1. [True or False] For parts (a) and (b), $X$ is a random variable with $E[X] = \mu$ and $\text{Var}(X) = \sigma^2$.

(a) □ $\mathbb{P}(X \geq 2\mu) \leq (\frac{\sigma}{\mu})^2$.

(b) □ Chebyshev’s inequality tells us that $\mathbb{P}(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}$. From this we can conclude that $\mathbb{P}(X - \mu \geq \alpha) \leq \frac{1}{2} \cdot \frac{\sigma^2}{\alpha^2}$.

(c) □ If we toss $n$ balls into $k \geq 3$ bins, the number of balls that land in the first 3 bins is distributed according to a binomial distribution.

Solution:

(a) True. Using Chebyshev’s inequality, $\mathbb{P}(X \geq 2\mu) \leq \mathbb{P}(|X - \mu| \geq \mu) \leq \frac{\sigma^2}{\mu^2} = (\frac{\sigma}{\mu})^2$.

(b) False. This requires $X$ to be symmetric about its mean. The probability of both the tails summed is bounded by Chebyshev’s, but we cannot assume anything about the "weight" of probability within each tail, and whether it is even.

(c) True. Each ball independently has a $\frac{3}{k}$ chance of landing within the first 3 bins. If $X$ is the random variable for number of balls in the first 3 bins, we have $X \sim \text{Binom}(n, \frac{3}{k})$.

2. [Short Answer] Let $Z$ be a random variable which takes values less than 4. We know that $E[Z] = 2$.

(a) Provide an upper bound on $\mathbb{P}(Z \leq 0)$.

(b) What is a requirement on the variance of $Z$ such that we can guarantee $\mathbb{P}(0 < Z < 4) \geq \frac{1}{2}$?

Solution:

(a) Apply Markov’s inequality to $4 - Z \geq 0$, so

$$\mathbb{P}(Z \leq 0) = \mathbb{P}(4 - Z \geq 4) \leq \frac{\mathbb{E}[4 - Z]}{4} = \frac{1}{2}$$

(b) By Chebyshev’s inequality, we have:

$$\mathbb{P}(0 < Z < 4) = 1 - \mathbb{P}(|Z - 2| \geq 2) \geq 1 - \frac{\text{Var}(Z)}{2^2}$$

So we need to set the variance such that $1 - \frac{\text{Var}(Z)}{4} \geq \frac{1}{2}$. After some manipulation we see that $\text{Var}(Z) \leq 2$ suffices.
3. **[Long Answer]** Let $X_i$, where $1 \leq i \leq n$, be i.i.d. with mean $\mu$ and variance $\sigma^2$. Let also $A_n = (X_1 + \cdots + X_n)/n$. Using Chebyshev’s inequality, find a 90% confidence interval for $\mu$.

**Solution:**

By linearity of expectation, we have:

$$
\mathbb{E}[A_n] = \mathbb{E}[(X_1 + \cdots + X_n)/n] = (\mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n])/n = n\mu/n = \mu
$$

We also have:

$$
\text{Var}(A_n) = \text{Var}((X_1 + \cdots + X_n)/n) = (\text{Var}[X_1] + \cdots + \text{Var}[X_n])/n^2 = n\sigma^2/n = \sigma^2/n
$$

So by Chebyshev’s, we know that:

$$
\mathbb{P}(|A_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}
$$

We choose $\varepsilon$ so that the right-hand side is 0.1. This gives:

$$
\frac{\sigma^2}{n\varepsilon^2} = 0.1
$$

$$
\varepsilon^2 = 10 \cdot \frac{\sigma^2}{n}
$$

$$
\varepsilon = \sqrt{10} \cdot \frac{\sigma}{\sqrt{n}}
$$

The 90% confidence interval is then $A_n \pm \sqrt{10} \cdot \frac{\sigma}{\sqrt{n}}$. 