STAT743 FOUNDATIONS OF STATISTICS (Part II)

Assignment 1 — Due in class at 12:30pm on February 1, 2018

Instructions:

1. Ensure that your full name and student number are clearly marked on each page of 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 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 This question describes two ways of generating normal random variates from uniforms.

a) Casella and Berger 5.50

b) Consider the following accept-reject algorithm

1. Generate 2 independent $\text{Unif}(0,1)$ random variables U_1 and U_2.

2. Let $X = -\ln U_1$ and $Y = -\ln U_2$.

3. If $2X \geqslant (Y-1)^2$ return Y.

(i) Show that X and Y are independent $\exp(1)$ random variables.

(ii) Use Bayes Theorem to show that

$$f_Y\left(y|2X \geqslant (Y-1)^2\right) \propto \mathrm{e}^{-y^2/2} \quad \text{for } y > 0.$$

(iii) Using a transformation show that

$$\int_0^\infty \mathrm{e}^{-y^2/2}\, dy = \frac{\Gamma(\frac{1}{2})}{\sqrt{2}} = \sqrt{\frac{\pi}{2}}$$

and hence give the density of the accepted observations.

(iv) Suppose that Y is generated using this method and U_3 is another $\text{Unif}(0,1)$ random variable. Define the random variable Z as

$$Z = \begin{cases} Y & \text{if } U_3 > 0.5 \\ -Y & \text{if } U_3 \leqslant 0.5 \end{cases}$$

Show that Z is a standard normal random variable.

Q. 2 Suppose we want to generate n random vectors from the bivariate normal distribution (see textbook pages 175–177) with $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$ and correlation $\rho = 0.7$. Explain how to use the following three methods to do this and implement the methods in R. Start the chains at the stationary distribution by setting your initial value using the `mvrnorm` in `library(MASS)`.

a) An independence Metropolis–Hastings algorithm generating the candidate random vectors as independent standard normal random variables.

b) A random walk Metropolis–Hastings algorithm where the candidate random vectors are generated from independent normals with variances equal to 1 centred at the current value of the chain.

c) A Gibbs Sampler.

Q. 3 Suppose that $X_1, \ldots, X_n$ are a random sample from a Normal population with unknown mean μ and variance σ^2. We wish to estimate the population standard deviation σ.

a) For a fixed value of n find the constant c_n such that $c_n S$ is an unbiased estimator of σ where S^2 is the usual unbiased estimator of σ^2.

b) Find the efficiency of the estimator from (a) as a function of the sample size n and draw a plot to describe how the efficiency varies as n increases.

c) Examine the ratio of the mean squared error of the unbiased estimator to that of the sample standard deviation, S, for different values of n.

d) What conclusions would you draw from the analyses in (b) and (c).

Q. 4 **a)** Casella and Berger 7.39

b) Casella and Berger 7.60, also show that the best unbiased estimator does not achieve the Cramér-Rao Lower Bound.