CPH 931 — Fall 2009 — Dr. Charnigo

Final Examination

This non-collaborative take-home final examination may be submitted any time between 12 Noon on Monday 14 December and 12 Noon on Friday 18 December. By non-collaborative I mean that you are not permitted to discuss the examination with anyone other than me, until after the deadline for submission. The examination is to be submitted in hard copy, to me in person or under my office door (CPH 203-B), or electronically, to {RJCharn2@aol.com} in either Word 2003 or PDF.

[50] 1. Refer to {http://clincancerres.aacrjournals.org/content/10/1/113/T2.expansion.html}, which provides results for a case-control study assessing whether there might be an association between hormone replacement therapy and lung cancer.

[10] a. Show how the authors calculated the estimated odds ratio of 0.78 and the accompanying 95% confidence interval of 0.61 to 1.01 in the "Univariate OR (95% CI)" column.

[10] b. An unsophisticated reader states, "The estimated risk of lung cancer among those with hormone replacement therapy is $232/505 = 0.459$, and the estimated risk of lung cancer among those without hormone replacement therapy is $267/513 = 0.520$. So the estimated relative risk is $0.459/0.520 = 0.883$." Explain why the above statement is in error.

[10] c. Let r denote the risk of developing lung cancer in the population to which this study generalizes. Using the authors' data, employ Bayes' Theorem to estimate the (unadjusted) relative risk as a function of r. What happens to your estimate of the relative risk as r approaches 0? Is there any value of r that reproduces the unsophisticated reader's estimate from part b?

[10] d. Another unsophisticated reader states, "Hormone replacement therapy reduces the risk of lung cancer by 22%." Explain why the above statement is in error, but then indicate why the above statement also has a kernel of truth to it.

[10] e. Most authors prefer to report "positive" results, since those are far easier to get published. Why do you suppose that these authors went to press with a confidence interval that just barely included 1 when they might have acquired a little more data and reduced the length of the confidence interval enough to exclude 1?

[50] 2. The file {CigaretteConsumption.xls} contains historical data on state-level variables thought to be related to annual cigarette sales (Sales, expressed in millions of dollars), including median age (Age), percent of residents with a high school education (HS), per capita income (Income, expressed in dollars), percent of African-American residents (Black), percent of female residents (Female), and average price per pack of cigarettes (Price, expressed in cents — yes, this is a very old data set!).

This exercise does not have parts a, b, c, d, and e. Rather, your task is to write a "Case Study" illustrating appropriate biostatistical methodology from CPH 930 and CPH 931. Be selective in what you present. Do not try to demonstrate every idea that you learned in CPH 930 and CPH 931. Rather, synthesize what you learned in CPH 930 and CPH 931 to address as best you can the question of what drives cigarette sales at the state level.

Note: You are limited to a maximum of 1000 words and three tables/figures.