For example, in groups that had CRP levels that were 2-fold higher, the ratio was 1.0758 (7.58% higher) and in groups that had CRP levels that were 1.5-fold higher, the ratio was 1.0436 or 4.36% higher.
2. Because both fibrinogen and CRP are markers of inflammation that likely signify on a multiplicative, rather than an arithmetic scale, I want to use an analysis that provides information on the percent increase in fibrinogen for a given percentage increase in CRP. Therefore, I prefer to use the analysis from problem 6, log-transforming both the predictor and the fibrinogen levels.