

Stat 518 - Homework 6

Due: Wednesday, November 8th

1) The web-site has the data from an article in the journal *Lancet* concerning the consumption of wine (in liters per person per year) and the number of heart disease deaths (per 1,000 men aged 55 to 64) in 18 industrialized countries.

a) By hand, calculate Kendall's τ for this data set.

b) Find a formula, using just the sample size n and the observed τ value, that gives an estimate of the probability that the observed slope for a randomly selected pair of observations will be positive. Estimate that probability for this example. (You may assume that there are no ties in coming up with the formula.)

c) Use S-Plus or SAS to perform a simple linear regression using the standard least squares method to estimate the heart disease death rate from the wine consumption. Check the assumptions and report the results of the regression.

d) Use S-Plus to perform the nonparametric linear regression method based on Kendall's τ to estimate the heart disease death rate from the wine consumption. (This method is described in the text starting on the bottom of page 335 and continues through the bottom of page 336.) Compare the result to that of part c.

e) Use S-Plus's kernel smoothing procedure to nonparametrically estimate the curve for this problem. Try several bandwidths and choose the one you think works the best. Briefly defend your choice and comment on its shortcomings. Compare its effectiveness to the methods in part c and d.

2a) Page 442 number 6 using S-Plus.

b) Construct an estimated power curve for this test of hypothesis using 100 simulation runs, $\alpha=0.05$, true mean = 5.6, and true standard deviations of 0.6, 1.2, 1.8, and 2.4.