STAT 7650: Homework 5
(Due: Thursday, 04/10/2025)

Note: Show all your work for the necessary steps to receive full credit.

Please turn in the HW on paper, hand-written and/or typed. For computational problems, return only the relevant parts of the output with comments/annotations. No need to turn in your code file or send it via email or print long lists of generated samples, but you are welcome to email it to me. However, from the code you are using to answer the problems, turn in the relevant code and output with explanation and justification, and the figures (if requested), preferably printed from the output.

Please disclose any us of AI in your solutions. Regardless of you use it or not, make sure you submit your own work, not copy from other source(s). Any suspicion of AI use will result in automatic 0 or substantial point loss in any question.

Q1. Suppose that $X \sim \text{Gamma}(3, 1)$ and $Y \sim \text{Gamma}(7, 1)$ are independent random variables.

(a) Letting $U = X$ and $V = X + Y$, find the joint distribution of (U, V) using Jacobian.

(b) Find, numerically, the 80th percentile of the conditional distribution of X, given $X + Y = 7$ using your result in part (a). (Hint: Use the R functions `integrate` and `uniroot`.)

(c) Derive the conditional distribution of X, given $X + Y = 7$, in part (a) above. Use your formula for the conditional distribution to find the exact 80th percentile (you may use a built-in quantile function in R here) and compare to your numerical approximation in part(a).

(d) Redo the calculation in part (a) using an AI tool. Compare the two approaches and discuss their efficiency and accuracy.

Q2. (Based on Problem 4.13 in Lange's Numerical Analysis for Statisticians.) The von Mises distribution is commonly used as a model for data that takes value on a circle. The density function for the von Mises distribution is

$$f(x) = \frac{\exp(\kappa \cos(x - \mu))}{2\pi I_0(\kappa)}, \quad -\pi \le x \le \pi, \tag{1}$$

where μ is a location parameter, $\kappa > 0$ is a concentration parameter, and the normalizing constant $I_0(\kappa)$ is the modified Bessel function

$$I_0(\kappa) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(\kappa \cos y)\, dy. \tag{2}$$

(a) Use the Laplace approximation technique to show that

$$I_0(\kappa) \approx (2\pi\kappa)^{-\frac{1}{2}} e^{\kappa}, \qquad \text{as } \kappa \to \infty. \tag{3}$$

(b) Argue that the von Mises distribution is approximately normal for large κ. (Hint: Use a Taylor approximation of $\cos z$ at $z = 0$.)

(c) Simulate samples from the von Mises distribution for $\kappa = 2.5$ with the Accept-Reject Method using a $N(0, \sigma^2)$ distribution as the envelope (or majorant) function. What would be the optimal σ^2 value to maximize the acceptance rate? **Hint:** Note the difference in the support of the current version of the von Mises distribution than the one in the slides.

(d) Leveraging an AI tool, simulate samples from the von Mises distribution for varying κ values. Visualize these samples on a circular plot to demonstrate the effect of the concentration parameter on the distribution's shape.

Q3.

(a) For any continuous distribution function F (not necessarily strictly increasing), show that if $U \sim \text{Uniform}(0,1)$, then $X = F^{-1}(U) = \inf\{x : F(x) \geq U\}$ has cumulative distribution function equal to F.

(b) Write the details or steps of the cdf inverse technique to generate samples from the symmetric triangle distribution in the class slides by explicitly using the inverse of the cdf (i.e. without utilizing the symmetry in the pdf).

(c) Utilizing AI-based mathematical tools, generate samples from a symmetric triangle distribution. Detail the steps of the cdf inverse technique explicitly, and compare the results with samples generated in part (b).

Q4. **Sampling from Multivariate Normal Distribution (from class slides):** Consider a p-dimensional normal distribution with mean vector $\boldsymbol{\mu}$ and a $p \times p$ variance-covariance matrix $\boldsymbol{\Sigma}$. We have seen techniques (e.g. Box-Muller) for sampling a vector $\mathbf{Z} = (Z_1, \ldots, Z_p)^\top$ of independent normal random variables (RVs), and now we seek a method to incorporate the dependence described by $\boldsymbol{\Sigma}$. Let $\boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}^\top$ represent the Cholesky decomposition of $\boldsymbol{\Sigma}$. Show that the vector $\mathbf{X} = \boldsymbol{\mu} + \mathbf{L}\mathbf{Z}$ follows the desired p-dimensional normal distribution.

Q5. We aim to sample from a one-dimensional random variable X that has a strictly increasing distribution function F. Unfortunately, an explicit formula for the inverse function F^{-1} is not accessible, but you can evaluate f and F at any given value.

(a) Explain how to utilize Newton's method in conjunction with the inverse cumulative distribution function (cdf) method to simulate random variables distributed according to F. Make sure to write the corresponding updating equation in the Newton's method.

(b) Apply the described method for simulating variables from a standard normal distribution, $N(0,1)$. Assess the efficiency of this sampling approach in comparison to the Box-Muller method and the `rnorm` function (i.e., tabulate the time required to generate, say 100000 of the draws, for each method to compare).

(c) Repeat part (b) utilizing an AI tool.

$$\begin{aligned}
&\max_{\{X_{ij}\}\geq 0} \ \mathbb{E}_{\theta\sim\pi}\left[\sum_{j\neq s} \left(P_{sj}(X_{sj},\theta) - c_{sj}(\theta)\right) X_{sj} \ - \ \sum_{i\neq s} \left(P_{is}(X_{is},\theta) + t_{is}(\theta)\right) X_{is}\right] \\
&\text{s.t.} \quad \sum_{j\neq s} X_{sj} \leq S_s, \qquad \sum_{i\neq s} X_{is} \leq D_s.
\end{aligned} \tag{4}$$

where X_{ij} is the flow from exporter i to importer j; s denotes the supported actor (e.g., the United States); $P_{ij}(\cdot, \theta)$ is the price received or paid under adversarial state θ; $c_{sj}(\theta)$ and $t_{is}(\theta)$ are uncertain costs and tariffs; S_s and D_s are the supply and demand limits, respectively; and π is the supported actor's ARA belief over θ.