CS 70 Discrete Mathematics and Probability Theory
Summer 2025 Tate

# HW 06

Due: Saturday, 8/9, 4:00 PM
Grace period until Saturday, 8/9, 6:00 PM
Remember to show your work for all problems!

## Sundry

Before you start writing your final homework submission, state briefly how you worked on it. Who else did you work with? List names and email addresses. (In case of homework party, you can just describe the group.) If you used an LLM, place transcripts of your chats here.

## 1 Deriving the Chernoff Bound

Note 17

We've seen the Markov and Chebyshev inequalities already, but these inequalities tend to be quite loose in most cases. In this question, we'll derive the *Chernoff bound*, which is an *exponential* bound on probabilities.

The Chernoff bound is a natural extension of the Markov and Chebyshev inequalities: in Markov's inequality, we utilize only information about $\mathbb{E}[X]$; in Chebyshev's inequality, we utilize only information about $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$ (in the form of the variance). In the Chernoff bound, we'll end up using information about $\mathbb{E}[X^k]$ for *all* $k$, in the form of the *moment generating function* of $X$, defined as $\mathbb{E}[e^{tX}]$. (It can be shown that the $k$th derivative of the moment generating function evaluated at $t = 0$ gives $\mathbb{E}[X^k]$.)

In several subparts, we'll ask you to express your answer as a single exponential function, which has the form $e^{f(t)} = \exp(f(t))$ for some function $f$.

Here, we'll derive the Chernoff bound for the binomial distribution. Suppose $X \sim \text{Binomial}(n, p)$.

(a) We'll start by computing the *moment generating function* of $X$. That is, what is $\mathbb{E}[e^{tX}]$ for a fixed constant $t > 0$? (Your answer should have no summations.)

*Hint*: It can be helpful to rewrite $X$ as a sum of Bernoulli RVs.

(b) A useful inequality that we'll use is that

$$1 - \alpha \leq e^{-\alpha},$$

for any $\alpha$. Since we'll be working a lot with exponentials here, use the above to find an upper bound for your answer in part (a) as a single exponential function. (This will make the expressions a little nicer to work with in later parts.)

(c) Use Markov's inequality to give an upper bound for $\mathbb{P}[e^{tX} \geq e^{t(1+\delta)\mu}]$, for $\mu = \mathbb{E}[X] = np$ and a constant $\delta > 0$.

Use this to deduce an upper bound on $\mathbb{P}[X \geq (1+\delta)\mu]$ for any constant $\delta > 0$. (Your bound should be a single exponential function, where $f$ should also depend on $\mu = np$ and $\delta$.)

(d) Notice that so far, we've kept this new parameter $t$ in our bound—the last step is to optimize this bound by choosing a value of $t$ that minimizes our upper bound.

Take the derivative of your expression with respect to $t$ to find the value of $t$ that minimizes the bound. Note that from part (a), we require that $t > 0$; make sure you verify that this is the case!

Use your value of $t$ to verify the following Chernoff bound on the binomial distribution:

$$\mathbb{P}[X \geq (1+\delta)\mu] \leq \exp(-\mu(1+\delta)\ln(1+\delta) + \delta\mu).$$

*Note: As an aside, if we carried out the computations without using the bound in part (b), we'd get a better Chernoff bound, but the math is a lot uglier. Furthermore, instead of looking at the binomial distribution (i.e. the sum of independent and identical Bernoulli trials), we could have also looked at the sum of independent but* not necessarily identical *Bernoulli trials as well; this would give a more general but very similar Chernoff bound.*

(e) Let's now look at how the Chernoff bound compares to the Markov and Chebyshev inequalities. Let $X \sim \text{Binomial}(n = 100, p = \frac{1}{5})$. We'd like to find $\mathbb{P}[X \geq 30]$.

(i) Use Markov's inequality to find an upper bound on $\mathbb{P}[X \geq 30]$.

(ii) Use Chebyshev's inequality to find an upper bound on $\mathbb{P}[X \geq 30]$.

(iii) Use the Chernoff bound from part (d) to find an upper bound on $\mathbb{P}[X \geq 30]$.

(iv) Now use a calculator to find the exact value of $\mathbb{P}[X \geq 30]$. How did the three bounds compare? That is, which bound was the closest and which bound was the furthest from the exact value?

# 2 Exponential Median

Note 21

(a) Prove that if $X_1, X_2, \ldots, X_n$ are mutually independent exponential random variables with parameters $\lambda_1, \lambda_2, \ldots, \lambda_n$, then $\min(X_1, X_2, \ldots, X_n)$ is exponentially distributed with parameter $\sum_{i=1}^{n} \lambda_i$.

*Hint*: Recall that the CDF of an exponential random variable with parameter $\lambda$ is $1 - e^{-\lambda t}$.

(b) Given that the minimum of three i.i.d exponential variables with parameter $\lambda$ is $m$, what is the probability that the difference between the median and the smallest is at least $s$? Note that the exponential random variables are mutually independent.

(c) What is the expected value of the median of three i.i.d. exponential variables with parameter $\lambda$?

*Hint:* Part (b) may be useful for this calculation.

## 3 Interesting Gaussians

Note 21

(a) If $X \sim N(0, \sigma_X^2)$ and $Y \sim N(0, \sigma_Y^2)$ are independent, then what is $\mathbb{E}\big[(X+Y)^k\big]$ for any *odd* $k \in \mathbb{N}$?

