**Questions 1 and 2** suppose that you are reading a scientific article in a journal with inadequate statistical review. The scientific question addressed by the article is the association between blood lipid profiles (especially total cholesterol), biomarkers of inflammation (fibrinogen), and mortality from cardiovascular disease. The authors were also interested in the role of race (as categorized by Caucasian and Noncaucasian) in the relationship between sex and the serum measurements of total cholesterol and fibrinogen.

The authors reported gathering data on 3,015 subjects, of whom 1,258 were male and 1,757 were female. The subjects were further characterized as 2,534 Caucasians, 481 Noncaucasians. The data analysis presented in the manuscript is limited to the means and standard errors of the serum measures within subgroups as given in the following table.

**Table 1. Means (standard errors) of serum cholesterol and fibrinogen according to patient sex and race.**

|  |  |  |
| --- | --- | --- |
|  | **Males** | **Females** |
| **Caucasians** | **Noncaucasians** | **Caucasians** | **Noncaucasians** |
| **Cholesterol (mg/dl)** | 197.5 (1.092) | 197.9 (2.557) | 222.8 (1.103) | 213.6 (2.321) |
| **Fibrinogen (mg/dl)** | 317.8 (2.126) | 333.7 (5.628) | 320.7 (1.627) | 349.4 (4.643) |

1. You desire to do a more careful evaluation of the evidence at hand for associations between sex and cholesterol. You therefore desire to compute estimates, 95% confidence intervals, and P values to address questions of associations within subgroups, associations adjusted for race, and effect modification. In addressing the following questions, provide a sentence that interprets your inferential statistics in a manner suitable for inclusion in a scientific journal article. Avoid statistical jargon. (You note that without the sample sizes by subgroup, you will not be able to use the exact statistical methods (i.e., t tests) that you might otherwise have, but you will be able to perform analyses based on large sample approximations and the fact that sample means are approximately normally distributed. The Stata function normal() will return the cumulative distribution function for the standard normal. Hence,

di normal(1.96)

will display 0.9750021.

* 1. Are mean cholesterol levels associated with sex in Caucasians? (Recall that the standard error of two independent statistics is the square root of the sum of the squares of the individual standard errors. Thus calculate the standard error for the difference in mean cholesterol using the standard errors for the males and females.)

**Mean cholesterol levels were found to be 25.3mg/dL lower (SE: 1.552mg/dL) in Caucasian males than females. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean cholesterol was between 22.59mg/dL and 28.342mg/dL lower in Caucasian males than Caucasian females. This difference is statistically significant at a 0.05 level of significance (p < 0.0001) indicating that we can, with confidence, reject the null hypothesis of no difference in mean cholesterol by sex in Caucasians.**

* 1. Are mean cholesterol levels associated with sex in Noncaucasians?

**Mean cholesterol levels were found to be 15.7mg/dL lower (SE: 3.453mg/dL) in non-Caucasian males than females. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean cholesterol was between 8.932mg/dL and 22.469mg/dL lower in non-Caucasian males than non-Caucasian females. This difference is statistically significant at a 0.05 level of significance (p < 0.0001) indicating that we can, with confidence, reject the null hypothesis of no difference in mean cholesterol by sex in non-Caucasians.**

* 1. Are mean cholesterol levels associated with sex after adjustment for race? Provide adjusted estimates using both importance and efficiency weights.

*An approach that can be used here is to find a weighted average of the measures of effect in each race group. Hence, you might use a weighted average of the estimates ΔC and ΔN you derived in parts a and b, respectively: Let the adjusted estimated be defined according to*

*Δadj = (wC × ΔC + wN × ΔN) / (wC + wN)*

*where wC and wN are relative weights to be applied to the two strata. (Note that the equation becomes simpler if we ensure that the relative weights sum to 1.) The SE of the adjusted estimate of effect is then found by using the properties of variances. Recall that when multiplying a random variable by a constant, Var(cX) = c2 Var(X). Hence, you can find the standard error of the adjusted estimate can be found by*

**

*Many options could be considered for choosing the weights. Two that might be considered include:*

* + - *Importance weights: We weight each stratum according to its relative importance in the population of interest. This could be estimated from our sample (84.05% of our sample was Caucasian, so we could assume that that was also the frequency in the general population of elderly adults) or taken from, say, US census data (86.37% of US residents aged 65 years or older are Caucasian).*
		- *Efficiency weights: Under the assumption of no effect modification, the most efficient analysis would be to weight each stratum in proportion to the inverse of the square of the standard error of the stratum specific estimate.*

