Homework Assignment #1

Assigned: Saturday 01/24/2015; Due: Saturday 02/07/2015 via Oncourse.
(total: 95 points)

Problem 1 (5 points) Let $(\Omega, \mathcal{F}, P)$ be a probability space. Using only the set operations and axioms of probability, show that for any two sets, not necessarily disjoint, $A \subseteq \Omega$ and $B \subseteq \Omega$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Problem 2 (5 points) Prove the following expression or provide a counterexample if it does not hold

$$P(A) = P(A|B) + P(A|B^c)$$

Problem 3 (15 points) Two players perform a series of coin tosses. Player one wins a toss if the coin turns heads and player two wins a toss if it turns tails. The game is played until one player wins n times. However, the game is interrupted when player one had m wins and player two had l wins, where $0 \leq m < n$ and $0 \leq l < n$.

a) (5 points) Assuming $n = 8$, $m = 4$, and $l = 6$, what is the probability that player one would win the game if the game was to be continued later.
b) (10 points) Derive the general expression or write an algorithm that player one will win the game if the game is to be continued later. Your expression should be a function of m, l, and n. If you are providing an algorithm, implement it and submit your code along with your pseudo-code. You may *not* simulate the game as your solution.

Problem 4 (10 points) Let $\Omega_X = \{a, b, c\}$ and $p_X(a) = 0.1$, $p_X(b) = 0.2$, and $p_X(c) = 0.7$. Let

$$f(x) = \begin{cases} 10 & x = a \\ 5 & x = b \\ 10/7 & x = c \end{cases}$$

a) (3 points) What is $E[f(x)]$?
b) (3 points) What is $E[1/p_X(x)]$?
c) (4 points) For an arbitrary pmf $p_X(x)$, what is $E[1/p_X(x)]$?

Problem 5 (15 points) A biased four-sided die is rolled and the down face is a random variable N described by the following pmf:

$$p_N(n) = \begin{cases} n/10 & n = 1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

Given the random variable N a biased coin is flipped and the random variable X is 1 or zero according to whether the coin shows heads or tails. The conditional pmf is

$$p_{X|N}(x|n) = \left(\frac{n+1}{2n}\right)^{x}\left(1-\frac{n+1}{2n}\right)^{1-x}$$

where $x \in \{0,1\}$.

a) (5 points) Find the expectation $E[N]$ and variance $V[N]$ of N
b) (5 points) Find the conditional pmf $p_{N|X}(n|x)$
c) (5 points) Find the conditional expectation $E[N|X=1]$, i.e. the expectation with respect to the conditional pmf $p_{N|X}(n|1)$

Problem 6 (10 points) Let X, Y, and Z be discrete random variables defined as functions on the same probability space $(\Omega, \mathcal{F}, P)$. Prove or disprove the following expression

$$P_{X|Y}(X=x|Y=y) = \sum_{z\in\Omega_Z} P_{X|YZ}(X=x|Y=y,Z=z)P_{Z|Y}(Z=z|Y=y)$$

Problem 7 (15 points) Consider the following game: A player is shown four closed doors and informed that the prize money is behind one of them (there is an equal chance that the money is behind each door). The player is asked to step in front of one of the doors, but then the following twist happens: Instead of revealing whether money is behind the door, the host chooses to open one of the remaining doors without the prize (the host picks the doors with equal probability). The player is then given an opportunity to step in front of some other door (he/she can choose to stay at the original door as well).

a) (9 points) Define the probability space for this game. In particular, what is the sample space (hint: enumerate all outcomes of the game), what is the probability mass function, and consequently, what is the probability of winning the prize if the player stays at the original door.
b) (6 points) Use random variables to solve this problem. What is the probability of winning the prize if the player stays at the original door and what is the probability of winning the prize if the player changes doors? Derive all expressions (no guesses) and clearly explain all random variables used in the derivation.

Problem 8 (10 points) The time (in hours) necessary to find and fix an electrical problem in a certain institution is a random variable, say X, whose density is given by

$$p_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

If the cost of the breakdown of duration x is x^3, what is the expected cost of an electrical breakdown?

Problem 9 (10 points) Let X be a continuous random variable with a cumulative distribution function $F_X(t)$. The mediam of of a random variable is defined as a value of t for which

$$F(t) = \frac{1}{2}$$

Find the median of the random variables with the following density functions

a) (5 points) $p_X(x) = e^{-x}$, $x > 0$.
b) (5 points) $p_X(x) = 1$, $0 \leq x \leq 1$.

Extra Problem

Extra Problem (20 points) High dimensional spaces.

a) (10 points) Show that in a high dimensional space, most of the volume of a cube is concentrated in corners, which themselves become very long "spikes." Hints: compute the ratio of the volume of a hypersphere of radius a to the volume of a hypercube of side $2a$ and also the ratio of the distance from the center of the hypercube to one of the corners divided by the perpendicular distance to one of the edges.

b) (10 points) Show that for points which are uniformly distributed inside a sphere in d dimensions where d is large, almost all of the points are concentrated in a thin shell close to the surface. Hints: compute the fraction of the volume of the sphere which lies at values of the radius between $a - \varepsilon$ and $0 < \varepsilon < a$; Evaluate this fraction for $\varepsilon = 0.01a$ and also for $\varepsilon = 0.5a$ for $d \in \{2, 3, 10, 100\}$.

Homework policies:

Your assignment must be typed; for example, in Latex, Microsoft Word, Lyx, etc. Images may be scanned and inserted into the document if it is too complicated to draw them properly. Submit a single pdf document or if you are attaching your code submit your code together with the typed (single) document as one .zip file.

All code (if applicable) should be turned in when you submit your assignment. Use Matlab, Python, R, or Java.

Policy for late submission assignments: Unless there are legitimate circumstances, late assignments will be accepted up to 5 days after the due date and graded using the following rule:

on time:	your score $\times$ 1
1 day late:	your score $\times$ 0.9
2 days late:	your score $\times$ 0.7
3 days late:	your score $\times$ 0.5
4 days late:	your score $\times$ 0.3
5 days late:	your score $\times$ 0.1

For example, this means that if you submit 3 days late and get 80 points for your answers, your total number of points will be $80 \times 0.5 = 40$ points.

All assignments are individual, except when collaboration is explicitly allowed. All the sources used for problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc. Academic honesty is taken seriously; for detailed information see Indiana University Code of Student Rights, Responsibilities, and Conduct.

Good luck!