Department of Mathematics and Statistics

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

CSE Exercises - Week 10

(1) Revisiting exercise 2 in week 9, do the following using the data in normal_data.mat:

(a) Plot the empirical DF together with the 95% confidence band given by the DKW inequality.

(b) Use the nonparametric bootstrap to estimate the standard error of $\hat{q}$. You should use the same number of bootstrap realizations (N) as for exercise 2(b)(ii) in week 9.

(c) Compare and comment on your estimates of the standard error using the two bootstrap methods, i.e the nonparametric bootstrap in part (b) above and the parametric bootstrap result from week 9.

(2) Revisiting exercise 3 in week 9, do the following using the data in beta_data.mat:

(a) Plot the empirical DF together with the 95% confidence band given by the DKW inequality.

(b) Use the nonparametric bootstrap to find approximate 95% confidence intervals for α, β and m, based on ML estimates $\hat{\alpha}$, $\hat{\beta}$ and $\hat{m}$.

(c) Compare and comment on the results from part (b) above with those obtained using the parametric bootstrap from exercise 3(c) in week 9.

(d) Now repeat for the larger data set with $n = 100$ given in beta_data2.mat, i.e. find approximate 95% confidence intervals for α, β and m based on ML estimates using both parametric and nonparametric bootstraps. Comment on your results vis-a-vis those for $n = 25$.

(3) Here is an example where the nonparametric bootstrap performs poorly.

Let $X_1, \ldots, X_n \overset{IID}{\sim} \text{Uniform}(0, \theta)$. Recall from exercise 1 in week 9 that the ML estimator for θ is $\hat{\theta} = X_{(n)} = \max\{X_1, \ldots, X_n\}$. Suppose that $\theta = 1$ and $n = 50$.

(a) Use Monte Carlo simulation to get an idea of the true distribution of $\hat{\theta}$. To do this, generate $N = 1000$ realizations, each containing 50 sample values from Uniform(0, 1). For each realization, find $\hat{\theta}$. Plot a density histogram for the 1000 $\hat{\theta}$ estimates.

(b) Now suppose that our observed data values are contained in the file uniform_data.mat. Use the parametric bootstrap to obtain $N = 1000$ bootstrap estimates, $\hat{\theta}_1, \ldots, \hat{\theta}_N$. Plot the density histogram for these estimates.

(c) Repeat part (b) using the nonparametric bootstrap.

(d) Compare and comment on the 3 density histograms obtained.