UC Berkeley
Department of Electrical Engineering and Computer Sciences

ELECTRICAL ENGINEERING 126: PROBABILITY AND RANDOM PROCESSES

Problem Set 5
Spring 2017

Issued: Thursday, February 23 **Due:** 8am Thursday, March 2

---

1. **Confidence Interval Comparisons**

   In order to estimate the probability of a head in a coin flip, p, you flip a coin n times and count the number of heads, S_n. You use the estimator $\hat{p} = S_n/n$.

   (a) You choose the sample size n to have a guarantee

   $$\Pr(|\hat{p} - p| \geq \epsilon) \leq \delta.$$

   Using Chebyshev Inequality, determine n with the following parameters:

   (i) Compare the value of n when $\epsilon = 0.05$, $\delta = 0.1$ to when $\epsilon = 0.1, \delta = 0.1$.

   (ii) Compare the value of n when $\epsilon = 0.1, \delta = 0.05$ to when $\epsilon = 0.1, \delta = 0.1$.

   (b) Now, we change the scenario slightly. You know that $p \in (0.4, 0.6)$ and would now like to determine the smallest n such that

   $$\Pr\left(\frac{|\hat{p} - p|}{p} \leq 0.05\right) \geq 0.95.$$

   Use the CLT to find the value of n that you should use.

2. **Chernoff Bound**

   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 $X_1, \ldots, X_n$ are i.i.d. Bernoulli random variables with $\Pr(X_i = 1) = 1/2$. Let $S_n = \sum_{i=1}^{n} X_i$.

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

   $$\Pr\left(S_n \geq \frac{n}{2}(1+\epsilon)\right) \leq \frac{1}{\epsilon^2 n}. \tag{1}$$

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

   $$\Pr\left(S_n \geq \frac{n}{2}(1+\epsilon)\right) \leq \exp\left(-\frac{\epsilon^2 n}{10}\right). \tag{2}$$

   We will derive (2) in the next steps. We should notice that if $\epsilon > 1$, we have $\Pr(S_n \geq (n/2)(1+\epsilon)) = 0$. 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 ($m \le k \le n$),

$$\Pr(M \mid S_n = k) \ge \left(\frac{k-m}{n-m}\right)^m,$$

and further, show that

$$\Pr(M \mid S_n \ge k) \ge \left(\frac{k-m}{n-m}\right)^m.$$

(c) For simplicity, we assume that $\epsilon n/4$ is an integer and let $m = \epsilon n/4$. Let G be the event that $S_n \ge (n/2)(1+\epsilon)$. Show that

$$\Pr(M \mid G) \ge \left(\frac{1}{2} + \frac{\epsilon}{4}\right)^m.$$

(d) Show that $\Pr(M) \ge \Pr(G)\Pr(M \mid G)$. Then show that

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

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

$$\ln\left(1 + \frac{\epsilon}{2}\right) > \frac{2}{5}\epsilon, \tag{3}$$

show that (2) holds. (You do not need to prove (3).)

(f) Compare (1) and (2) and argue why the Chernoff bound is better than the Chebyshev inequality.

3. Chernoff Bound Application: Load Balancing

Here, we will give an application for the Chernoff bound derived in the previous question. However, we will need a slightly more general version of the bound that works for any Bernoulli random variables. If $X_1, \ldots, X_n$ are i.i.d. Bernoulli, with $\Pr(X_i = 1) = p$, and $S_n = \sum_{i=1}^n X_i$, then the following bound holds for $0 \le \varepsilon \le 1$:

$$\Pr(S_n > (1+\varepsilon)np) \le \exp\left(-\frac{\varepsilon^2 np}{3}\right). \tag{4}$$

You may take (4) as a fact (or try to prove it on your own if you want!).

Here is the setting: there are k servers and n users. The simplest load balancing scheme is simply to assign each user to a server chosen uniformly at random (think of the users as "balls" and we are tossing them into server "bins"). By using the union bound, show that with probability at least $1 - 1/k^2$, the maximum load of any server is at most $n/k + 3\sqrt{\ln k}\sqrt{n/k}$.

