

**Bios 312: Modern Regression Analysis**  
Spring 2013  
Homework #2

**Written answers due:**

Qstn 1. Consider the following summary data on cholesterol levels by smoking status and gender. The primary research question is to determine if cholesterol levels differ by smoking status. Secondly, the researchers want to describe if the effect of smoking on cholesterol differs by gender. The researchers have decided that they are interested in measuring the difference in mean cholesterol levels (rather than the ratio or some other summary measure).

Table 1: Mean cholesterol (ug/dl) by smoking status and gender

	Men		Women		Combined	
	n	Mean	n	Mean	n	Mean
Smokers	100	200	1000	180	1100	181.8
Non-smokers	1000	240	10000	200	11000	203.6

- 1.1) What is the average effect of smoking on cholesterol in males (the 'male stratum')?
- 1.2) What is the average effect of smoking on cholesterol in females (the 'female stratum')?
- 1.3) Is there any effect modification by gender? If yes, what is the magnitude of the effect modification?
- 1.4) Calculate the weighted average of the stratum specific effects. Specifically, use the average effects found in 1.1 and 1.2 and the number of males and number of females as the weights.
- 1.5) What is the average effect of smoking in the combined sample?
- 1.6) Are there symptoms of gender confounding the association between smoking status and cholesterol in this data? If yes, what are the symptoms? If not, why is gender not a confounder?
- 1.7) Is gender a precision variable? Why or why not?
- 1.8) Table 1 is missing a key summary statistic that would be needed to answer the primary research question. What is it missing? If you had this statistic, how would you summarize the results from Table 1?

Qstn 2. Consider the following summary data on cholesterol levels by smoking status and gender. The primary research question is to determine if cholesterol levels differ by smoking status. Secondly, the researchers want to describe if the effect of smoking on cholesterol differs by gender. The researchers have decided that they are interested in measuring the difference in mean cholesterol levels (rather than the ratio or some other summary measure).

Table 2: Mean cholesterol (ug/dl) by smoking status and gender

	Men		Women		Combined	
	<u>n</u>	<u>Mean</u>	<u>n</u>	<u>Mean</u>	<u>n</u>	<u>Mean</u>
Smokers	200	200	2000	160	2200	163.6
Non-smokers	2000	240	20000	200	22000	203.6

- 2.1) What is the average effect of smoking on cholesterol in males (the 'male stratum')?
- 2.2) What is the average effect of smoking on cholesterol in females (the 'female stratum')?
- 2.3) Is there any effect modification by gender? If yes, what is the magnitude of the effect modification?
- 2.4) Calculate the weighted average of the stratum specific effects. Specifically, use the average effects found in 1.1 and 1.2 and the number of males and number of females as the weights.
- 2.5) What is the average effect of smoking in the combined sample?
- 2.6) Are there symptoms of gender confounding the association between smoking status and cholesterol in this data? If yes, what are the symptoms? If not, why is gender not a confounder?
- 2.7) Is gender a precision variable? Why or why not?

Qstn 3. Consider the following summary data on cholesterol levels by smoking status and gender. The primary research question is to determine if cholesterol levels differ by smoking status. Secondly, the researchers want to describe if the effect of smoking on cholesterol differs by gender. The researchers have decided that they are interested in measuring the difference in mean cholesterol levels (rather than the ratio or some other summary measure).

Table 3: Mean cholesterol (ug/dl) by smoking status and gender

	Men		Women		Combined	
	<u>n</u>	<u>Mean</u>	<u>n</u>	<u>Mean</u>	<u>n</u>	<u>Mean</u>
Smokers	350	260	50	200	400	(missing)
Non-smokers	50	280	350	200	400	(missing)

3.1 Give the effect of smoking on cholesterol in males, females, and in the combined sample. (You will need to calculate the means for smokers and non-smokers in the combined sample).

3.2 Are there symptoms of gender confounding the association between smoking status and cholesterol in this data? If yes, what are the symptoms? If not, why is gender not a confounder?

3.3 Is there evidence of effect modification by gender? If yes, what is the magnitude of the effect modification?

3.4 Is gender a precision variable? Why or why not?

3.5 Based on the data in Table 3 only, do you conclude that there is an association between smoking and mean cholesterol? Why or why not?

3.6 Suppose that we fit a linear regression model with cholesterol as the outcome with predictors for gender and smoking status. From this model, what would be the estimated effect of smoking on cholesterol controlling for gender? Note that you do not need to fit a regression model to answer this question; it can be answered from the information in the table alone.