**Importance Weights:**

**After adjusting for race mean cholesterol levels were found to be 23.769mg/dL lower (SE: 1.416mg/dL) in males than in females of the same race. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean cholesterol was between 20.993mg/dL and 26.544mg/dL lower in males than in females of the same race. This difference is statistically significant at a 0.05 level of significance (p < 0.0001) indicating that we can, with confidence, reject the null hypothesis of no difference in mean cholesterol by sex after adjusting for race.**

**Efficiency Weights:**

**After adjusting for race mean cholesterol levels were found to be 23.687mg/dL lower (SE: 1.416mg/dL) in males than in females of the same race. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean cholesterol was between 20.912mg/dL and 26.461mg/dL lower in males than in females of the same race. This difference is statistically significant at a 0.05 level of significance (p < 0.0001) indicating that we can, with confidence, reject the null hypothesis of no difference in mean cholesterol by sex after adjusting for race.**

* 1. Does race modify the association between mean cholesterol level and sex?

**The difference in mean cholesterol by sex was found to be 9.6mg/dL lower (SE: 3.786mg/dL) in Caucasians than non-Caucasians. Based on a 95% confidence interval we find that the observed difference in the association of mean cholesterol and sex across race groups would not be unusual if the true difference in effect were such that Caucasians had mean difference in cholesterol by sex between 2.179mg/dL and 17.021mg/dL lower than that in non-Caucasians. This difference is statistically significant at a 0.05 level of significance (p = 0.0112) indicating that we can, with confidence, reject the null hypothesis of no effect modification by race in the association between cholesterol level and sex.**

1. You also desire to do a more careful evaluation of the evidence at hand for fibrinogen. You therefore answer the questions of problem 1 using the statistics for fibrinogen.
	1. Are mean fibrinogen levels associated with sex in Caucasians

**Mean fibrinogen levels were found to be 2.9mg/dL lower (SE: 2.677mg/dL) in Caucasian males than females. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean fibrinogen was between 8.147mg/dL lower and 2.347mg/dL higher in Caucasian males than Caucasian females. This difference is not statistically significant at a 0.05 level of significance (p = 0.2787) indicating that we cannot, with confidence, reject the null hypothesis of no difference in mean fibrinogen by sex in Caucasians.**

* 1. Are mean fibrinogen levels associated with sex in Noncaucasians?

**Mean fibrinogen levels were found to be 15.7mg/dL lower (SE: 7.296mg/dL) in non-Caucasian males than females. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean fibrinogen was between 1.398mg/dL and 30.000mg/dL lower in non-Caucasian males than non-Caucasian females. This difference is statistically significant at a 0.05 level of significance (p = 0.0314) indicating that we can, with confidence, reject the null hypothesis of no difference in mean fibrinogen by sex in non-Caucasians.**

* 1. Are mean fibrinogen levels associated with sex after adjustment for race?

**Importance Weights:**

**After adjusting for race mean fibrinogen levels were found to be 4.942mg/dL lower (SE: 2.533mg/dL) in males than in females of the same race. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean fibrinogen was between 9.907mg/dL lower and 0.023mg/dL higher in males than in females of the same race. This difference is not statistically significant at a 0.05 level of significance (p = 0.0511) indicating that we cannot, with confidence, reject the null hypothesis of no difference in mean fibrinogen by sex after adjusting for race.**

**Efficiency Weights:**

**After adjusting for race mean fibrinogen levels were found to be 4.419mg/dL lower (SE: 2.513mg/dL) in males than in females of the same race. Based on a 95% confidence interval we find that the observed difference would not be unusual if the true difference in mean fibrinogen was between 9.345mg/dL lower and 0.507mg/dL higher in males than in females of the same race. This difference is not statistically significant at a 0.05 level of significance (p = 0.0787) indicating that we cannot, with confidence, reject the null hypothesis of no difference in mean fibrinogen by sex after adjusting for race.**

* 1. Does race modify the association between mean fibrinogen level and sex?

**The difference in mean fibrinogen by sex was found to be 12.8mg/dL higher (SE: 7.772mg/dL) in Caucasians than non-Caucasians. Based on a 95% confidence interval we find that the observed difference in the association of mean cholesterol and sex across race groups would not be unusual if the true difference in effect were such that Caucasians had mean difference in cholesterol by sex between 2.433mg/dL lower and 28.033mg/dL higher than that in non-Caucasians. This difference is not statistically significant at a 0.05 level of significance (p = 1.9004) indicating that we cannot, with confidence, reject the null hypothesis of no effect modification by race in the association between fibrinogen level and sex.**

**Questions 3 – 5** relate to the planning of a phase III clinical trial of a dietary intervention intended to improve cardiovascular health in a population of elderly adults by lowering serum cholesterol. Because we anticipate using an elderly patient population similar to that used in the cardiovascular health study, we will use the data in inflamm.txt (on the class web pages) to obtain estimates of the variances and correlations necessary to obtain power and sample size.