4. Convergence in Probability

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) $Y_n = \prod_{i=1}^n X_i$.
(c) $Y_n = \max\{X_1, X_2, \dots, X_n\}$.
(d) $Y_n = (X_1^2 + \cdots + X_n^2)/n$.

5. **Almost Sure Convergence**

In this question, we will explore almost sure convergence and compare it to convergence in probability. Recall that a sequence of random variables (X_n) converges **almost surely** (abbreviated a.s.) to X if $\Pr(\lim_{n\to\infty} X_n = X) = 1$.

(a) Suppose that, with probability 1, the sequence (X_n) oscillates between two values $a \neq b$ infinitely often. Is this enough to prove that (X_n) does *not* converge almost surely? Justify your answer.

(b) Suppose that Y is uniform on $[-1, 1]$, and X_n has distribution

$$\Pr(X_n = (y + n^{-1})^{-1} \mid Y = y) = 1.$$

Does X_n converge a.s.?

(c) Define random variables X_n in the following way: first, set each X_n to 0. Then, for each $k \in \mathbb{N}$, pick j uniformly randomly in $\{2^k, \dots, 2^{k+1} - 1\}$ and set $X_j = 2^k$. Does the sequence (X_n) converge a.s.?

(d) Does the sequence (X_n) from the previous part converge in probability to some X? If so, is it true that $E[X_n] \to E[X]$?

6. **Compression of a Random Source (Optional)**

Consider n i.i.d. random variables with $X_i \sim \text{Bernoulli}(p)$. Additionally define the entropy of a random variable X as $H(X) = -\sum_{x\in\mathcal{X}} p(X = x) \log P(X = x)$. That is, we define $H(X) = E_p[\log \frac{1}{p(X)}]$. In this problem, we will show that a random source whose symbols are drawn according to the distribution of X_i can be compressed to $H(X)$ bits per symbol.

(a.) Show that

$$-\frac{1}{n} \log p(X_1, X_2, \dots, X_n) \to H(X_1)$$

Now, define A as the set of all sequences $(x_1, x_2, \dots, x_n) \in \mathcal{X}^n$ such that:

$$2^{-n(H(X_1)+\epsilon)} \leq p(x_1, x_2, \dots, x_n) \leq 2^{-n(H(X_1)-\epsilon)}$$

(b.) Show that $P((x_1, x_2, \dots, x_n) \in A) > 1 - \epsilon$

(c.) Show that $(1 - \epsilon)2^{n(H(X_1)-\epsilon)} \leq |A| \leq 2^{n(H(X_1)+\epsilon)}$

7. **Lognormal Distribution and the Moment Problem (Optional)**

This question seeks to answer the following question: if two distributions have the same moments of all orders, are they necessarily the same? An equivalent way to phrase the problem is: if the moments exist, do they completely determine the distribution?

(a) Suppose that Z is a standard Gaussian and let $X = e^Z$. Calculate the density of X. (This is known as the **lognormal** distribution.)

(b) Let $f_X(x)$ denote the density of the lognormal density. Define

$$f_a(x) = f_X(x)(1 + a\sin(2\pi \log x)), \qquad x > 0, \quad -1 \leq a \leq 1.$$

Argue that $f_a(x)$ is a valid density function and show that $f_X(x)$ and $f_a(x)$ have the same moments of all orders by showing that

$$\int_0^\infty x^k f_X(x) \sin(2\pi \log x)\, \mathrm{d}x = 0, \qquad k \in \mathbb{N}.$$

(c) Explicitly calculate the moments of the lognormal distribution.

(d) Now, let Y_b $(b > 0)$ be a discrete random variable with distribution

$$\Pr(Y_b = be^n) = c_n b^{-n} e^{-n^2/2}, \qquad n \in \mathbb{Z},$$

where c_n is chosen to normalize the distribution:

$$\sum_{n=-\infty}^{\infty} c_n b^{-n} e^{-n^2/2} = 1.$$

Show that Y_b has the same moments as X. This provides a discrete counterexample to the moment problem.