1 Independence

For each one of the statements below, give either a proof or a counterexample showing that the statement is not always true.

1. If events A and B are independent, then the events A and B^c (complement) are also independent. (2 marks)
2. Let A, B, and C be events associated with a common probabilistic model, and assume that $0 < P(C) < 1$. Suppose that A and B are conditionally independent given C. Then, A and B are conditionally independent given C^c. (2 marks)
3. Let X and Y be independent random variables. Then, $\text{var}(X + Y) \geq \text{var}(X)$. (1 mark)

2 Duplicates

A box has n balls numbered from 1 to n. Suppose you keep picking a ball randomly each time and put it back in the box before the next pick.

1. Let X be the random variable denoting the first time at which you have seen a ball twice. Find the PMF of X. (2 marks)
2. Let T_i be the random variable corresponding to the time taken for seeing a new ball, after you have seen i different balls. Find the PMF of T_i. (2 marks)
3. Let T be the random variable corresponding to the first time at which you have encountered all the n balls. Find $\mathbb{E}T$. (1 mark)

3 Randomized Coloring

Given a (undirected) graph $G = (V, E)$, and a 3-color assignment $a : V \to \{R, G, B\}$ is an assignment of colors R, G, B to the vertices of the graph. Given an assignment a, the set of monochromatic edges $E(a) = \{(u, v) \in E : a(u) = a(v)\}$, is the set of edges that has same colors for endpoints. Let a be randomly chosen, ie for every $v \in V$, it is chosen to be R, G, B uniformly and independent of the other vertices.

1. For any edge $e \in E$, let X_e be the random variable which is 1 when e is monochromatic and 0 otherwise. Show that the set of random variables $\{X_e\}_{e \in E}$ are pairwise independent. Show that they are not independent. (1 mark)
2. Let Y be the random variable corresponding to the number of non-monochromatic edges. That is $Y = |E \setminus E(a)|$. Find $\mathbb{E}[Y]$. (1 mark)
3. Show that there cannot be a graph for which all 3-color assignments make $< 2|E|/3$ edges non-monochromatic. That is for any graph G, there exists an assignment $a : V \to \{R, G, B\}$ such that the number of non-monochromatic edges is at least $2|E|/3$. (1 mark)
4. Show that: $P(Y \geq |E|/2) \geq 1/3$. (1 mark)
5. Devise a method (which by obtaining multiple independent copies of Y by randomly choosing a's independently) that can find an assignment for which the number of non-monochromatic edges is at least $|E|/2$ with probability at least $99/100$. (1 mark)

4 Continuous Random Variables

The random variable X is exponential with parameter 1. Given the value x of X, the random variable Y is exponential with parameter equal to x (and mean $1/x$).

1. Find the joint PDF of X and Y. (1 mark)
2. Find the marginal PDF of Y. (1 mark)
3. Find the conditional PDF of X, given that $Y = 2$. (1 mark)
4. Find the conditional expectation of X, given that $Y = 2$. (1 mark)
5. Find the conditional PDF of Y, given that $X = 2$ and $Y \geq 3$. (1 mark)

5 Signal Classification

Consider the communication of binary-valued messages over some transmission medium. Specifically, any message transmitted between locations is one of two possible symbols, 0 or 1. Each symbol occurs with equal probability. It is also known that any numerical value sent over this wire is subject to distortion; namely, if the value X is transmitted, the value Y received at the other end is described by $Y = X + N$ where the random variable N represents additive noise that is independent of X. The noise N is normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 4$.

1. Suppose the transmitter encodes the symbol 0 with the value $X = -2$ and the symbol 1 with the value $X = 2$. At the other end, the received message is decoded according to the following rules:

   (a) If $Y \geq 0$, then conclude the symbol 1 was sent.

   (b) If $Y < 0$ then conclude the symbol 0 was sent.

   Determine the probability of error for this encoding/decoding scheme. Answer can be in terms of Gaussian integrals (CDF of Normal Distribution). (2 marks)

2. In an effort to reduce the probability of error, the following modifications are made. The transmitter encodes the symbols with a repeated scheme. The symbol 0 is encoded with the vector $X = [-2, -2, -2]^T$ and the symbol 1 is encoded with the vector $X = [2, 2, 2]^T$. The vector $Y = [Y_1, Y_2, Y_3]^T$ received at the other end is described by $Y = X + N$. The vector $N = [N_1, N_2, N_3]^T$ represents the noise vector where each N_i is a random variable assumed to be normally distributed with mean $\mu = 0$ and variance $\sigma^2 = 4$. Assume each N_i is independent of each other and independent of the X_i's. Each component value of Y is decoded with the same rule as in part (1.). The receiver then uses a majority rule to determine which symbol was sent. The receiver's decoding rules are:

   (a) If 2 or more components of Y are greater than 0, then conclude the symbol 1 was sent.

   (b) If 2 or more components of Y are less than 0, then conclude the symbol 0 was sent.

   Determine the probability of error for this modified encoding/decoding scheme. (3 marks)