Homework 11

625.433

1. (50 pts.) The data set entitled 'Salmon' contains 40 annual counts of the numbers of recruits and spawners in a salmon population. The units are thousands of fish. Recruits are fish that enter the catchable population. Spawners are fish that are laying eggs. Spawners die after laying eggs.

   The classic Beaverton-Holt model for the relationship between spawners and recruits is

   $$R = \frac{1}{\beta_1 + \beta_2/S},$$

   where R and S are the numbers of recruits and spawners, respectively. This model may be fit using linear regression with the transformed variables $\frac{1}{R}$ and $\frac{1}{S}$.

   Now consider the problem of maintaining a sustainable fishery. The total population abundance will only stabilize if $R = S$. The total population will decline if fewer recruits are produced than the number of spawners who died producing them. If too many recruits are produced, the population will also decline eventually because there is not enough food for them all. Thus, only some middle level of recruits can be sustained indefinitely in a stable population. This stable population level is the point where the 45-degree line intersects the curve relating R and S.

   (a) Fit the Bevearton-Holt model and find a point estimate for the stable population level where $R = S$. Use the bootstrap to obtain a corresponding 95% confidence interval and a standard error (estimated variance) for your estimate, from two methods: bootstrapping the residuals and bootstrapping the cases. Histogram each bootstrap distribution, and comment on the differences in your results.

   (b) Do you believe there is any bias in your estimate? Defend your answer using the bootstrap.

2. (30 pts.) Generate 100 samples $X_1, X_2, \ldots, X_{20}$ from a normal population $N(\theta, 1)$. with $\theta = 1$.

   (a) For each sample compute the bootstrap and jackknife estimate for variance for $\hat{\theta} = \overline{X}$ and compute the mean and standard deviation of these variance estimates over the 100 samples.

   (b) Repeat (a) for the statistic $\hat{\theta} = \overline{X}^2$, and compare the results. Give an explanation for your findings.

3. (20 pts.) Find the number of bootstraps necessary to accurately calculate a 95% confidence interval in the difference of two population means. That is, simulate n values of X, where $X_i \sim N(\mu_1, \sigma_1^2)$ and m values of Y, where $Y \sim N(\mu_2, \sigma^2)$ and calculate a 95% confidence interval in $\mu_1 - \mu_2$. Now calculate a 95% confidence interval for $\mu_1 - \mu_2$ using the bootstrap techniques learned in this module. How large does B have to be until you accurately and dependably reproduce the confidence interval you originally calculated?