Introduction to Biostatistics (171:161)
Breheny

Assignment 9
Due: Tuesday, April 8

1. In population A, the standard deviation of LDL cholesterol is 20 mg/dL, and in population B, the standard deviation is 40 mg/dL. An investigator collects random samples of 10 individuals and uses Student's curve to calculate a 95% confidence interval from each population. Which of the following is true: (i) the confidence interval for population A will definitely be wider (ii) the confidence interval for population A will probably be wider (iii) the two confidence intervals should be about equally wide (iv) the confidence interval for population B will probably be wider (v) the confidence interval for population B will definitely be wider

2. People with diabetes often sustain vascular damage to their retina (diabetic retinopathy), which can lead to blindness. Levels of vascular endothelial growth factor (VEGF) in the eye are highly correlated with retinopathy, and often used as a marker of the severity of the damage. In one study published in the *New England Journal of Medicine*, researchers performed a treatment called laser photocoagulation on seven patients with diabetic retinopathy. They measured the ocular VEGF levels before and after treatment and found that VEGF levels were lower by an average of 4.2 ng/mL, with a standard deviation of 2.85 ng/mL.

   (a) What is wrong with the following argument? "4.2 is only about one and a half standard deviations away from 0. Differences need to be 2 or more standard deviations away in order to be significant. Therefore, this evidence is not statistically significant."

   (b) Construct a 95% confidence interval for the average decrease in ocular VEGF levels following treatment.

   (c) Conduct a paired t-test of the hypothesis that the treatment has no effect on VEGF levels.

   (d) Does laser photocoagulation seem to help or hurt patients with diabetic retinopathy? Or is difficult to say, because this study provides little evidence of an effect in either direction?

3. ⌨ This problem involves the same crossover study of oat bran diets and serum cholesterol levels that was in assignment 6.

   (a) Perform a paired t-test of the hypothesis that oat bran consumption has no effect on serum cholesterol.

   (b) Construct a 95% confidence interval for the average amount by which an oat bran diet could be expected to reduce a man's serum LDL cholesterol levels (as compared to a corn flake diet).

   (c) Compared with the binomial test (the test you performed in assignment 6), name an advantage of the paired t-test.

   (d) Compared with the binomial test, name a disadvantage of the paired t-test.

1

4. ⌨ You have developed a new, cholesterol-lowering drug as a treatment for patients with high cholesterol. You are planning a crossover study in which each patient receives both the new drug and a placebo. You expect that the standard deviation is around 45 mg/dl, that your new drug will reduce cholesterol levels by about 10 mg/dl on average, and that the placebo will have no effect on patients' cholesterol levels.

   (a) If your sample size is 50, what is your power?

   (b) If instead of 50, you enroll 100 subjects in your study, what is the power?

   (c) What sample size do you need to achieve 80% power?

5. (a) While performing a χ^2-test, a researcher obtains a χ^2-statistic of 4.70. What is his two-sided p-value?

   (b) While performing a χ^2-test, a researcher obtains a two-sided p-value of 0.16. What was her χ^2-statistic?

6. Suppose you work at the Gap in the Coral Ridge Mall. Your coworker is bored and bets you that you can't predict the number of customers that you will have in the next hour to within 5. Which is better for your chances of winning the bet: if this bet takes place on a Monday during the summer, or a Saturday two weeks before Christmas?