CS 174 Homework Assignment 3 (due Wednesday, Feb. 20)

1. Choose N coins from an infinite collection of coins, where N is a Poisson random variable with parameter λ. Each coin has a probability p of coming up heads (i.e., the same value for all coins). Toss each of the N coins once. Show that the total number of heads is a Poisson random variable with parameter λp.

2. Let X and Y be independent Poisson random variables with parameters λ and μ respectively. Let $Z = X + Y$. What is the conditional probability distribution $P(X|Z = n)$?

3. Let X be a Poisson random variable with parameter λ. Compute the mode of the Poisson distribution (the point of maximal probability). When is the mean equal to the mode? Is it possible to have $P(X = 1) > P(X = 0)$?

4. Let $X_1, X_2, \ldots, X_n$ be independent random variables, and let X_i be Bernoulli with parameter p_i. Show that $Y = \sum_{i=1}^{n} X_i$ has mean and variance given by:
$$\mathrm{E}[Y] = \sum_{i=1}^{n} p_i \qquad \mathrm{Var}[Y] = \sum_{i=1}^{n} p_i(1 - p_i).$$
Show that, for $\mathrm{E}[Y]$ fixed, $\mathrm{Var}[Y]$ is a maximum when $p_1 = p_2 = \cdots = p_n$. That is to say, the variation in the sum is greatest when individuals are most alike. Is this contrary to intuition?

5. Consider a random graph $G = (V, E)$ on n vertices, where edges are chosen independently with probability $P(n)$. Derive an upper bound on the probability that such a random graph G has no triangles.

6. This problem will help you to understand the inclusion-exclusion principle (page 1 of Note 3). Consider once again the balls-and-bins experiment, in which we toss m labeled balls into n labeled bins. We are going to use inclusion-exclusion to figure out precisely the probability that there are no empty bins.

   (a) Let E_i be the event that bin i is empty, and for any given set of k bins $i_1, i_2, \ldots, i_k$, let $p_{i_1 i_2 \ldots i_k}$ denote the probability that all these bins are empty, i.e., $p_{i_1 i_2 \ldots i_k} = \Pr[E_{i_1} \wedge E_{i_2} \wedge \ldots \wedge E_{i_k}]$. Show that $p_{i_1 i_2 \ldots i_k} = (1 - \frac{k}{n})^m$.

   (b) As in Note 3, define $S_k = \sum_{i_1, i_2, \ldots, i_k} p_{i_1 i_2 \ldots i_k}$. Show that $S_k = \binom{n}{k}(1 - \frac{k}{n})^m$.

   (c) Now let q denote the probability that there are no empty bins. Deduce from Theorem 1 of Note 3 that
$$q = 1 - \binom{n}{1}\left(1 - \frac{1}{n}\right)^m + \binom{n}{2}\left(1 - \frac{2}{n}\right)^m - \binom{n}{3}\left(1 - \frac{3}{n}\right)^m + \cdots \pm \binom{n}{n-1}\left(1 - \frac{n-1}{n}\right)^m.$$

(d) Now let us consider the special case $m = n \ln n + cn$. (Recall from our discussion of the coupon collecting problem in Note 1 that this is around the threshold required for no empty bins.) Show that, in this case, for any fixed value of k,

$$S_k = \binom{n}{k}\left(1 - \frac{k}{n}\right)^m \sim \frac{1}{k!}\mathrm{e}^{-kc}.$$

Hint: If k is fixed, then it should be clear that $\binom{n}{k} \sim \frac{n^k}{k!}$. And you already know how to handle things like $\left(1 - \frac{k}{n}\right)^m$.

(e) Deduce that when $m = n \ln n + cn$, for any constant c,

$$\Pr[\text{there are no empty bins}] \sim \mathrm{e}^{-\mathrm{e}^{-c}} \qquad \text{as } n \to \infty.$$

(f) Recall that we proved on page 3 of Note 1 that $\mathrm{E}[X] \sim \mathrm{e}^{-c}$, where the r.v. X denotes the number of empty bins. Applying Markov's inequality immediately gives us the bound $\Pr[X \geq 1] \leq \mathrm{e}^{-c}$, and hence

$$\Pr[\text{there are no empty bins}] \geq 1 - \mathrm{e}^{-c} \qquad \text{as } n \to \infty.$$

The exact asymptotic value we have just proved in part (e) (with quite a bit more work!) is stronger than this bound, especially when c is small. To see this, take $c = 0$ (i.e., we are throwing exactly $n \ln n$ balls) and show that the above bound from Note 1 is useless while the exact value computed in part (e) gives a significant probability of finding no empty bins. What happens for large values of c? How about negative values?