ECE 361 Probability for Engineers (Fall, 2016)
Homework 4

Assigned:	Thursday October 13, 2016	
Due:	Thursday October 20, 2016	**(at the beginning of class 11am)**
Returned:	Monday October 24, 2016	**(in recitation)**

Your full name ______________________________

Your Drexel student ID ______________________________

ECE 361 homework policies:

1. There will be eight homework assignments, roughly one per week, but check the Course Calendar for exact dates.
2. Each homework will count equally towards your overall homework grade, although the lowest homework will be dropped.
3. Homework must be done INDIVIDUALLY. Although you are free to discuss course content in general with your classmates and peers, you are expected to NOT discuss particulars about homework problems with them or anyone else.
4. Similarly, although you are encouraged to read external sources (online or offline), it is expected that your homework handed in reflects YOUR work, and not work that you found elsewhere.
5. Homework will always be due at the BEGINNING of lecture. One of the Teaching Assistants will be in the lecture hall at the beginning of class to collect the assignments. At 11:10am (ten minutes into the start of lecture) the TA will leave the lecture with the assignments.
6. No submissions will be accepted after that time. No exceptions.
7. Homework will be returned to you during recitations.
8. Homework assignments should have your full NAME, your STUDENT ID, the ASSIGNMENT NUMBER, should be STAPLED, and should be written CLEARLY, with your final ANSWER clearly indicated, and all supporting WORK provided. Sloppy, unclear, or illegible work will not be graded.

Please answer the following questions:

1. (3 points) An undirected simple graph G of *order* n and *size* m has n vertices $V = [n]$ and m edges $E = (e_1, \ldots, e_m)$, with each e_i a distinct unordered pair of vertices. A graph is often denoted as $G = (V, E)$. A graph on n vertices may have anywhere from $m = 0$ to $m = M_n \equiv \binom{n}{2}$ edges. A graph with $m = 0$ is called an empty graph, while a graph with $m = M_n$ edges is a complete graph. Given a graph, we define the *neighbors* of a vertex $v \in [n]$ as those vertices $N(v)$ connected by an edge to v. Formally, $N(v) = \{u \in [n] : \{u, v\} \in E\}$ is the set of neighbors of node v, and every node $v \in [n]$ has a neighbor set, possibly empty, meaning there are no edges incident at that vertex. Finally, the last graph-theoretic notion we require is *degree*: the degree of a vertex is the number of neighbors, i.e., $d(v) = |N(v)|$, and again, every node $v \in [n]$ has a degree. Note $d(v) \in \{0, \ldots, n-1\}$, where $d(v) = 0$ means vertex v is *isolated* (has no neighbors) and $d(v) = n - 1$ means vertex v is connected to every possible other vertex.

   A *random* graph with parameters (n, p), with $n \in \mathbb{N}$ and $p \in (0, 1)$, is defined in terms of a collection of independent and identically distributed Bernoulli random variables $X = (X_1, \ldots, X_{M_n})$, with $X_e \sim \mathrm{Ber}(p)$ for $e \in M_n$. We construct the random graph from the vector of RVs X as follows:

$$X_e = \begin{cases} 1, & \text{graph contains edge } e \\ 0, & \text{graph does NOT contain edge } e \end{cases} \tag{1}$$

   Thus, we effectively flip a (biased) coin once for each *possible* edge in the graph, i.e., for each unordered pair of vertices, and add an edge connecting those two vertices when that coin flip shows a head, which happens with probability p. Example, for $n = 4$ there are $M_4 = \binom{4}{2} = 6$ possible edges, and each such edge is added independently with probability p.

   Consider a random graph with order $n = 100$ nodes constructed as above with edge probability $p = 1/3$. Please answer the following questions:

   - Consider vertex 1. Let Y_1 be the RV giving the *degree* of vertex 1. What type of RV is Y_1?
   - Compute $\mathbb{E}[Y_1]$ and $\mathrm{Var}(Y_1)$.
   - Find $\mathbb{P}(Y_1 > 40)$. *Hint: use a computer or a table.*

2. (3 points) Recall the previous problem. Consider a random graph with $n \geq 2 \in \mathbb{N}$ nodes and edge probability $p \in (0, 1)$, and let Y_1 be the RV giving the degree of vertex 1. Let A be the event that $\{1, 2\}$ is an edge in the graph. Give the conditional distribution of Y_1 given A.