1. **(Obtaining estimates for use in sample size calculations when using mean cholesterol)** When making inference about cholesterol using means (and differences of means), the formula for *V* will typically involve the standard deviation *σ* of measurements made within a treatment group. The following estimates should be used as needed to answer all other questions. Using the inflamm.txt dataset available on the class web pages.
	1. Ideally, we want the standard deviation of cholesterol at baseline and the standard deviation of cholesterol measured after two years of treatment. However, as we only have ready access to a single cross-sectional measurement, we will have to use that data to estimate both SDs. What is your best estimate of the standard deviation of cholesterol within the sample? Report using four significant digits.

**Using the inflammatory biomarkers dataset, the best estimate for the standard deviation is the root MSE from the linear regression model of cholesterol by age using robust standard errors.**

**SD = 39.17mg/dL**

* 1. Assuming that the correlation ρ of cholesterol measurements made two years apart on the same individual is ρ = 0.40, what is the standard deviation of the change in cholesterol measurements made after three years within the population? Report using four significant digits.

**Using the estimate for standard deviation above the best estimate for the standard deviation of paired measurements is given by:**

$$SD= \sqrt{2σ^{2}\left(1-ρ\right)}= \sqrt{2×39.17^{2}\left(1-0.40\right)}$$

**SD = 42.91mg/dL**

* 1. We could also consider an analysis that would adjust for age and sex. In such a setting, we would want an estimate of the SD within groups that are homogenous for age and sex. What is your best estimate of the standard deviation of cholesterol within groups that had constant age and sex? Report using four significant digits. (Hint: Recall that the output from a regression model will provide an estimate of a common SD within groups as the “root mean squared error”. So you will need to perform a regression that allows each age-sex combination to have its own mean. A linear regression modeling age continuously along with sex would be one approach.)

**Using the inflammatory biomarkers dataset, the best estimate for the standard deviation in groups that are homogenous in age and sex is given by the Root MSE from the linear regression model of cholesterol by age and sex using robust standard errors.**

**SD = 37.49mg/dL**

1. **(A two arm study of change in cholesterol after 2 years of treatment with adjustment for age and sex)** Suppose we randomly assign *N* subjects to receive either the new treatment or a control strategy. We use a randomization ratio of 1 subject on the new treatment to 1 subject on control. We use as our measure of treatment effect the mean change in cholesterol at the end of treatment for patients on the new treatment and mean change in cholesterol at the end of treatment for patients on control. The null hypothesis is that the difference in means is 0 mg/dL, and we want to detect whether the new treatment will result in an average change in cholesterol that is 10 mg/dL lower than might be expected on control.. We intend to perform a hypothesis test in which
* we adjust for age and sex,
* the one-sided level of significance is α = 0.025,
* the desired statistical power is β = 0.80 or 0.90,
* the measure of treatment effect is *θ = (μ T,2 - μ T,0 ) – (μ C,2 - μ C,0 )* (the mean change in cholesterol in the patients receiving the new treatment for 2 years of treatment minus the mean change in cholesterol in the patients treated with control for two years), and
* the average variability contributed by each subject to the estimated treatment effect (the difference in sample means) is *V= 8σ 2(1-ρ).* (Again, use a correlation of 0.4.)
* the comparison between alternative and null hypotheses is *Δ = θ1 - θ0*.
1. What sample size will provide 80% power to detect the design alternative?

$$N= \frac{δ\_{αβ}^{2}V}{∆^{2}}= \frac{\left(Z\_{1-α}+ Z\_{β}\right)^{2}\left(8σ^{2}\left(1-ρ\right)\right)}{(θ\_{1}- θ\_{0})^{2}}$$

$$N= \frac{\left(1.960+0.842\right)^{2}\left(8×37.49^{2}\left(1-0.40\right)\right)}{(-10-0)^{2}}=529.57$$

**N = 530 subjects**

1. What sample size will provide 90% power to detect the design alternative?

$$N= \frac{δ\_{αβ}^{2}V}{∆^{2}}$$

$$N= \frac{\left(Z\_{1-α}+ Z\_{β}\right)^{2}\left(8σ^{2}\left(1-ρ\right)\right)}{(θ\_{1}- θ\_{0})^{2}}$$

$$N= \frac{\left(1.960+1.282\right)^{2}\left(8×37.49^{2}\left(1-0.40\right)\right)}{(-10-0)^{2}}=708.95$$

