University of Manitoba
Department of Statistics

STAT 4600: Computational Statistics

Assignment 6 – Part I

Due: Tuesday, March 27

Question 1:

Consider the density function $f(x) = h(x)/3$, where

$$h(x) = \begin{cases} 2-(x+1)^2 & \text{for } -1 < x \leq 0, \\ 1 & \text{for } 0 < x \leq 1, \\ (x-2)^2 & \text{for } 1 < x \leq 2. \end{cases}$$

(A) Construct an accept-reject sampler for f based on a rectangle.

(B) Construct an accept-reject sampler based on the candidate density $g(x) = \frac{2}{9}(2-x)$.

(C) Find, with justification, a better candidate density of the form

$$g(x) = a + bx \qquad\qquad \text{for } -1 < x < 2,$$

and construct a third accept-reject simulator for f based on it.

(D) Visually confirm, using histograms, that the three simulators produce samples according to f and compare their performance via acceptance probabilities.

Question 2:

Consider the integrals

$$I_1 = \int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi},$$

and

$$I_2 = \int_{\mathbb{R}} \frac{dx}{1+x^2} = \pi.$$

Now, consider approximating these integrals via Monte Carlo and an importance sampling strategy. Specifically, consider

(A) approximating I_1 using a Cauchy importance function;

(B) approximating I_2 using a Normal importance function.

One of the above strategy fails. Identify clearly which one that is and justify your answer. Also, run these Monte Carlo experiments and produce appropriate output to confirm your claim.

University of Manitoba
Department of Statistics

STAT 4600: Computational Statistics

Assignment 6 – Part II

Due: Tuesday, March 27

Question 3:

The Weibull family of distributions is often used to model lifetimes. Specifically, the Weibull density with parameters $a > 0$ and $b > 0$ is given by

$$f(x; a, b) = \frac{a}{b}\left(\frac{x}{b}\right)^{a-1} e^{-(x/b)^a},$$

for $x > 0$. Note that the above reduces to an exponential distribution when $a = 1$.

Now, in a specific application, 23 devices were subject to life testing, leading to the following failure times (in hours) for 20 of the devices:

8.71,	1.57,	4.22,	6.38,	1.72,	7.70,	8.43,	2.99,	7.45,	4.71,
9.25,	6.05,	7.32,	4.99,	4.08,	7.59,	5.52,	2.63,	6.24,	5.90.

Due to a recording error, one of the failures is known to have occurred between 3 and 4 hours after the start of the experiment, but is not otherwise known precisely. Also, two devices had still not failed at the conclusion of the experiment, after 10 hours.

(A) What is an appropriate likelihood function for $\theta = (a, b)'$ in this case?

(B) Write an R function that calculates the MLEs of a and b from observed data (including potential non-failures and any interval censured observations).

Use your function to find the MLE of a and b from the above data.

(C) Using the standard (nonparametric) bootstrap, construct confidence intervals for each of a and b of approximate level 95%.

Note: Have a look at the two examples involving the Gamma distribution