R Tutorial 7

Instructions:

- Answer all questions.
- Ensure that your findings and results are clearly stated and thoroughly discussed. Please support your arguments using suitable R code with the relevant outputs, interpretations, plots and graphs whenever possible. You should support your argument using appropriate theory that is appropriately referenced.
- The R commands that you use in obtaining your results for all questions must be documented in a R script file. These scripts must be clearly commented. Ensure that any output is clearly stated and interpreted separately from the code as additional comments.
- Include the task name, your name and surname, and your student number in your R script file.
- You MUST label each answer by question number and, where a question has multiple parts, label each part of the question CLEARLY.
- On completion of your assignment, please submit onto RUconnected. If there are any issues uploading onto RUconnected, you may email your submission to: s.izally@ru.ac.za. Please submit your R script file and any other saved data files and plots mentioned in the questions below. Your student number should be included in the name of each file that you submit.
- Each student must complete an individual assignment. You will be assessed based on the quality and/or correctness of the R code, its outputs, and your explanations and interpretations. Acknowledge any help you may have received. Feel free to note any help you may have given to other students in the course.
- This assignment must be submitted by Tuesday, 10 September 2024 by 17:00. Late submissions will be penalized.
- Please note the Rhodes University and the Rhodes University Department of Statistics plagiarism policies.

Questions:

1. Consider Y with probability function

Y	1	2	3	4
$P(Y=y)$	0.4	0.3	0.2	0.1

   (a) Simulate 1 000 values from this distribution using the sample function in R. Store this in a variable called "draws". Use set.seed(1234). **Do not copy the output in your Word document when you submit.**

   (b) Use the values in part (a) to estimate the following in R:

      i. $E(Y)$.
      ii. $E\left(\frac{1}{Y}\right)$.
      iii. $E\left(Y^2\right)$.
      iv. $E\left(Y^2-1\right)$.
      v. $Var(Y)$.

   (c) Use the barplot(table()) command in R to plot the 1 000 draws of X. Label the x-axis as "y", the y-axis as "Frequency" and give the graph the following title: "Barplot of 1 000 discrete draws of Y".

2. A fair coin is repeatedly tossed. Estimate the probability that you will obtain a heads for the third time on the 10^{th}toss. Using R, do the following:

   (a) Use the sample command to simulate a coin toss 10 times and call it coinToss. Use set.seed(1234).

   (b) Add up all the "Heads" obtained from coinToss. Use the sum function in R.

   (c) Evaluate coinToss at its 10^{th} index.

   (d) Count the first 9 tosses where only two heads were obtained. Use the sum function in R.

(e) Use the logical and operator "&&" in R to check that only two heads were obtained in the first 9 tosses **and** a head was obtained on the 10^{th} toss.

(f) Use the mean and replicate function in R to replicate the experiment in (e), 10 000 times. This will estimate the probability of obtaining a heads for the third time on the 10^{th} toss.

3. A fair coin is tossed three times. Use the sample, mean and replicate function for 10 000 replicates and set.seed(1234) to estimate:

   (a) The probability of zero heads obtained.

   (b) The probability of one heads obtained.

   (c) The probability of two heads obtained.

   (d) The probability of three heads obtained.