**N = 710 subjects (even number so that subjects can be allocated 1:1 to treatment and control arms)**

1. How would the sample size for 90% power change if you had not decided to adjust for age and sex?

$$N= \frac{δ\_{αβ}^{2}V}{∆^{2}}$$

$$N= \frac{\left(Z\_{1-α}+ Z\_{β}\right)^{2}\left(8σ^{2}\left(1-ρ\right)\right)}{(θ\_{1}- θ\_{0})^{2}}$$

$$N= \frac{\left(1.960+1.282\right)^{2}\left(8×39.17^{2}\left(1-0.40\right)\right)}{(-10-0)^{2}}=773.75$$

**N = 774 subjects**

1. What would be the effect on your sample size computation if you had decided to analyze only the final cholesterol measurement adjusted for age and sex (i.e., not the change)? (A qualitative answer is sufficient.)

**Fewer samples would be required to detect a difference in means than a difference in mean changes because the measurements are weakly positively correlated.**

1. What would be the effect on your sample size computation if you had decided to use an Analysis of Covariance model that adjusted for age, sex, and the baseline cholesterol level? (A qualitative answer is sufficient.)

**Fewer samples would be required if an ANCOVA model were used because it allows for a gain in precision from the correlated measurements.**

1. **(A two arm study of cholesterol after 2 years of treatment and the effect of dichotomizing the data)** Suppose we choose to provide the new treatment to *N* subjects. We use as our measure of treatment effect the proportion of subjects having cholesterol below 200 mg/dL at the end of treatment. We are guessing that the new treatment will result instead in an average cholesterol of 135 mm Hg. We intend to perform a hypothesis test in which
* the one-sided level of significance is α = 0.025,
* the desired statistical power is β = 0.90,
* we presume that the proportion *pC* of subjects on the control arm with serum cholesterol below 200 mg/dL will be the same as was observed in the CHS inflamm.txt data set.
* we presume that the treatment will tend to lower serum cholesterol by 10 mg/dL on average, so the proportion *pT* of subjects on the treatment arm with serum cholesterol below 200 mg/dL will be the same as was observed in the CHS inflamm.txt data set for cholesterol levels below 210 mg/dL.
* the measure of treatment effect is *θ1 = pT, - pC* (the difference in the proportion of subjects receiving the new treatment who have cholesterol lower than 200 mg/dL minus the corresponding proportion on the control arm after 2 years of treatment). Under the null hypothesis, we assume there would be no difference between the treatment arms.,
* the average variability contributed by each subject to the estimated treatment effect (the sample proportion) is *V=2( pT,(1- pT, ) + pC (1 - pC ))*(most often, we would compute this under the alternative hypothesis in this setting),
* the comparison between alternative and null hypotheses is *Δ = θ1 - θ0 = θ1*.
1. Using the inflammatory biomarkers dataset, what is your estimate of the proportion *pC* of subjects on the control arm with serum cholesterol below 200 mg/dL at the end of treatment?

**The proportion of subjects in the control arm with serum cholesterol below 200mg/dL at the end of treatment is estimated by the proportion of subjects with serum cholesterol below 200 mg/dL in the inflammatory biomarkers dataset.**

**pC = 0.3957**

1. Using the inflammatory biomarkers dataset, what is your estimate of the proportion *pT* of subjects on the treatment arm with serum cholesterol below 200 mg/dL at the end of treatment? (This is assumed to be equal to the number having cholesterol levels below 210 mg/dL in the CHS data.)

**The proportion of subjects in the treatment arm with serum cholesterol below 200mg/dL at the end of treatment is estimated by the proportion of subjects with serum cholesterol below 210 mg/dL in the inflammatory biomarkers dataset. This is assuming that the treatment will tend to lower serum cholesterol by 10mg/dL on average.**

**pT = 0.4942**

1. What sample size will provide 90% power to detect the design alternative?

$$N= \frac{δ\_{αβ}^{2}V}{∆^{2}}$$

$$N= \frac{\left(Z\_{1-α}+ Z\_{β}\right)^{2}\left(2\left(p\_{T}\left(1-p\_{T}\right)+ p\_{C}\left(1- p\_{C}\right)\right)\right)}{(p\_{T}- p\_{C}-0)^{2}}$$

$$N= \frac{\left(1.960+1.282\right)^{2}\left(2\left(0.494\left(1-0.494\right)+0.396\left(1-0.396\right)\right)\right)}{(0.494-0.396)^{2}}=1059.35$$

**N = 1060 subjects**

1. What advantages or disadvantages does this study design have over the study design used in problem 4b?

**This study design requires a much larger sample size than the study design proposed in problem 4b, however, it represents a clinically relevant treatment effect.**