Homework 7

1. Analysis of mean of serum cholesterol and according to patient sex and race
2. Mean cholesterol levels were found to be 25.3 mg/dL lower in male Caucasians than in female Caucasians (standard error of 1.55 mg/dL). We found a significant difference to be able to reject the null hypothesis of no difference in mean cholesterol across groups defined by gender in Caucasians (P<.0001). Based on a 95% confidence interval, we find that the observed difference in mean cholesterol is not atypical of settings in which the true difference in mean cholesterol were such that the male Caucasians had cholesterol levels of 28.34 mg/dL to 22.26 mg/dL lower than their female counterparts.
3. Mean cholesterol levels were found to be 15.7 mg/dL lower in male Noncaucasians than in female Noncaucasians (standard error of 3.45 mg/dL). We found a significant difference to be able to reject the null hypothesis of no difference in mean cholesterol across groups defined by gender in Noncaucasians (P<.0001). Based on a 95% confidence interval, we find that the observed difference in mean cholesterol is not atypical of settings in which the true difference in mean cholesterol were such that the male Noncaucasians had cholesterol levels of 22.47 mg/dL to 8.93 mg/dL lower than their female counterparts.
4. **Using importance weights**: After adjustment for race, mean cholesterol levels were found to be 23.77 mg/dL lower in males than in females (standard error of 1.416 mg/dL). We found a significant difference to be able to reject the null hypothesis of no difference in mean cholesterol across groups defined by gender (P<.0001). Based on a 95% confidence interval, we find that the observed difference in mean cholesterol is not atypical of settings in which the true difference in mean cholesterol were such that the males had cholesterol levels of 26.55 mg/dL to 20.99 mg/dL lower than the females.

**Using efficiency weights**: After adjustment for race, mean cholesterol levels were found to be 23.69 mg/dL lower in males than in females (standard error of 1.416 mg/dL). We found a significant difference to be able to reject the null hypothesis of no difference in mean cholesterol across groups defined by gender (P<.0001). Based on a 95% confidence interval, we find that the observed difference in mean cholesterol is not atypical of settings in which the true difference in mean cholesterol were such that the males had cholesterol levels of 26.47 mg/dL to 20.91 mg/dL lower than the females.

1. The difference in mean cholesterol across groups defined by gender was found to be 9.6 mg/dL lower in Caucasians than in Noncaucasians (standard error of 3.786 mg/dL). Such a difference was found sufficiently extreme to be able to rule out a null hypothesis of no effect modification by race in the association between cholesterol level and sex (P=0.011). Based on a 95% confidence interval, we find that the observed difference in the association between cholesterol and sex across the race groups is not atypical of settings in which the true difference in effect were such that Caucasians had mean difference in cholesterol across gender groups 17.02 mg/dL to 2.18 mg/dL lower than that in Noncaucasians.
2. Analysis of mean of serum fibrinogen and according to patient sex and race
3. Mean fibrinogen levels were found to be 2.9 mg/dL lower in male Caucasians than in female Caucasians (standard error of 2.677 mg/dL). We did not find a significant difference to be able to reject the null hypothesis of no difference in mean fibrinogen across groups defined by gender in Caucasians (P=0.2787). Based on a 95% confidence interval, we find that the observed difference in mean fibrinogen is not atypical of settings in which the true difference in mean fibrinogen were such that the male Caucasians had fibrinogen levels of 8.147 mg/dL lower to 2.347 mg/dL higher than their female counterparts.
4. Mean fibrinogen levels were found to be 15.7 mg/dL lower in male Noncaucasians than in female Noncaucasians (standard error of 7.296 mg/dL). We found a significant difference to be able to reject the null hypothesis of no difference in mean fibrinogen across groups defined by gender in Noncaucasians (P=0.0314). Based on a 95% confidence interval, we find that the observed difference in mean fibrinogen is not atypical of settings in which the true difference in mean fibrinogen were such that the male Noncaucasians had fibrinogen levels of 30.00 mg/dL to 1.40 mg/dL lower than their female counterparts.
5. **Using importance weights**: After adjustment for race, mean fibrinogen levels were found to be 4.94 mg/dL lower in males than in females (standard error of 2.53 mg/dL). We did not find a significant difference (by a thin margin) to be able to reject the null hypothesis of no difference in mean fibrinogen across groups defined by gender (P=0.051). Based on a 95% confidence interval, we find that the observed difference in mean fibrinogen is not atypical of settings in which the true difference in mean fibrinogen were such that the males had fibrinogen levels of 9.91 mg/dL to 0.02 mg/dL lower than the females.

