Due: Thursday, August 25th, 2016
Problem to be turned in: 3

Due: Thursday, August 25th, 2016
Problem to be turned in: 7

1. Consider an experiment described by a Poisson($\frac{1}{2}$) distribution and answer the following questions.
   (a) What is the probability the experiment will produce a result of 0?
   (b) What is the probability the experiment will produce a result larger than 1?
2. Let $\lambda > 0$. For the problems below, assume the probability space is a Poisson(λ) distribution.
   (a) Let k be a non-negative integer. Calculate the ratio $\frac{P(\{k+1\})}{P(\{k\})}$.
   (b) Use (a) to calculate the mode of a Poisson(λ).
3. Suppose that the number of earthquakes that occur in a year in California has a Poisson distribution with parameter λ. Suppose that the probability that any given earthquake has magnitude at least 6 on the Richter scale is p.
   (a) Given that there are exactly n earthquakes in a year, find an expression (in terms of n and p) for the conditional probability that exactly one of them is magnitude at least 6.
   (b) Find an expression (in terms of λ and p) for the probability that there will be exactly one earthquake of magnitude at least 6 in a year.
   (c) Find an expression (in terms of n, λ, and p) for the probability that there will be exactly n earthquakes of magnitude at least 6 in a year.
4. Consider n vertices labeled $\{1, 2, \ldots, n\}$. Corresponding to each distinct pair $\{i, j\}$ we perform an independent Bernoulli (p) experiment and insert an edge between i and j with probability p. The graph constructed this way is denoted as $G(n, p)$.
   (a) Let $1 \leq i \leq n$. We say j is a neighbour of i if there is an edge between i and j. For some $1 \leq k \leq n$ determine the probability that i has k neighbours ?
   (b) Let $\lambda > 0$ and n large enough so that $0 < p = \frac{\lambda}{n} < 1$ and let $A_k = \{$ vertex 1 has k neighbours$\}$ what is the
   $$\lim_{n\to\infty} P(A_k)?$$
5. Suppose there are thirty balls in an urn, ten of which are black and the remaining twenty of which are red. Suppose three balls are selected from the urn (without replacement).
   (a) What is the probability that the sequence of draws is red-red-black?
   (b) What is the probability that the three draws result in exactly two red balls?
6. Consider a room of one hundred people – forty men and sixty women.
   (a) If ten people are selected from the room, find the probability that exactly six are women. Calculate this probability with and without replacement and compare the decimal approximations of your two results.
   (b) If ten people are selected from the room, find the probability that exactly seven are women. Calculate this probability with and without replacement and compare the decimal approximations of your two results.
   (c) If 100 people are selected from the room, find the probability that exactly sixty are women. Calculate this probability with and without replacement and compare the two answers.

(d) If 100 people are selected from the room, find the probability that exactly sixty-one are women. Calculate this probability with and without replacement and compare the two answers.

7. For the problems below, assume a HyperGeo(N, r, m) distribution.

   (a) Calculate the ratio $\frac{P(\{k+1\})}{P(\{k\})}$.
   (Assume that $max\{0, m-(N-r)\} \leq k \leq min\{r, m\}$ to avoid zero in the denominator).

   (b) Use (a) to calculate the mode of a HyperGeo(N, r, m).

8. Biologists use a technique called "capture-recapture" to estimate the size of the population of a species that cannot be directly counted. The following exercise illustrates the role a hypergeometric distribution plays in such an estimate.

   Suppose there is a species of unknown population size N. Suppose fifty members of the species are selected and given an identifying mark. Sometime later a sample of size twenty is taken from the population and it is found that four of the twenty were previously marked. The basic idea behind mark-recapture is that since the sample showed $\frac{4}{20} = 20\%$ marked members, that should also be a good estimate for the fraction of marked members of the species as a whole. However, for the whole species that fraction is $\frac{50}{N}$ which provides a population estimate of $N \approx 250$.

   Looking more deeply at the problem, if the second sample is assumed to be done at random without replacement and with each member of the population equally likely to be selected, the resulting number of marked members should follow a HyperGeo$(N, 50, 20)$ distribution.

   Under these assumptions use the formula for the mode calculated in the previous exercise to determine which values of N would cause a result of four marked members to be the most likely of the possible outcomes.