- (1) An article in the New York Times (February 17, 1999) about the PSA blood test for detecting prostate cancer stated that, of men who had this disease, the test fails to detect prostate cancer in 1 in 4 (so called false-negative results), and of men who did not have it, as many as two-thirds receive false-positive results. Let $C(\overline{C})$ denote the event of having (not having) prostate cancer and let +(-) denote a positive (negative) test result.
 - (a) Which is true: P(-|C) = 1/4 or P(C|-) = 1/4? $P(\overline{C}|+) = 2/3$ or $P(+|\overline{C}) = 2/3$?
 - (b) What is the sensitivity of this test?
 - (c) Of men who take the PSA test, suppose P(C) = 0.01. Find the cell probabilities in the 2 × 2 table for the joint distribution that cross classifies Y = diagnosis(+, -) with $X = \text{true disease status } (C, \overline{C})$.
 - (d) Using (c), find the marginal distribution for the diagnosis.
 - (e) Using (c) and (d), find P(C|+), and interpret.
- (2) According to recent UN figures, the annual gun homicide rate is 62.4 per one million residents in the United States and 1.3 per one million residents in the UK.
 - (a) Compare the proportion of residents killed annually by guns using the (i) difference of proportions, (ii) relative risk.
 - (b) When both proportions are very close to 0, as here, which measure is more useful for describing the strength of association? Why?
- (3) Data posted at the FBI website (www.fbi.gov) stated that of all blacks slain in 2005, 91% were slain by blacks, and of all whites slain in 2005, 83% were slain by whites. Let Y denote race of victim and X denote race of murderer.
 - (a) Which conditional distribution do these statistics refer to, Y given X, or X given Y.
 - (b) Calculate and interpret the odds ratio between X and Y.
 - (c) Given that a murderer was white, can you estimate the probability that the victim was white? What additional information would you need to do this? (Hint: How could you use Bayes' Theorem?)
- (4) The following table comes from one of the first studies of the link between lung cancer and smoking, by Richard Doll and A. Bradford Hill. In 20 hospitals in London, UK, patients admitted with lung cancer in the previous year were queried about their smoking behavior. For each patient admitted, researchers studied the smoking behavior of a non-cancer control patient at the same hospital of the same sex and within the same 5-year grouping on age. A smoker was defined as a person who had smoked at least one cigarette a day for at least a year.

Have Smoked	Lung Cancer	
	Cases	Controls
Yes	688	650
No	21	59
Total	709	709

Based on data reported in Table IV, R. Doll and A. B. Hill, *Br. Med. J.*, 739–748, September 30, 1950.

- (a) Identify the response variable and the explanatory variable.
- (b) Identify the type of study this was.
- (c) Can you use these data to compare smokers with nonsmokers in terms of the proportion who suffered lung cancer? Why or why not?
- (d) Summarize the association, and explain how to interpret it.