(b) Let $f_{\mu,\sigma}(x)$ be the density of a $N(\mu, \sigma^2)$ random variable, and let $X$ be distributed according to $\alpha f_{\mu_1,\sigma_1}(x) + (1-\alpha) f_{\mu_2,\sigma_2}(x)$ for some $\alpha \in [0,1]$. Compute $\mathbb{E}[X]$ and $\mathrm{Var}(X)$. Is $X$ normally distributed?

## 4 Uniform Estimation

Note 17
Note 21

Let $U_1, \ldots, U_n$ be i.i.d Uniform$(-\theta, \theta)$ for some unknown $\theta \in \mathbb{R}$, $\theta > 0$. We wish to estimate $\theta$ from the data $U_1, \ldots, U_n$.

(a) Why would using the sample mean $\overline{U} = \frac{1}{n}\sum_{i=1}^{n} U_i$ fail in this situation?

(b) Find the PDF of $U_i^2$ for $i \in \{1, \ldots, n\}$.

(c) Consider the following variance estimate:

$$V = \frac{1}{n}\sum_{i=1}^{n} U_i^2.$$

Show that for large $n$, the distribution of $V$ is close to one of the famous ones, and provide its name and parameters.

(d) Use part (c) to construct an unbiased estimator for $\theta^2$ that uses all the data.

(e) Let $\sigma^2 = \mathrm{Var}(U_i^2)$. We wish to construct a confidence interval for $\theta^2$ with a significance level of $\delta$, where $0 < \delta < 1$.

*Note*: A $(1-\delta)$ confidence interval has a *significance level* of $\delta$.

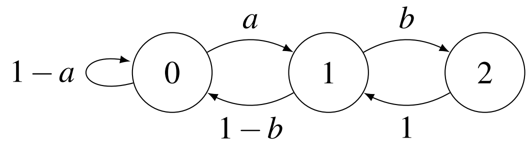
(i) Without any assumption on the magnitude of $n$, construct a confidence interval for $\theta^2$ with a significance level of $\delta$ using your estimator from part (d).

(ii) Suppose $n$ is large. Construct an approximate confidence interval for $\theta^2$ with a significance level of $\delta$ using your estimator from part (d). You may leave your answer in terms of $\Phi$ and $\Phi^{-1}$, the normal CDF and its inverse.

## 5 Analyze a Markov Chain

Note 22

Consider a Markov chain with the state diagram shown below where $a, b \in (0,1)$.




Here, we let $X(n)$ denote the state at time $n$.

(a) Is this Markov chain irreducible? Is this Markov chain aperiodic? Justify your answers.

(b) Calculate $\mathbb{P}[X(1)=1, X(2)=0, X(3)=1, X(4)=2 \mid X(0)=0]$.

(c) Calculate the invariant distribution. Do all initial distributions converge to this invariant distribution? Justify your answer.

# 6 A Bit of Everything

Note 22

Suppose that $X_0, X_1, \ldots$ is a Markov chain with finite state space $S = \{1, 2, \ldots, n\}$, where $n > 2$, and transition matrix $P$. Suppose further that

$$P(1,i) = \frac{1}{n} \quad \text{for all states } i \text{ and}$$
$$P(j, j-1) = 1 \quad \text{for all states } j \neq 1,$$

with $P(i,j) = 0$ everywhere else.

(a) Prove that this Markov chain is irreducible and aperiodic.

(b) Suppose you start at state 1. What is the distribution of $T$, where $T$ is the number of transitions until you leave state 1 for the first time?

(c) Again starting from state 1, what is the expected number of transitions until you reach state $n$ for the first time?

(d) Again starting from state 1, what is the probability you reach state 2 before you reach state $n$?

(e) Compute the stationary distribution of this Markov chain.

# 7 Playing Blackjack

Note 22

Suppose you start with \$1, and at each turn, you win \$1 with probability $p$, or lose \$1 with probability $1-p$. You will continually play games of Blackjack until you either lose all your money, or you have a total of $n$ dollars.

(a) Formulate this problem as a Markov chain.

(b) Let $\alpha(i)$ denote the probability that you end the game with $n$ dollars, given that you started with $i$ dollars.

Notice that for $0 < i < n$, we can write $\alpha(i+1) - \alpha(i) = k(\alpha(i) - \alpha(i-1))$. Find $k$.

(c) Using part (b), find $\alpha(i)$, where $0 \leq i \leq n$. (You will need to split into two cases: $p = \frac{1}{2}$ or $p \neq \frac{1}{2}$.)

*Hint*: Try to apply part (b) iteratively, and look at a telescoping sum to write $\alpha(i)$ in terms of $\alpha(1)$. The formula for the sum of a finite geometric series may be helpful when looking at the case where $p \neq \frac{1}{2}$:

$$\sum_{k=0}^{m} a^k = \frac{1 - a^{m+1}}{1 - a}.$$

Lastly, it may help to use the value of $\alpha(n)$ to find $\alpha(1)$ for the last few steps of the calculation.

(d) As $n \to \infty$, what happens to the probability of ending the game with $n$ dollars, given that you start with $i$ dollars, with the following values of $p$?

(i) $p > \frac{1}{2}$

(ii) $p = \frac{1}{2}$

(iii) $p < \frac{1}{2}$