1. (10 points) (6 + 4)

Researchers notice that a mutation in a gene predisposes individuals to a kind of radiation-induced cancer. The researchers theorize that the gene is involved in repairing damage from radiation, and that the mutation disables the gene. To explore their theory, the researchers obtain cells growing in a laboratory that have the mutation. They take eight different clumps of the cells, and randomize the clumps to treatment with radiation or no radiation (four in each group). They then examine a marker of damage from radiation in each cell in each clump, recording whether or not there appears to be damage. The researchers run the same experiment in clumps of cells that do not have the mutation. They explain that the cells that do not have the mutation are a "control". The researchers ask you to analyze the results.

(a) Propose a "reasonable" model to analyze the data.

(b) Propose how you plan to conclude whether mutation plays a role in repairing radiation damage.

2. (14 points) (7 + 7) Suppose that $X_1, \ldots, X_{2n}$ are i.i.d. $U[0, 1]$. Let $Y_i = X_{2i-1} + X_{2i}$ for $1 \leq i \leq n$.

(a) Find the limiting distribution of Y_1.

(b) Find the limiting distribution of $\sqrt{n}(2 - Y_{(n)})$ as $n \to \infty$.

3. (20 points) (5 + 5 + 10) Suppose that $(X_{1i}, X_{2i}) \overset{i.i.d.}{\sim} N_2(\theta, I_2)$ for $1 \leq i \leq n$, where the parameter space is restricted to $\Theta := \{\theta = (\theta_1, \theta_2) : \theta_1, \theta_2 \geq 0\}$. Consider the following hypothesis testing problem:

$$H_0 : \theta = (0, 0) \qquad \text{versus} \qquad H_1 : \theta \in \Theta \backslash \{(0, 0)\}.$$

(a) Find the MLE of θ (when $\theta \in \Theta$).

(b) Find an expression for the likelihood ratio statistic $\Lambda_n \in (0, 1]$ in this case.

(c) Find the asymptotic distribution of $-2 \log \Lambda_n$, under H_0 [Hint: You may want to consider the cases where $(\overline{X}_1, \overline{X}_2)$ belongs to each of the four quadrants separately.]

4. (10 points) (3 + 3 + 2 + 2)

Suppose we generate $U \sim [0, 3]$. Let V denote the value of the integer nearest to U (so V takes values in $\{0, 1, 2, 3\}$). Let X denote the rounding error i.e. the absolute distance between U and V.

(a) What is the distribution of V?

(b) What is the distribution of the X.

(c) Are X and V independent?

(d) Are X and U independent?

5. (10 points) (5 + 5)

Consider the following Bayesian model

$$Y_1, \ldots, Y_n \overset{\text{iid}}{\sim} \text{Uniform}([0, \theta]) \text{ and } \theta \sim \text{Pareto}(\beta, \lambda,)$$

where the pdf of the Pareto distribution is given by

$$\pi(\theta; \beta, \lambda) = \frac{\beta \lambda^\beta}{\theta^{(\beta+1)}}, \quad \theta > \lambda, \quad \beta, \lambda > 0.$$

Moreover, for this exercise you may assume $\beta > 1$.

(a) Use the Bayes formula to derive the posterior density of θ as explicitly as possible.

(b) Compute the prior and posterior means of θ.

6. (14 points) $(7+7)$

Suppose $(X_1, \ldots, X_n)$ are i.i.d. from a Normal distribution with $\mathbb{E}X_i = \mathrm{Var}(X_i) = \theta$, where $\theta > 0$ is unknown.

(a) Find the MLE for θ explicitly.

(b) Find the asymptotic distribution of your MLE.

7. (10 points) $(4+6)$

Suppose X_1, X_2 are i.i.d. $N(0,1)$.

(a) Find the joint distribution of $X_1 + X_2$ and $X_1 - X_2$.

(b) Show that $2X_1X_2$ has the same distribution as $X_1^2 - X_2^2$.

8. (10 points) $(2+3+5)$

For every $n \geq 1$, let A_n be an $n \times n$ symmetric matrix with non negative entries. Let $R_n(i) := \sum_{j=1}^n A_n(i,j)$ denote the ith row/column sum of A_n. Assume that

$$\lim_{n\to\infty} |R_n(i) - 1| = 0.$$

Let $\lambda_n \geq 0$ denote an eigenvalue with the largest absolute value, and let $\mathbf{x} := (x_1, \ldots, x_n)$ denote its corresponding eigenvector.

(a) Show that $\frac{1}{n}\sum_{i,j=1}^n A_n(i,j) \to 1$.

(b) Show that $\lambda_n |x_i| \leq \max_{1\leq j\leq n} |x_j| R_n(i)$.

(c) Using parts (a) and (b), show that $\lambda_n \to 1$.