Homework of Week 6

Deadline: 9:00am, November 22(Thursday), 2018

1. Consider the probability that every bin receives exactly one ball when n balls are thrown randomly into n bins.

   - Give an upper bound on this probability using the condition-free Poisson approximation.
   - Determine the exact probability of this event.

2. Let $X_i^{(m)}$ be the number of balls in bin i when m balls are independently and uniformly thrown at random into n bins, and $Y_i^{(m)}, 1 \leq i \leq n$, are independent Poisson random variables each having expectation m/n. Assume that f is a nonnegative function.

   - Prove that if $E[f(X_1^{(m)}, ...X_n^{(m)})]$ is monotonically increasing in m, then $E[f(X_1^{(m)}, ...X_n^{(m)})] \leq 2E[f(Y_1^{(m)}, ...Y_n^{(m)})]$. (Hint: Show that $E[f(Y_1^{(m)}, ...Y_n^{(m)})] \geq E[f(X_1^{(m)}, ...X_n^{(m)})]Pr(\sum Y_i^{(m)} \geq m)$ and $Pr(\sum Y_i^{(m)} \geq m) \geq 1/2$.)
   - (**Bonus score 5 points**) If $E[f(X_1^{(m)}, ...X_n^{(m)})]$ is monotonically decreasing in m, then $E[f(X_1^{(m)}, ...X_n^{(m)})] \leq 2E[f(Y_1^{(m)}, ...Y_n^{(m)})]$.

3. Bloom filters can be used to estimate set differences. Suppose Alice has a set X and Bob has a set Y, both with n elements. For example, the sets might represent their 100 favorite songs. Alice and Bob create Bloom filters of their sets respectively, using the same number of bits m and the same k hash functions. Determine the expected number of bits where our Bloom filters differ as a function of m, n, k and $|X \bigcap Y|$. Explain how this could be used as a tool to find people with the same taste in music more easily than comparing lists of songs directly.

4. Do Bernoulli experiment for 20 trials, using a new 1-Yuan coin. Record the result in a string $s_1 s_2 ... s_i ... s_{20}$, where s_i is 1 if the i^{th} trial gets Head, and otherwise is 0.