UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: PROBABILITY AND RANDOM PROCESSES

Problem Set 5
Fall 2015

Issued: Thursday, October 1, 2015 **Due:** 9:00am Thursday, October 8, 2015

Problem 1. Let $X_1, X_2, \ldots, X_n$ be n i.i.d. exponential random variables with mean 1.

(a) Find the pdf of $Y = X_1 + X_2 + \cdots + X_n$. *Hint:* find the transform of Y and think about the inverse transform.

(b) Let $X_M = \max_i X_i$ be the maximum of these random variables. What is the joint distribution of X_M and the rest of the random variables?

Problem 2. Consider ten independent and identically distributed random variables $X_1, X_2, \ldots, X_{10}$, with each X_i uniformly distributed over the interval [0,1].

(a) Use the Markov inequality to bound the probability

$$\Pr(X_1 + X_2 + \cdots + X_{10} \geq 7). \tag{1}$$

(b) Repeat part (a) using the Chebyshev inequality.

(c) Approximate the probability in part (a) using the Central Limit Theorem.

(d) Use Python or any other programming language to simulate the probability in (1). Among the bounds or approximations you get in part (a), (b), and (c), which is the closest to the answer you get in the simulations?

Problem 3. Using Central Limit Theorem, compute the number of people to poll in a public opinion survey to estimate the fraction of the population that will vote in favor of a proposition within a fraction of α, with probability at least $1 - \beta$. Specifically , let p be the fraction of the whole population that will vote in favor of the proposition, n be the number of people that we poll, and S_n be the number of people who vote in favor of a proposition among the n people that we poll. We want to find a lower bound of n such that

$$\Pr(|\frac{S_n}{n} - p| \geq \alpha) \leq \beta.$$

Problem 4. Let X_i, $1 \le i \le n$ be a sequence of i.i.d. random variables distributed uniformly in $[-1, 1]$. Show that the following sequences converge in probability to some limit.

(a) $Y_n = X_n/n$

(b) $Y_n = (X_n)^n$

(c) $Y_n = \prod_{i=1}^{n} X_i$

Problem 5. Consider the Markov chain of Figure 1, where $a, b \in (0, 1)$.

(a) Find the invariant distribution.

(b) Calculate $\Pr(X(1) = 1, X(2) = 0, X(3) = 0, X(4) = 1 | X(0) = 0)$.

(c) Show that the Markov chain is aperiodic.

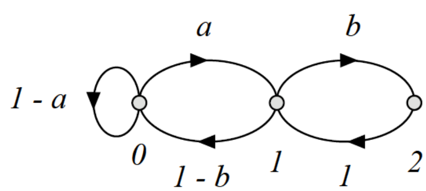

Figure 1: Markov Chain for Problem 5

Problem 6. Consider the Markov chain in Figure 2. Suppose that $X(0) = 1$. Calculate the expected number of times that the state $X(t)$ becomes 1 before being absorbed in state 3. ($X(0) = 1$ is included in this number.)

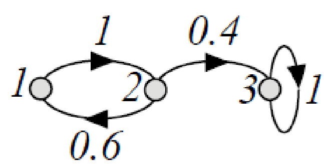

Figure 2: Markov Chain for Problem 6

Problem 7. In class, we learned some inequalities such as the Markov inequality, the Chebyshev inequality, and the Chernoff bound. In this problem, we will derive an inequality, which is a special case of Chernoff bound, using a simple counting method.

Suppose $X_1, \dots, X_n$ are i.i.d. Bernoulli random variables with $\Pr(X_i = 1) = 1/2$.

(a) First, use the Chebyshev inequality to show that for any $\epsilon > 0$,

$$\Pr\left(\sum_{i=1}^{n} X_i \ge \frac{n}{2}(1+\epsilon)\right) \le \frac{1}{\epsilon^2 n}. \tag{2}$$

The special case of Chernoff bound that we will derive is as follows: for any $\epsilon > 0$,

$$\Pr\left(\sum_{i=1}^{n} X_i \geq \frac{n}{2}(1+\epsilon)\right) \leq \exp\{-\frac{\epsilon^2 n}{10}\}. \tag{3}$$

We will derive (3) in the next steps. We should notice that if $\epsilon > 1$, we have $\Pr(\sum_{i=1}^{n} X_i \geq \frac{n}{2}(1+\epsilon)) = 0$. Therefore, we only need to consider the cases when $0 < \epsilon \leq 1$.

(b) Let M be the event that $X_1 = X_2 = \cdots = X_m = 1$, $m < n$. Show that for an integer k ($m \leq k \leq n$),

$$\Pr(M | \sum_{i=1}^{n} X_i = k) \geq (\frac{k-m}{n-m})^m,$$

and further, show that

$$\Pr(M | \sum_{i=1}^{n} X_i \geq k) \geq (\frac{k-m}{n-m})^m.$$

(c) For simplicity, we assume that $\frac{\epsilon n}{4}$ is an integer and let $m = \frac{\epsilon n}{4}$. Let G be the event that $\sum_{i=1}^{n} X_i \geq \frac{n}{2}(1+\epsilon)$. Show that

$$\Pr(M|G) \geq (\frac{1}{2} + \frac{\epsilon}{4})^m.$$

(d) Show that $\Pr(M) \geq \Pr(G)\Pr(M|G)$. Then show that

$$\Pr(G) \leq (1 + \frac{\epsilon}{2})^{-m}.$$

(e) Combining the fact that for any $0 < \epsilon \leq 1$,

$$\ln(1 + \frac{\epsilon}{2}) > \frac{2}{5}\epsilon, \tag{4}$$

show that (3) holds. (You do not need to prove (4).)

(f) Compare (2) and (3) and argue why the Chernoff bound is better than the Chebyshev inequality.