**Biost 518: Applied Biostatistics II**

**Biost 515: Biostatistics II**

Emerson, Winter 2014

**Homework #7**

February 24, 2014

1. The association between gender and cholesterol was assessed for Caucasians and Noncaucasians. The difference in mean cholesterol between genders was calculated for each race group. Corresponding standard error estimates and 95% confidence intervals for the estimated difference in mean cholesterol were also calculated. Finally, Z-statistics and p-values were calculated to test if there is an association between sex and cholesterol within each race group. The formulas used for calculation are shown below.

The association between cholesterol level and sex was also assessed after adjusting for race. Adjusted estimates for difference in mean cholesterol between genders were calculated using efficiency and importance weights. Standard error estimates and 95% confidence intervals were also calculated for the weighted estimates. Finally, Z-statistics and p-values were calculated to test if there is an association between sex and cholesterol after controlling for race. The formulas used for calculating the efficiency and importance weights, weighted estimate, and standard error are below. The formulas used to calculate the confidence interval and Z statistic are the same as above. The notation is as follows: w is the calculated weight, Δ is the difference in mean cholesterol across genders, C represents Caucasians and NC represents Noncaucasians, and p is the proportion of individuals in the sample that are a particular race.

Race was also assessed as an effect modifier by testing if the difference in mean cholesterol level between race groups was equal between genders. The above formulas were used.

* 1. The difference in mean cholesterol levels between male and female Caucasians is 25.3 mg/dl with males having lower average cholesterol. The standard error for this estimate is 1.55 mg/dl. It would not be unusual to observe mean cholesterol 22.3 mg/dL to 28.3 mg/dL higher in Caucasian females compared to Caucasian males. The p-value associated with the test of equality in mean cholesterol between Caucasian men and women is less than 0.001, so we find that there is an association between sex and cholesterol among Caucasians.
	2. The difference in mean cholesterol levels between male and female Noncaucasians is 15.7 mg/dl with males having lower average cholesterol. The standard error for this estimate is 3.45 mg/dl. It would not be unusual to observe mean cholesterol between 8.94 mg/dL to 22.5 mg/dL higher in Noncaucasian females compared to Noncaucasian males. The p-value associated with the test of equality in mean cholesterol between Caucasian men and women is less than 0.001, so we find that there is an association between sex and cholesterol among Noncaucasians.
	3. **Importance weights**: The weight for Caucasians is 0.8405 and the weight for Noncaucasians is 0.1595. The estimate for difference in mean cholesterol between men and women after adjusting for race is 23.77 mg/dl, with men having lower cholesterol level. The estimated standard error is 1.414 mg/dl. It would not be unusual to observe a mean cholesterol 21.0 mg/dl to 26.5 mg/dl higher in women compared to men, after adjusting for race. The p-value associated with the test of equality in mean cholesterol between men and women after adjusting for race is less than 0.001, so we reject the null hypothesis that there is no association between sex and Cholesterol after adjusting for race.

	**Efficiency weights:** The weight for Caucasians is 0.8320 and the weight for Noncaucasians is 0.1680. The estimate for difference in mean cholesterol between men and women after adjusting for race is 23.69 mg/dl, with men having lower cholesterol level. The estimated standard error is 1.414 mg/dl. It would not be unusual to observe a mean cholesterol 20.92 mg/dl to 26.64 mg/dl higher in women compared to men, after adjusting for race. The p-value associated with the test of equality in mean cholesterol between men and women after adjusting for race is less than 0.001, so we reject the null hypothesis that there is no association between sex and Cholesterol after adjusting for race.
	4. The difference in mean cholesterol across race groups was found to be 9.60 mg/dl lower in males than females. The corresponding standard error is 3.78 mg/dl. It would not be unusual to observe a difference in mean cholesterol across race groups 17.01 mg/dl to 2.19 mg/dl lower in males compared to females. The p-value associated with the testing the equality of difference, between genders, in mean cholesterol across race groups is 0.0111. Therefore, we can reject the null hypothesis and find that race does modify the association between mean cholesterol and sex.
1. The association between gender and fibrinogen was assessed, adjusting for race. Race was also assessed as an effect modifier. The methods used are the same as in question 1.

