UNIVERSITY OF THE FREE STATE
DEPARTMENT OF MATHEMATICAL STATISTICS AND ACTUARIAL SCIENCE
STSM 2634

Assignment 1

Full marks: 50

Date: 26 February, 2024

Deadline: 02 March, 2024

FOLLOW THESE INSTRUCTIONS METICULOUSLY, OTHERWISE MARKS WILL BE SUBTRACTED:

• Save the answer file in MS-Word format as **'Assignment1_student number'** as the file name. Your **programming (code), and the output must be included in your answers**. Write the explanation after the code and the output as necessary.

• You have freedom to write the code in your own way. To get full marks, your code should be executable and producing the correct output. PARTIAL GRADING is possible depending on the quality of the code. You have freedom to write the code any way you like. **Smarter coding will get better marks. If the code fails to run, you will get 0 mark.**

• After you are finished with this assignment, please press F12, save the MS-Word document as a .pdf file. After you have saved the pdf document, you submit your **FINAL pdf file (with 'Assignment1_student number' as file name) to the Blackboard assignment submission link**.
This is the document that will be marked.

• You are allowed to use the class notes, or any other help from the internet.

• All computations must be done with the help of suitable R functions. Manual or calculator-based answers will not be accepted.

• 50% deduction in marks for submissions after the deadline.

Q1. (i) Draw **a random sample of real numbers of size 1000**, **with replacement**, from the interval (1, 10) (i.e., excluding the endpoints). The population size (the set from where you draw the sample) should be 449. Hence, choose the range accordingly. Calculate the sample mean. You do not have to print all the random samples. Store the sample mean in a variable, say, x1. Then again draw **another set of 1000 random samples** from the same set, with replacement, calculate the sample mean, and store the sample mean in a different variable, say, x2.

Repeat the process of drawing random samples for 5 times. So, you should have 5 sample mean values stored in variables x1, x2, x3, x4, x5. [2×5 = 10]

(ii) This process of resampling with replacement is the fundamental idea of what we call Bootstrapping. Now calculate the grand mean and the standard deviation of the sample means. [5]

[15]

Q2. (i) Draw **a random sample of integers of size 10 000**, **with replacement**, from the interval [1, 10] (i.e., including the endpoints). You do not have to print all the random samples. Calculate the sample mean. Store the sample mean in a variable, say, y1. Then again draw another set of 10 000 random samples from the same set, with replacement, calculate the sample mean, and store the sample mean value in a different variable, say, y2.

Repeat the process of drawing random samples for 5 times. So, you should have 5 sample mean values stored in variables y1, y2, y3, y4, y5. [2×5 = 10]

(ii) Now calculate the grand mean and the standard deviation of the sample means. Is the standard deviation smaller than that in Q1? Explain why.

[5]

[15]

Q3. (i) Create three vectors X, Y, Z with **all negative integers (any value you like)** and each vector has 3 elements. [6]

(ii) Combine the three vectors to become a 3×3 matrix A where each column represents a vector. [3]

(iii) Then obtain the transpose of the matrix A, denoted by A', and multiply it with the original matrix A to obtain the product $A'A$. Calculate the trace and the determinant of the product $A'A$. Is it possible to have a negative trace? Explain your answer. [6]

[15]

Q4. Create the following matrix with suitable R functions. Manual entries will not be accepted.

	[,1]	[,2]	[,3]
[1,]	1	5	2017
[2,]	3	4	2017
[3,]	5	3	2017
[4,]	7	2	2017
[5,]	9	1	2017

[5]

--- End ---