1. Toss a fair coin $n$ times. What is the probability that we get at least $3n/4$ heads? Let $S_n$ denote the number of heads.

(i) Use Chebyshev’s inequality to show that $P[S_n \geq \frac{3}{4}n] \leq \frac{4}{n}$. That is, the probability converges to zero at least linearly in $n$.

(ii) Apply the CLT to argue that asymptotically, $P[S_n \geq \frac{3}{4}n]$ is $P[g \geq \sqrt{n}/8]$, where $g \sim N(0,1)$. Conclude that we may instead expect the probability to converge to zero exponentially in $n$, $P[S_n \geq \frac{3}{4}n] \leq \frac{1}{\sqrt{2\pi}}e^{-n/8}$.

(iii) We want non-asymptotic bounds. Explain why the Berry-Esseen nonasymptotic version of the CLT is less informative than Chebyshev’s inequality in this application.

(iv) Apply a Chernoff bound for Poisson trials to derive a nonasymptotic tail bound on $P[S_n \geq \frac{3}{4}n]$ that converges to zero exponentially fast in $n$.

2. Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with mean $\mu$ and finite variance. Show that

$$E \left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| = O \left( \frac{1}{\sqrt{n}} \right) \quad \text{as } n \to \infty.$$ 

Hint: use the tail sum formula.

3. Let $X$ be a Binomial random variable with parameters $n$ and $p$.

(i) Use the CLT to provide a tail bound for $P[|X - E[X]| > \delta E[X]]$ that holds asymptotically as $n \to \infty$.

(ii) Use a Chernoff bound for Poisson trials to determine a tail bound of the same form.

4. Consider a collection $X_1, \ldots, X_n$ of $n$ independent integers chosen uniformly from the set $\{0, 1, 2\}$. Let $X = \sum_{i=1}^{n} X_i$ and $0 < \delta < 1$. Derive a Chernoff bound for $P(X \geq (1+\delta)n)$ and $P(X \leq (1-\delta)n)$.

5. Consider a collection $X_1, \ldots, X_n$ of $n$ independent geometrically distributed random variables with mean $2$. Let $X = \sum_{i=1}^{n} X_i$ and $\delta > 0$.

(i) Derive a bound on $P(X \geq (1+\delta)(2n))$ by applying a Chernoff bound to a sequence of $(1+\delta)(2n)$ fair coin tosses.

(ii) Directly derive a Chernoff bound on $P(X \geq (1+\delta)(2n))$ using the moment generating function for geometric random variables. The form of the bound should be simple.

(iii) Which bound is better in your opinion, and why?

6. [required only for CSCI6220] Consider $n$ balls thrown randomly into $n$ bins. Let $X_i = \mathbf{1}$ if the $i$th bin is empty and 0 otherwise. Let $X = \sum_{i=1}^{n} X_i$. Let $Y_i$ for $i = 1, \ldots, n$ be independent Bernoulli random variables that are 1 with probability $p = (1-1/n)^n$. Let $Y = \sum_{i=1}^{n} Y_i$.

(i) Show that $E[X_1 X_2 \cdots X_k] \leq E[Y_1 Y_2 \cdots Y_k]$ for any $k \geq 1$.

(ii) Show that $E[e^{tX}] \leq E[e^{tY}]$ for all $t \geq 0$. (Hint: use the expansion for $e^x$ and compare $E[X^k]$ to $E[Y^k]$.)

(iii) Derive a Chernoff bound for $P(X \geq (1+\delta)E[X])$. 