Biost 517: Applied Biostatistics I
Emerson, Fall 2010

Homework #8 Key
November 29, 2010

Written problems: To be handed in at the beginning of class on Friday, December 3, 2010.

Questions for Biost 514 and Biost 517:

Problems 1 and 2 refer to data on the usefulness of post-treatment PSA levels for prognosis in hormonally treated prostate cancer (the file psa.txt on the class web pages). We consider several alternative strategies to assess whether there is an association between time in remission and nadir PSA level. In all problems, provide relevant descriptive statistics and as complete statistical inference as possible (i.e., provide point estimates, confidence intervals, and p values where possible, along with a statement of your scientific/statististical conclusions). (Note that these problems are additional alternatives to the analyses performed in homework #7.)

1. Base your analysis on a comparison of the probability of remaining in remission for 36 months as a function of nadir PSA levels above a threshold of 2 ng/ml.

Ans: Fifty men with hormonally treated prostate cancer were followed for cancer relapse for an average of 52.1 months (range 24 – 75 months). During the period of follow-up, relapse was observed in 36 patients.

The nadir PSA was observed to be greater than 2 ng/ml in 19 subjects, and to be less than or equal to 2 ng/ml in 31 subjects. Using Kaplan-Meier estimates and Greenwood's formula, the estimated probability of relapse-free survival at 3 years was 15.8% (SE 8.37%) in the group having higher nadir PSA, and 59.9% (SE 9.04%) in the group having lower nadir PSA. The comparison of the two groups defined by nadir PSA thus estimates that the group with higher nadir PSA had a probability of relapse within 36 months that was 44.1% higher (absolute difference) than that in the low nadir PSA group. Such an observation is statistically different from 0 (two-sided P = 0.0003), with a 95% confidence interval suggesting that the observed result is not unusual if the true absolute difference in relapse probabilities at 3 years were anywhere from 19.9% higher to 68.2% higher for the group with nadir PSA greater than 2 ng/ml than for the group with the lower nadir PSA.

2. Base your analysis on the probability that a randomly chosen individual who relapsed within 24 months might have a higher nadir PSA than a randomly chosen individual who did not relapse within 24 months.

Ans: Because no subjects were censored prior to 24 months, we know for each subject whether he relapsed within 24 months. We are thus free to divide this cross-sectional sample into groups defined by whether the patient relapses within 24 months.

The study data consist of measurements made on 50 consecutive hormonally treated prostate cancer patients. Relapse within 24 months was observed in 22 patients, and the remaining 28 patients stayed in remission for at least 24 months. The probability that a randomly chosen individual who relapsed within 24 months might have a higher nadir PSA than a randomly chosen individual who remained in remission for 24 months is estimated to be 88.2%. Such an observation is statistically different from the 50% probability that would

be anticipated if nadir PSA were totally independent of time in remission (two-sided P < .0001 by the Wilcoxon rank sum test). *(Note that when sticking strictly to the rank sum statistic, I do not get a scientific estimate that gives me an idea about clinical importance. Were I to use this statistic in real-life (perhaps with a gun pointed at my head), I would in fact have described the distribution of nadir PSA values for the two groups with means, geometric means, medians, etc., and then provided the wording that I gave above. But I would far rather have known that a scientifically useful summary of the distribution was statistically significant.)*

Problem 3 and 4 refers to the laboratory data for the clinical trial of methotrexate in PBC as stored on the class web pages. In this homework we will consider the serum bilirubin value five years post randomization (week 260) for those patients for whom data is available. These measurements will be compared to patient's baseline (week 1) measurement. In all problems, provide relevant descriptive statistics and as complete statistical inference as possible (i.e., provide point estimates, confidence intervals, and p values where possible, along with a statement of your scientific/statististical conclusions). (Note that these problems are alternatives to the analyses performed in homework #6.)

3. Use the sign test to perform an analysis to assess whether the placebo group tended to have lower bilirubin levels 5 years post randomization (week 260) than they did at the time of randomization (week 1). How does this analysis differ from problem 2 of homework #6? (This analysis can be effected using the Stata command

   `signtest` *`var1=var2`*

   which considers the proportion of positive differences is less than or greater than the proportion of negative differences – differences of zero are ignored.)

<u>Ans:</u> One hundred thirty-three (133) subjects were randomized to receive placebo, and all of those patients had bilirubin measurements available at time of randomization. However, only 116 of those subjects had bilirubin measurements available at 5 years. Analysis of the change in bilirubin levels in the placebo group is restricted to those 116 patients with data available at 5 years.

Over the 5 year treatment period, the mean change in bilirubin levels in the placebo group showed an increase of 0.308 mg/dl (SD 1.34 mg/dl, range -1.2 to 10 mg/dl), with 25th, 50th, and 75th percentiles of -0.1 mg/dl, 0 mg/dl, and 0.25 mg/dl, respectively, where a positive number indicates an increase in bilirubin. The observed median difference of 0.0 mg/dl was not statistically significantly different from 0 (P= 0.151 in a two-sided sign test), with a 95% confidence interval suggesting that the observed result is not unusual if the true median difference in bilirubin measurements within an individual over 5 years were anywhere from 0.0 mg/dl to 0.10 mg/dl increase. *(Note that no 95% CI is readily obtained from Stata in the "signtest" command, though one is possible by "inverting" the sign test. In doing so, we would perform hypothesis tests of other values (i.e., nonzero) until we just barely had a one-sided P value of 0.025 and then again until we just barely had a one-sided P value of 0.975. When I did this in an iterative search (not too difficult because bilirubin was only measured to a precision of 0.01 mg/dl at most), I find a 95% CI of 0.00 mg/dl to 0.10 mg/dl. That is, when executing*

signtest diffbili = x if week==1 & tx==0

any value of x less than 0.00 had an upper one-sided P value < 0.025, and any value greater than 0.10 had an upper one-sided P value > 0.975. Alternatively, we could have used

centile diffbili if week==1 & tx==0

which does exactly the same thing, only more easily.)

4. Use the sign rank test to perform an analysis to assess whether the placebo group tended to have lower bilirubin levels 5 years post randomization (week 260) than they did at the time of randomization (week 1). How does this analysis differ from problem 3 above? (This analysis can be effected using the Stata command

   `signrank var1=var2`

   which considers whether the ranks of the absolute differences are evenly split between the positive differences and the negative differences– differences of zero are ignored.)

Ans: One hundred thirty-three (133) subjects were randomized to receive placebo, and all of those patients had bilirubin measurements available at time of randomization. However, only 116 of those subjects had bilirubin measurements available at 5 years. Analysis of the change in bilirubin levels in the placebo group is restricted to those 116 patients with data available at 5 years.

Over the 5 year treatment period, the mean change in bilirubin levels in the placebo group showed an increase of 0.308 mg/dl (SD 1.34 mg/dl, range -1.2 to 10 mg/dl), with 25th, 50th, and 75th percentiles of -0.1 mg/dl, 0 mg/dl, and 0.25 mg/dl, respectively, where a positive number indicates an increase in bilirubin. The observed data is statistically significantly different from what might be expected if the distribution of bilirubin values were exactly the same at baseline and after 5 years of follow-up (P= 0.0278 in a two-sided signed rank test). *(Note that by examining the sum of the ranks versus expected, it does appear that there is some sort of tendency toward higher values over the 5 year period, but it is hard to differentiate from the Stata output whether this is driven primarily by a tendency toward more subjects with increases or by a tendency for any such increases to be larger in absolute value than the decreases or both. The actual functional tested by the signed rank test the median of something called the Walsh averages-do you find that knowing that is helpful?)*