Stat 201A, Fall 2012
HOMEWORK 5 (due Tuesday 9/11)

1. In a population of 800 voters, 200 favor Candidate A. In a simple random sample of 30 voters, there is at least a 90% chance that the proportion favoring Candidate A is in the range ________________________

Fill in the blank with the best numerical range you can find.

2. A coin that lands heads with probability p is tossed repeatedly. Let H_n be the number of heads in the first n tosses. For integer $k \geq 0$, find the correlation between H_n and H_{n+k}. For fixed n, describe the behavior of your answer as p and k vary, and explain whether the behavior makes intuitive sense to you.

3. A fair die is rolled n times. Let S be the number of times the face with six spots appears, and let F be the number of times the face with five spots appears. Let $D = S - F$. Find $E(D)$ and $SD(D)$.

[You're going to have to find a covariance term. One quick way is to notice that S, F, and $S + F$ all have recognizable famous distributions.]

4. A "without replacement" version of the geometric. A standard deck has 52 cards, of which 4 are aces. The deck is well shuffled, and cards are dealt one by one without replacement till the first ace appears. Let X be the number of cards dealt.

a) Find $E(X)$.

b) Find $SD(X)$.

[Do *not* find the distribution of X and leave your formulas in terms of that distribution; I won't even accept it if you then run the formulas through R to get numerical answers. Notice that X is a count, and recall how to write a count as a sum. This will help in both parts. Also look in Homework 1 for a probability calculation closely related to what you need here.]

5. This problem is the start of a sequence in which you will prove a limit theorem.

Each of n students in a class submits a homework set. Suppose the papers are returned at random to the students, one paper per student. Let M_n be the number of students who get back their own homework.

a) For each i and j in the range 1 through n, with $i \neq j$, find P(student i gets back his/her own homework) and P(both students i and j get back their own homeworks).

b) Give an intuitive justification for why the distribution of M_n should be approximately Poisson(1) for large n.

c) Find $E(M_n)$ and $SD(M_n)$, and show that their limits as n gets large agree with the Poisson limit in **b**.

6 (continuing the previous problem). The goal now is to formally derive the Poisson limit above. For this you will need a preliminary result that you probably already know. Let $A_1, A_2, \ldots, A_n$ be events. Then the *inclusion-exclusion* formula is

$$P(\bigcup_{i=1}^{n} A_i) \;=\; \sum_{i=1}^{n} P(A_i) - \sum\sum_{i<j} P(A_iA_j) + \sum\sum\sum_{i<j<k} P(A_iA_jA_k) - cdots + (-1)^{n+1}P(A_1A_2\ldots A_n)$$

Prove this formula by induction. Recall that the case $n = 2$ is proved in Lemma 1.6 of the text.

You don't have to prove that again.

7 (continuing the previous problem). In this problem you will derive $\lim_{n\to\infty} P_n(0)$, where $P_n(0) = P(M_n = 0)$. Consider the complement of the event $\{M_n = 0\}$ and see how it is related to the events $A_i =$ "student i gets back his/her own homework". Then use the previous problem to get an exact formula for $P_n(0)$. Finally, show that $\lim_{n\to\infty} P_n(0)$ agrees with the Poisson(1) prediction you made a couple of problems ago.

8 (finishing up the proof sequence). Now fix k in the range $0 < k \leq n$. Let $P_n(k) = P(M_n = k)$ be the probability that exactly k of the n students get their own homeworks back. Explain why

$$P_n(k) \quad = \quad \binom{n}{k}\frac{(n-k)!}{n!} \cdot P_{n-k}(0)$$

It might help to note that $m!P_m(0)$ is the number of arrangements of m homeworks so that there are no matches at all; these are called "derangements."

Now plug in (appropriately) the exact formula you derived in the previous problem, and show that for each fixed k, $\lim_{n\to\infty} P_n(k)$ agrees with the Poisson(1) prediction you made a few problems ago.