1. Let $\{X_n\}_{n\geq 0}$ and $\{Y_n\}_{n\geq 0}$ be two independent Markov chains on the finite state spaces S and T, respectively, with transition probability matrices $\mathbf{P} = ((p_{ij}))$ and $\mathbf{Q} = ((q_{ij}))$, respectively. Define a new stochastic process $Z_n = (X_n, Y_n), n \geq 0$.

   (a) Show that $\{Z_n\}_{n\geq 0}$ is also a Markov chain on the state space $S \times T$ and write down its transition probability matrix.

   (b) If $\{X_n\}_{n\geq 0}$ and $\{Y_n\}_{n\geq 0}$ are both irreducible, then will $\{Z_n\}_{n\geq 0}$ always be irreducible? Justify your answer.

[5+10 = 15]

2. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables having a common distribution function F with $\mathbb{E}(X_1) = \mu$. Assume that $F(\cdot)$ is continuous at μ. Define the empirical cumulative distribution function

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_i \leq t), \quad \text{for all } t \in \mathbb{R}.$$

For $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, show that $F_n(\bar{X}_n)$ is a consistent estimator of $F(\mu) = \mathbb{P}(X_1 \leq \mu)$.

[15]

3. Let X_1, X_2, X_3 be independent and identically distributed positive random variables. Define $U_i = X_i/(X_1 + X_2 + X_3)$ for $i = 1, 2, 3$.

   (a) Give a choice of (a_1, a_2, a_3) so that $\text{Var}(a_1U_1 + a_2U_2 + a_3U_3) = 0$.

   (b) Find the correlation matrix of (U_1, U_2, U_3).

   (c) Find a_1, a_2 and a_3 satisfying $a_1^2 + a_2^2 + a_3^2 = 1$ such that $\text{Var}(a_1U_1 + a_2U_2 + a_3U_3)$ is maximum.

[3+6+6=15]

4. Let $\mathbf{X} = (X_1, \ldots, X_4)^\top \sim N_4(\mathbf{0}, \mathbf{I}_4)$, where $\mathbf{I}_4$ is the 4×4 identity matrix. Define

$$Q_1 = \frac{3X_1^2 + X_2^2 + X_3^2 + X_4^2 + 2X_2X_3 + 2X_2X_4 + 2X_3X_4}{3} \quad \text{and}$$
$$Q_2 = \frac{2X_2^2 + 2X_3^2 + 2X_4^2 - 2X_2X_3 - 2X_2X_4 - 2X_3X_4}{3}.$$

(a) Find the distributions of Q_1 and Q_2.

(b) Show that Q_1 and Q_2 are independent.

[10+5 = 15]

5. Suppose that we have a coin with an unknown probability of head p. To estimate $\theta = p^2$, we can adopt two strategies (S1) and (S2):

(S1) Flip the coin $2n$ times. Let S_{2n} be the number of heads. Consider $U_n = \left(\frac{S_{2n}}{2n}\right)^2$ as an estimator of θ.

(S2) Flip the coin twice and define $Y_i = 1$ if a pair of heads is obtained and $Y_i = 0$ otherwise. Repeat the experiment n times and use $V_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ as an estimator of θ.

(a) Show that both U_n and V_n are consistent estimators of θ.

(b) Find asymptotic distributions of $\sqrt{n}(U_n - \theta)$ and $\sqrt{n}(V_n - \theta)$.

(c) Which of these two estimators would you prefer? Justify your answer.

[5+7+3 = 15]

6. Suppose that M observations are generated from a bivariate normal distribution $N_2(\mu, \mu, \sigma_1^2, \sigma_2^2, \rho)$, where all parameters are unknown. If we only know the numbers of observations in the first quadrant and the third quadrant (M_1 and M_3, say), construct an exact test for $H_0 : \mu = 0$ against $H_1 : \mu > 0$. Justify your answer.

[15]

7. Let X_1, X_2 and X_3 be independent and identically distributed continuous random variables with a strictly increasing distribution function F. For any $x \in \mathbb{R}$, define

$$\psi(x) = \mathbb{P}\left(\min_{i=1,2,3} X_i \leq x \leq \max_{i=1,2,3} X_i\right).$$

Show that $\psi(x)$ is maximum when x is the median of the distribution.

[15]

8. An experiment is to be planned for comparing 5 drugs, and n patients of similar health conditions are available as experimental subjects. The drugs are to be allocated to the patients, one drug per patient, and observations are to be taken to get a total of n observations. Suppose that τ_i is the treatment effect of the i-th $(i = 1, 2, \ldots, 5)$ drug under the usual linear model for this setup, and we assume independent and homoscedastic errors.

   (a) If $n = 35$, suggest a suitable design for this experiment so that the average variance of the best linear unbiased estimators of the treatment contrasts $\tau_i - \tau_j$ will be minimized, the average being taken over all $i \neq j$ $(i, j = 1, 2, \ldots, 5)$.

   (b) If $n = 36$, what will be the allocation numbers for the treatments if you want to minimize the average variance of the best linear unbiased estimators of the treatment contrasts $\tau_i - \tau_j$, the average being taken over all $i \neq j$ $(i, j = 1, 2, \ldots, 5)$?

   (c) If $n = 36$ and the first drug is a control drug whereas the other four are test drugs, what will be the allocation numbers for the treatments if you want to minimize the average variance of the best linear unbiased estimators of four treatment contrasts $\tau_1 - \tau_j$, the average being taken over all $j = 1, 2, 3, 4$?

[5+5+5=15]