STAB52 Fall 2024 Name (F, L): Zay Lin Aung

Quiz 7 Student ID #: 1010388171

We consider some probability space $(\Omega, \mathbb{P})$. Let $X_1, \dots, X_n$ be independent and identically distributed random variables whose common distribution is uniform over $\{-1, 1\}$. We define $S_n = X_1 + \cdots + X_n$ be the sum.

1. (6 points) Find $\mathbb{E}[X_1]$ and $\mathbb{V}[X_1]$ ($= \mathrm{Var}(X_1)$) and use these to find an upper bound on
$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right),$$
for $\epsilon > 0$.

   We will now see another way to obtain a tighter (i.e., smaller) upper bound on the previous quantity.

2. (4 points) For $\lambda > 0$, find $\mathbb{E}[e^{\lambda X_1}]$ and prove that the random variables $e^{\lambda X_1}, \dots, e^{\lambda X_n}$ are jointly independent.
   Hint: Show that the joint pmf is the product of marginal pmfs.

3. (3 points) Deduce from the last question that
$$\mathbb{E}[e^{\lambda S_n/n}] \leq e^{\lambda^2/(2n)}.$$
   We take as given that the hyperbolic cosine function $\cosh\colon x \mapsto \frac{e^x + e^{-x}}{2}$ satisfies $\cosh(x) \leq e^{x^2/2}$ for any $x \in \mathbb{R}$.

4. (7 points) Prove that for any $\epsilon > 0$ and $\lambda > 0$
$$\mathbb{P}\left(\frac{S_n}{n} \geq \epsilon\right) \leq e^{\frac{\lambda^2}{2n} - \lambda\epsilon},$$
   and then,
$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right) \leq 2e^{-n\epsilon^2/2}.$$
   Hint: Note that X_1 and $-X_1$ have the same distribution and that, for any random variable Y, $\{|Y| \geq \epsilon\}$ is equivalent to $\{Y \geq \epsilon\} \cup \{-Y \geq \epsilon\}$.

Remark: Comparing the two bounds (in Questions 1 and 4), you will see that the bound in Question 4 decays much faster (i.e., it is much tighter) as $n\epsilon^2$ increases.

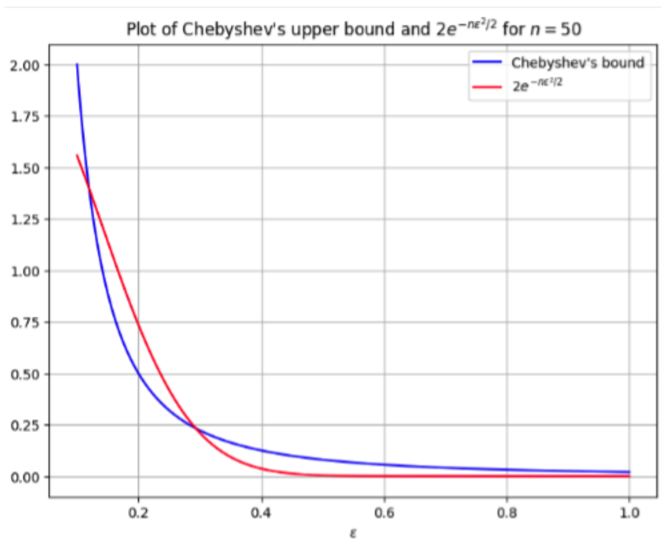

Figure 1: Upper bounds on $\mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq \epsilon\right)$ from Chebyshev's inequality and Question 4.

1. $X_1 \sim \text{Uniform}\{-1, 1\}$

$$E[X_1] = (-1)\tfrac{1}{2} + (1)(\tfrac{1}{2}) = -\tfrac{1}{2} + \tfrac{1}{2} = 0$$

$$E[X_1^2] = (-1)^2(\tfrac{1}{2}) + (1)^2(\tfrac{1}{2}) = \tfrac{1}{2} + \tfrac{1}{2} = 1$$

$$\text{Var}[X_1] = E[X_1^2] - E[X_1]^2 = 1 - 0 = 1$$

$$P\left(\left|\frac{S_n}{n}\right| \geqslant \varepsilon\right)$$

$$= P\left(\frac{|S_n|}{n} \geqslant \varepsilon\right) \quad (\because n \geqslant 0)$$

$$= P(|S_n| \geqslant \varepsilon n)$$

However, $\mu_n = E[S_n] = E[X_1 + \dots + X_n]$

$$= E[X_1] + \dots + E[X_n] \quad (\text{by linearity})$$

$$= 0 + \dots + 0 \quad (X_1, \dots X_n \text{ i.i.d})$$

$$= 0$$

$$\text{Var}[S_n] = \text{Var}[X_1 + \dots + X_n]$$

$$= \text{Var}[X_1] + \dots + \text{Var}[X_n] \quad (X_1, \dots X_n \text{ i.i.d})$$

$$= 1 + \dots + 1$$

$$= n$$

$$\therefore P(|S_n| \geqslant \varepsilon n) = P(|S_n - \mu| \geqslant \varepsilon n) \leq \frac{\sigma^2}{(\varepsilon n)^2} = \frac{n}{\varepsilon^2 n^2} = \frac{1}{n\varepsilon^2}$$

