South China University of Technology

Introduction to Statistical Programming With R

Final Examination

Part A contains 10 short answer questions, each worth 3 marks. Part B contains 10 multiple choice questions, each worth 3 marks. Part C contains 4 written questions, each worth 10 marks.

Total Marks Available = **100**. Answer all questions.

Part A – Short Answer (30 Marks)

In each case below, write down the response (if any) that you would see in the R console window, if the given commands were typed into the console.

1. ```
   > x <- c(1, 1, 2, 3, 5, 8, 13)
   > x[x > 4]
   ```
2. ```
   > x <- c(1, 1, 2, 3, NA, 8, 13)
   > x[x == NA] <- 5
   > x
   ```
3. ```
   > x <- rep(1:3, 4)
   > x
   ```
4. ```
   > x <- rep(-(1:3), 4)
   > x
   ```
5. `> 51 %/% 6`
6. `> 51 %% 6`
7. `> x <- rep(1:3, 2);    y <- seq(1,6);     x + y`
8. `> x <- 3^3000`
9. `> for (i in seq(22, 30, 3)) print(i %% 7)`
10. `> x <- rep(3:5, rep(2, 3));   x[x==4] <- NA;    x > 3`

Part B – Multiple Choice (30 Marks)

For questions 1 through 10, write down the letter corresponding to the correct response.

1. Find the decimal representation of the number 0.101_2.

   (a) 0.400 (b) 0.325 (c) 0.750 (d) 0.875 (e) 0.625

2. Assuming 4-digit binary accuracy, find the representation of the number 5/7.

   (a) .1110 (b) .0111 (c) .1001 (d) .1101 (e) .1011

3. Use Newton's method with an initial guess of $x = -1.5$ to determine the real valued solution of the equation

   $$x^3 - x + 1 = 0.$$

   (a) -2.500 (b) -1.500 (c) -1.325 (d) -1.115 (e) -0.800

Suppose 6 independent uniform random numbers have been generated:

```
0.15   0.92   0.07   0.49   0.63   0.17
```

Use these numbers to answer questions 4 through 7.

4. Use the above numbers to simulate 6 independent Bernoulli random variables with $p = 0.4$.

   (a) `1  9  0  4  6  1`   (b) `2  9  1  5  6  2`   (c) `1  1  0  0  0  1`
   (d) `1  0  1  0  0  1`   (e) `0  1  1  2  3  3`

5. Now, simulate 2 Binomial random variables with $n = 3$ and $p = 0.4$.

   (a) `10  11`   (b) `12  13`   (c) `2   1`   (d) `0   3`   (e) `2   3`

6. Now, simulate 6 independent uniform numbers on the interval $[3, 5]$.

   (a) `3.30 4.84 3.14 3.98 4.26 3.34`   (b) `2.45 4.76 2.21 3.47 3.89 2.51`
   (c) `6.45 8.76 6.21 7.47 7.89 6.51`   (d) `6.30 7.84 6.14 6.98 7.26 6.34`
   (e) `3.15 3.92 3.07 3.49 3.63 3.17`

7. Now, simulate 6 independent exponential random variables with rate 1.

   (a) `3.897 1.083 2.659 0.992 0.046 1.772`
   (b) `0.163 2.526 0.073 0.673 0.994 0.186`
   (c) `3.897 1.083 2.659 0.713 0.462 1.772`
   (d) `3.151 3.922 3.073 3.494 3.635 3.176`
   (e) `0.863 1.526 2.073 0.673 0.994 0.186`

8. Consider the function

```
unknown.fn <- function() {
    x <- rexp(100000)
    mean(sin(x))
}
```

   This function computes a numerical estimate of

   (a) $\sum_{i=1}^{\infty} \sin(u_i)$ (b) $\int_0^\infty \sin(e^{-x})dx$ (c) $\int_0^1 \sin(u)du$
   (d) $\int_0^1 \sin(x)e^{-x}dx$ (e) $\int_0^\infty \sin(x)e^{-x}dx$

9. What is the output from the following commands?

```
x <- seq(2,20,3)
x[x>15 | x < 6]
```

   (a) `2  4 16 18 20`
   (b) `2  20`
   (c) `2  4  20`
   (d) `2  5 17 20`
   (e) `numeric(0)`

10. Determine the output that results from typing

```
rep(rep(seq(-1,1),rep(2,3)),2)
```

(a) `-1 -1  1  1 -1 -1  1  1 -1 -1  1  1 -1 -1  1  1`
(b) `-1  0  1 -1  0  1 -1  0  1 -1  0  1`
(c) `-1 -1  0  0  1  1`
(d) `-1 -1  1  1 -1 -1  1  1`
(e) `-1 -1  0  0  1  1 -1 -1  0  0  1  1`

Part C – Written Questions (40 Marks)

1. Write the lines of R code required to compute the amount of interest to be paid on a loan of \$1000 held for 3 years at an interest rate of 4% compounded annually.

2. Consider the random number generator

$$\begin{aligned} x_j &= 171\ x_{j-1} \bmod 30269 \\ u_j &= x_j/30269, \quad j = 1, 2, \ldots \end{aligned}$$

   where x_0 is an initial seed.

   Write an R function called `myunif()` which takes `x0` and `n` as input arguments and returns a vector of length n which contains the values $u_1, u_2, \ldots, u_n$.

3. Use the rejection sampling method in a function called `rquad()` which should simulate n random variables with the density function

$$f(x) = 3x^2, \quad 0 < x < 1$$

   and 0, otherwise. The function should take `n` as an argument and it should return a vector of length n containing the simulated data values.

4. Consider the *distribution function*

$$F(x) = P(X \leq x) = \begin{cases} 1 - (1+x)^{-\alpha}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

   (a) Write an R function called `ppareto()` which takes a vector `x` and the scalar `alpha` as arguments and which returns a vector containing the probabilities $F(x)$ evaluated at each element of `x`.

   (b) Write an R function called `rpareto()` which takes `n` and `alpha` as arguments and which returns a vector containing n pseudorandom variates from the above distribution.