Statistical Computing and Simulation

Spring 2026

Final Exam, Due June 10/2026

1. There are several ways for acquiring random numbers, e.g., using irrational numbers.

   (a) Find a way to download the decimal places of irrational numbers π, *e*, φ (Golden ratio) or $\sqrt{2}$ in R or other software?

   (b) Check via graphic tools if the first 10,000 digits after the decimal point of π, *e*, φ, or $\sqrt{2}$ violate the assumption of random numbers.

   (c) Apply at least one formal test to verify if the first 10,000 digits after the decimal point of π, *e*, φ, or $\sqrt{2}$ satisfy the assumption of uniform(0,1) and independence.

(Note: If the last two digits of your student ID number have a remainder of 0, 1, 2, or 3 when divided by 4, then select the corresponding irrational numbers π, *e*, φ, or $\sqrt{2}$.)

2. We can use "arima.sim" in R to generate random numbers from ARIMA models, as well as applying Cholesky decomposition.

   (a) Generate 500 random numbers from AR(2) with parameter values $(\phi_1, \phi_2) = (\theta, \theta)$ for θ = 0, 0.1, and 0.2, given that μ = 0 and σ = 1. Calculate the correlation coefficients between x_i vs. x_{i+1} and x_i vs. x_{i+2} .

   (b) Generate 500 random vectors from $\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & a & a \\ a & 1 & a \\ a & a & 1 \end{pmatrix}\right)$, with $a = \frac{\theta}{1-\theta}$ and θ = 0, 0.1, and 0.2. Calculate the correlation coefficients between x_i vs. x_{i+1} and x_i vs. x_{i+2}. Compare the results of (a) and (b) and comment on what you found from the simulation.

3. (a) To compare teaching, twenty schoolchildren were divided into two groups: ten taught by conventional methods and ten taught by an entirely new approach. The following are the test results:

Conventional	65	79	90	75	61	85	98	80	97	75
New	90	98	73	79	84	81	98	90	83	88

   Are the two teaching methods equivalent in result? You need to use permutation test, (parametric and non-parametric) bootstrap, and parametric test, and then compare their differences in testing.

(b) Write a small program to perform the "Permutation test" and test your result on the correlation of DDT vs. eggshell thickness in class, and the following data:

X	585	1002	472	493	408	690	291
Y	0.1	0.2	0.5	1.0	1.5	2.0	3.0

Check your answer with other correlation tests, such as regular Pearson and Spearman correlation coefficients.

(c) Using simulation to construct critical values of the Mann-Whitney-Wilcoxon test in the case that $2 \leq n_1, n_2 \leq 10$, where n_1 and n_2 are the number of observations in two populations. The number of replications shall be at least 10,000.

(Note: If the last two digits of your student ID number have a remainder of 0, 1, or 2 when divided by 3, then select the problem (a), (b), or (c).)

4. Experiment with as many variance reduction techniques as you can think of to apply the problem of evaluating $P(X > 2.5)$ for $X \sim$ Cauchy.

5. (a) Let x be 150 equally spaced points on $[0,2\pi]$ and let $y_i = \sin x_i + \varepsilon_i$ with $\varepsilon_i \sim N(0,0.04)$. Apply at least 3 linear smoothers and compare the differences, with respect to mean squares error (i.e., bias^2 and variance) from 1,000 simulation runs.

(b) Let x be 150 equally spaced points on $[0,2\pi]$ and let $y_i = \cos x_i + \varepsilon_i$ with $\varepsilon_i \sim N(0,0.04)$. Apply at least 3 linear smoothers and compare the differences, with respect to mean squares error (i.e., bias^2 and variance) from 1,000 simulation runs.

(Note: If the last digit of your student ID number is even or odd number, then select the problem (a) or (b).)