(by chebyshev's inequality)

---

2. $$E[e^{\lambda X_1}] = e^{\lambda(-1)}\tfrac{1}{2} + e^{\lambda(1)}\tfrac{1}{2}$$

$$= \frac{e^{-\lambda} + e^{\lambda}}{2}$$

$$= \cosh(\lambda)$$

Since $X_1, \dots X_n$ are i.i.d,

$$f_{X_1, X_2, \dots X_n}(x_1, x_2, \dots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

Let $Y_i = e^{\lambda X_i}$ for $i = 1, 2, \dots, n$

Since Y_i only depends on X_i which are i.i.d, Y_i is also i.i.d.

Thus, $$f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = \prod_{i=1}^{n} f_{Y_i}(y_i)$$

3. $E[e^{\lambda \frac{S_n}{n}}]$

$$= E[e^{\lambda \frac{X_1 + \dots + X_n}{n}}]$$

$$= E[e^{\frac{\lambda X_1}{n}} e^{\frac{\lambda X_2}{n}} \dots e^{\frac{\lambda X_n}{n}}]$$

$$= E[e^{\frac{\lambda}{n}X_1}] E[e^{\frac{\lambda}{n}X_2}] \dots E[e^{\frac{\lambda}{n}X_n}] \quad (e^{\frac{\lambda}{n}X_1}, \dots, e^{\frac{\lambda}{n}X_n} \text{ jointly independent})$$

$$= \cosh(\tfrac{\lambda}{n}) \dots \cosh(\tfrac{\lambda}{n})$$

$$= \frac{\cosh^n(\frac{\lambda}{n})}{n}$$

We know $\cosh(\frac{\lambda}{n}) \leq e^{(\frac{\lambda}{n})^2/2}$

$$\cosh(\tfrac{\lambda}{n}) \leq e^{\frac{\lambda^2}{2n^2}}$$

$$\cosh^n(\tfrac{\lambda}{n}) \leq e^{\frac{\lambda^2 n}{2n^2}} = e^{\frac{\lambda^2}{2n}}$$

$$\therefore E[e^{\frac{\lambda S_n}{n}}] \leq e^{\frac{\lambda^2}{2n}}$$

4.

$$P(\frac{S_n}{n} \geq \varepsilon) = P(\frac{\lambda S_n}{n} \geq \lambda\varepsilon) = P(e^{\frac{\lambda S_n}{n}} \geq e^{\lambda\varepsilon}) \leq \frac{E[e^{\frac{\lambda S_n}{n}}]}{e^{\lambda\varepsilon}} \leq \frac{e^{\frac{\lambda^2}{2n}}}{e^{\lambda\varepsilon}} = e^{\frac{\lambda^2}{2n} - \lambda\varepsilon}$$

(by Markov's inequality)

$$P(|\tfrac{S_n}{n}| \geq \varepsilon) = P((\tfrac{S_n}{n} \geq \varepsilon) \cup (-\tfrac{S_n}{n} \geq \varepsilon))$$

$$= P(\tfrac{S_n}{n} \geq \varepsilon) + P(-\tfrac{S_n}{n} \geq \varepsilon) \quad \text{(sum rule)}$$

Since S_n and $-S_n$ have same distributions,
by Markov's inequality,

$$P(|\tfrac{S_n}{n}| \geq \varepsilon) \leq e^{\frac{\lambda^2}{2n} - \lambda\varepsilon} + e^{\frac{\lambda^2}{2n} - \lambda\varepsilon}$$

$$= 2e^{\frac{\lambda^2}{2n} - \lambda\varepsilon}$$

We can find tighter bound by minimizing the exponent,

$$\frac{d}{d\lambda}(\frac{\lambda^2}{2n} - \lambda\varepsilon) = \frac{\lambda}{n} - \varepsilon = 0$$

$$\frac{\lambda}{n} = \varepsilon$$

$$\lambda = n\varepsilon$$

$$\therefore P(|\tfrac{S_n}{n}| \geq \varepsilon) \leq 2e^{\frac{\lambda^2}{2n} - \lambda\varepsilon} = 2e^{\frac{n^2\varepsilon^2}{2n} - (n\varepsilon)\varepsilon} = 2e^{\frac{n\varepsilon^2 - 2n\varepsilon^2}{2}} = 2e^{-\frac{n\varepsilon^2}{2}}$$