**Using efficiency weights**: After adjustment for race, mean fibrinogen levels were found to be 4.42 mg/dL lower in males than in females (standard error of 2.51 mg/dL). We did not find a significant difference to be able to reject the null hypothesis of no difference in mean fibrinogen across groups defined by gender (P=0.079). Based on a 95% confidence interval, we find that the observed difference in mean fibrinogen is not atypical of settings in which the true difference in mean fibrinogen were such that the males had fibrinogen levels of 9.34 mg/dL lower to 0.51 mg/dL higher than the females.

1. The difference in mean fibrinogen across groups defined by gender was found to be 12.8 mg/dL lower in Caucasians than in Noncaucasians (standard error of 7.77 mg/dL). We did not find a significant difference to be able to reject the null hypothesis of no effect modification by race in the association between fibrinogen and sex (P=0.0996). Based on a 95% confidence interval, we find that the observed difference in the association between fibrinogen and sex across the race groups is not atypical of settings in which the true difference in effect were such that Caucasians had mean difference in fibrinogen across gender groups 28.03 mg/dL lower to 2.43 mg/dL higher than that in Noncaucasians.
2. Obtaining estimates for use in sample size calculations when using mean cholesterol
3. The best estimate of standard deviation of cholesterol within the sample is 39.29 mg/dL.
4. The standard deviation of the change in cholesterol measurements made after three years within the population is 43.04 mg/dL. This is obtained by assuming equal variances between the two timepoints.
5. The best estimate of the standard deviation of cholesterol within groups that have constant age and sex is 37.49 mg/dL. This is obtained from the root mean squared error output of linear regression of cholesterol on age and sex.
6. Two arm study of change in cholesterol after 2 years of treatment with adjustment for age and sex
7. A sample size of 530 will provide 80% power to detect the design alternative.
8. A sample size of 709 will provide 90% power to detect the design alternative.
9. If we do not adjust for age and sex, a sample size of 779 will provide 90% power to detect the design alternative.
10. If we analyze only the final cholesterol measurement adjusted for age and sex, the sample size needed will decrease. This is because they are not highly correlated. From lecture 8 slides 50 and 52, for values of correlation that are lower than 0.5 it is better to throw away the baseline.
11. If we use an Analysis of Covariance model that adjusted for age, sex, and the baseline cholesterol level, the sample size needed will decrease. This is because in a randomized clinical trial, the Analysis of Covariance produces a uniformly better efficiency. It is generally advantageous to use Analysis of Covariance since it is the most efficient estimator.
12. Two arm study of cholesterol after 2 years of treatment and the effect of dichotomizing the data
13. The estimate of the proportion pc of subjects on the control arm with serum cholesterol below 200 mg/dL at the end of treatment is 0.3957.
14. The estimate of the proportion pt of subjects on the treatment arm with serum cholesterol below 200 mg/dL at the end of the treatment (or number having cholesterol levels below 210 mg/dL in the CHS data) is 0.4943.
15. A sample size of 1058 will provide 90% power to detect the design alternative.
16. This study design is disadvantageous since we require a larger sample size to produce equivalent amount of power. To the extent that it is important to measure the treatment effect of the proportion of subjects having cholesterol below 200 mg/dL at the end of treatment, this study design may answer the most relevant scientific question. However, if such a scientific threshold did not exist, the new have lost information from dichotomizing which is reflected by the larger sample size required.