3. (3 points) Let (X_1, X_2) be two independent and identically distributed geometric random variables, each with parameter $p \in (0, 1)$. That is, $X_1 \sim \mathrm{Geo}(p)$ and $X_2 \sim \mathrm{Geo}(p)$. Let $A_n = \{X_1 + X_2 = n\}$ for $n \in \mathbb{N}$ be their total. You are told A_n is true and asked to guess the most likely value for X_1, i.e., given (n, p) find the value i^* as a function of n, p such that $\mathbb{P}(X_1 = i^* | X_1 + X_2 = n) > \mathbb{P}(X_1 = i | X_1 + X_2 = n)$ for all $i \neq i^*$. *Hint: first express* $\mathbb{P}(X_1 + X_2 = n)$ *as a simple expression of* n *and* p*. Verify that adding up this expression over all* $n \in \{2, 3, 4 \ldots\}$ *yields one. Next, use Bayes's rule to express* $\mathbb{P}(X_1 = i | X_1 + X_2 = n)$ *in terms of* $\mathbb{P}(X_1 + X_2 = n | X_1 = i)$, $\mathbb{P}(X_1 = i)$, *and* $\mathbb{P}(X_1 + X_2 = n)$.

4. (3 points) Consider a collection of N fair M-sided dice, for $N \in \mathbb{N}$ and $M \geq 2$. We roll all the dice simultaneously, and put aside all those that show face M. We then re-roll all the remaining dice, and again put aside all those that show face M. We repeat this procedure until all dice show face M. Define the RV X_N with support $\mathbb{N}$ as the number of times we rolled dice (not the number of dice rolled) starting with N dice. Find $\mathbb{E}[X_5]$ for $M = 6$. In words, find the average number of rolls required until five fair six-sided dice each show six, under the above procedure. You may give either an analytical expression or simulate this process on a computer and report simulation results. If you use a computer, use at least 1000 trials and average the results.

5. (3 points) A soft drink manufacturer runs a promotion where each bottle cap shows a random number between 1 and $N = 10$, and a prize is offered to any consumer holding a collection of N bottle caps with each of the N possible numbers included in the set. Let T_N be the RV denoting the number of soft drinks purchased until the complete set is obtained. Give $\mathbb{E}[T_N]$. *Hint: think of opening one can each time unit and thus T_N is the random time at which you first have a complete set. Define N stages, labeled $1, \ldots, N$, where in stage n we have obtained $n-1$ distinct bottle caps and are looking for the nth distinct bottle cap. Let $(X_1, \ldots, X_N)$ be RVs where X_n is the number of bottles opened in stage n. Observe $T = X_1 + \cdots + X_N$. Use linearity of expectation. Read about the Nth harmonic number and the Euler-Mascheroni constant.*
6. (3 points) Consider N distinct N-sided dice. The N dice are rolled simultaneously and we let the RV $X_N \in \{1, \ldots, N\}$ denote the number of distinct values shown on the N dice. Find $\mathbb{P}(X_N = 1)$, $\mathbb{P}(X_N = 2)$, $\mathbb{P}(X_N = N-1)$, and $\mathbb{P}(X_N = N)$ for generic $N \geq 2$. *Hint: try $N = 2$, $N = 3$, and $N = 4$ to build intuition for the general case. I advise you number the dice from 1 to N so they are distinguishable, and observe there are N^N possible outcomes when they are all rolled. The event $\{X_N = 1\}$, for example, has a probability given by the fraction of the N^N outcomes with 1 distinct value shown. Use the concepts of combinations, permutations, and multinomial coefficients studied earlier. Verify your proposed formulas for general N with calculations by hand for small N such as $N \in \{2, 3, 4\}$.*
7. (3 points) Recall the previous problem. Find $\mathbb{P}(X = n)$ for each $n \in [N]$ for $N = 6$. *Hint: you may use a computer for this problem.*