STAT743 FOUNDATIONS OF STATISTICS (Part II)

Assignment 4 Due in class at 12:30pm on **Monday, April 9** 2018

Instructions:

1. Ensure that your full name and student number are clearly marked on your solutions.
2. Solutions need not be typed (although they can be) but must be readable.
3. Start each question on a new page and clearly indicate where each part of the question begins.
4. Submit questions in the same order as given below.
5. You are expected to show all details of your solution and any results taken from my notes or the textbook must be clearly and properly referenced.
6. You must include a commented listing of your R code for any question which requires the use of computing.
7. No extensions to the due date and time will be given except in extreme circumstances and late assignments will not be accepted. Assignments are due at the **start** of class.
8. Students are reminded that submitted assignments must be entirely their own work. Submission of someone else's solution under your name is academic misconduct and will be dealt with as such. Penalties for academic misconduct can include a 0 for the assignment, an F for the course with an annotation on your transcript and/or dismissal from your program of study.

Q. 1 Suppose that $X_1, \ldots, X_n$ is a random sample from an exponential distribution with rate parameter λ and we are interested in constructing confidence intervals for λ.

a) Derive an exact confidence interval for λ based on the distribution of the maximum likelihood estimator $\hat{\lambda}$.

b) Derive an asymptotic confidence interval based on inversion of the likelihood ratio test.

c) Use the Delta method to find the asymptotic variance of $\hat{\lambda}$ (σ_n in my notes) in terms of λ and hence construct a Wald confidence interval for λ.

d) Derive an asymptotic score confidence interval and show that it is equivalent to a Wald confidence interval if we estimate σ_n with its mle but different from the Wald interval that does not estimate σ_n as found in part (c).

e) Suppose that the following data come from an exponential distribution.

18.5	3.5	5.3	3.3	9.4	10.9	21.6	53.4	5.3	33.7

Construct each of the four intervals for this dataset and compare the asymptotic intervals with the exact interval. Which of the asymptotic methods best approximates the exact method?

Q. 2 Casella and Berger 10.31

Q. 3 Let $X_1, \ldots, X_n$ be *iid* Bernoulli(θ) and suppose that interest is in estimation of θ^2.

a) Find the maximum likelihood estimator of θ^2 and show that it is a biased estimator.

b) Show that the jackknife bias estimate of the mle is

$$\hat{b}_{\text{jack}} = \frac{S(n-S)}{n^2(n-1)}$$

where $S = \sum X_i$.

c) The jackknife can be used to reduce the bias in an estimator by defining the *jackknife bias-adjusted estimator*

$$\hat{\theta}_{\text{jack}} = \hat{\theta} - \hat{b}_{\text{jack}}$$

Find the jackknife bias-adjusted estimator for this example and show that it is the minimum variance unbiased estimator of θ^2.

[*In general the bias-adjusted estimator only reduces the bias, this is a special case in which the bias is completely removed by the adjustment.*]

Q. 4 Suppose that $X_1, \ldots, X_n$ are *iid* from a population with cdf F_1 and $Y_1, \ldots, Y_m$ is an independent sample from cdf F_2 and that both populations have support on positive values only. The parameter of interest is

$$\theta \;=\; t(F_1, F_2) \;=\; \frac{\int x\, dF_1(x)}{\int y\, dF_2(y)} \;=\; \frac{\mu_1}{\mu_2}$$

Let $\hat{\theta} \;=\; \overline{X}/\overline{Y}$ be the estimator of θ.

a) Show that the influence functions are

$$L_{t,1}(x; F_1, F_2) \;=\; \frac{x-\mu_1}{\mu_2}$$

$$L_{t,2}(y; F_1, F_2) \;=\; -\frac{(y-\mu_2)\theta}{\mu_2}$$

b) Show that the infinitesimal jackknife standard error is given by

$$\mathrm{se}(\hat{\theta}) = \frac{\sqrt{n^{-2}\sum(x_i - \overline{x})^2 + m^{-2}\hat{\theta}^2\sum(y_j - \overline{y})^2}}{\overline{y}}$$

c) The following data give the energy expenditure in a day measured for a group of $n = 9$ obese women and $m = 13$ lean women.

Obese Women

9.21	11.51	12.79	11.85	9.97	8.79	9.69	9.68	9.19

Lean Women

7.53	7.48	8.08	8.09	10.15	8.40	10.88	6.13	7.90	7.05	7.48	7.58	8.11

Estimate the bias and the variance of the ratio of means using

(i) the infinitesimal jackknife (influence values),

(ii) the regular jackknife, and

(iii) the non-parametric bootstrap.

Include `R` code to do this **without** using the `boot` library.