	1. The difference in mean fibrinogen levels between male and female Caucasians is 2.90 mg/dl with males having lower average fibrinogen. The standard error for this estimate is 2.68 mg/dl. It would not be unusual to observe mean fibrinogen 8.15 mg/dL to 2.35 mg/dL lower in Caucasian males compared to Caucasian females. The p-value associated with the test of equality in mean fibrinogen between Caucasian men and women is 0.279, so we fail to reject the null hypothesis. That is, we do not have enough evidence to say there is a relationship between fibrinogen levels and gender for Caucasians.
	2. The difference in mean fibrinogen levels between male and female Noncaucasians is 15.7 mg/dl with males having lower average fibrinogen. The standard error for this estimate is 7.30 mg/dl. It would not be unusual to observe mean fibrinogen 30.0 mg/dL to 1.40 mg/dL lower in Noncaucasian males compared to Noncaucasian females. The p-value associated with the test of equality in mean fibrinogen between Noncaucasian men and women is 0.0315. We reject the null hypothesis and can say with high confidence that there is an association between sex and fibrinogen levels for Noncaucasians, with men tending to have lower fibrinogen levels.
	3. **Importance weights**: The weight for Caucasians is 0.8405 and the weight for Noncaucasians is 0.1595. The estimate for difference in mean fibrinogen between men and women after adjusting for race is 4.94 mg/dl, with men having lower fibrinogen level. The estimated standard error is 2.536 mg/dl. It would not be unusual to observe a mean fibrinogen level 9.911 mg/dl lower to 0.0306 mg/dl higher in men compared to women, after adjusting for race. The p-value associated with the test of equality in mean fibrinogen between men and women after adjusting for race is 0.0514, so we fail to reject the null hypothesis. That is, we do not have enough evidence to say there is an association between sex and fibrinogen after adjusting for race.

**Efficiency weights**: The weight for Caucasians is 0.8812 and the weight for Noncaucasians is 0.1188. The estimate for difference in mean fibrinogen between men and women after adjusting for race is 4.421 mg/dl, with men having lower fibrinogen level. The estimated standard error is 2.516 mg/dl. It would not be unusual to observe a mean fibrinogen level 9.352 mg/dl lower to 0.5103 mg/dl higher in men compared to women, after adjusting for race. The p-value associated with the test of equality in mean fibrinogen between men and women after adjusting for race is 0.0789, so we fail to reject the null hypothesis. That is, we do not have enough evidence to say there is an association between sex and fibrinogen after adjusting for race.

* 1. The difference in mean fibrinogen across race groups was found to be 12.8 mg/dl higher in males than females. The corresponding standard error is 7.78 mg/dl. It would not be unusual to observe a difference in mean cholesterol across race groups 2.442 mg/dl lower to 28.04 mg/dl higher in males compared to females. The p-value associated with the testing the equality of difference, between genders, in mean cholesterol across race groups is 0.100, so we fail to reject the null hypothesis. That is, we do not have enough evidence to say race modifies the association between mean fibrinogen levels and sex.
	2. The estimated standard deviation of cholesterol for the sample is 39.29 mg/dl. However, 47 subjects did not have cholesterol measurements.
	3. The estimated standard deviation of the change in cholesterol measurements made after three years in the population is 55.27 mg/dl.
	4. The estimated standard deviation for groups that are homogenous for age and sex is 37.49 mg/dl.
	5.
	6.
	7.
	8. If we only considered the final measurements, our sample size would become smaller. This would happen because the correlation between cholesterol measurements for an individual is small (les than 0.5) so removing the baseline measurements is actually more statistically precise.
	9. If we used the ANCOVA model that adjusted for age, sex, and the baseline cholesterol level the sample size would decrease and be smaller than those found in 5(a-d). This is because ANCOVA has uniformly better efficiency.
	10. The proportion of subjects on the control arm with serum cholesterol below 200 mg/dL at the end of treatment is 0.3957.
	11. The proportion of subjects on the treatment arm with serum cholesterol below 200 mg/dL at the end of treatment is 0.4943.
	12.
	13. A disadvantage to this study compared to the one found in 4(b) is this study design requires a much larger sample size. Because we are more interested in detecting the difference between two groups, dichotomizing the data better answers the scientific questions. However, another disadvantage is that by dichotomizing the data we lose information and have less statistical precision.