1. Markov's Inequality

a. For a **non-negative** random variable X, $P(X \geq a) \leq \frac{E(X)}{a}$

b. Proof:
$E(X) = \sum_x xP(X = x)$
$\geq \sum_{0<x<a} 0 \cdot P(X = x) + \sum_{x \geq a} aP(X = x)$
$= a \sum_{x \geq a} P(X = x)$
$= aP(X \geq a)$

2. Chebychev's Inequality

a. For **any** random variable X, $P(\mid X - \mu \mid \geq a) \leq \frac{Var(X)}{a^2}$

b. Proof:
Let $Y = (X - \mu)^2$.
$E(Y) = E((X - \mu)^2) = Var(x)$
$\mid X - \mu \mid \geq a \implies \mid X - \mu \mid^2 \geq a^2 \implies Y \geq a^2$
$P(\mid X - \mu \mid \geq a) = P(Y \geq a^2) \leq \frac{E(Y)}{a^2} = \frac{Var(X)}{a^2}$

3. Weak Law of Large Numbers(WLLN)

a. $X_1, \cdots, X_n$ are i.i.d. random variables with $E(X_i) = \mu$.
Then for $\epsilon > 0$,
$P(\mid \frac{X_1 + \cdots X_n}{n} - \mu \mid \geq \epsilon) \to 0$, as $n \to \infty$.

b. Informally, WLLN states that as we have more and more samples, average approaches to expectation.

c. Proof:
Let $X = \frac{X_1 + \cdots X_n}{n}$.
$P(\mid \frac{X_1 + \cdots X_n}{n} - \mu \mid \geq \epsilon) = P(\mid X - \mu \mid \geq \epsilon)$
$\leq \frac{Var(X)}{\epsilon^2} = \frac{Var(X_1 + \cdots + X_n)}{n^2} \cdot \frac{1}{\epsilon^2} = \frac{n \cdot Var(X_i)}{n^2 \epsilon^2}$
$= \frac{Var(X_i)}{n\epsilon^2} \to 0$, as $n \to \infty$.

4. Central Limit Theorem (CLT)

a. $X_1, \cdots, X_n$ are i.i.d. random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$.
Define $S_n = \frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}}$. Then,
$S_n \to N(0, 1)$, as $n \to \infty$.

b. Informally, CLT states that as we have more and more samples, sum converges to normal distribution.

c. 95% Confidence Interval (estimate mean):
Let $A_n = \frac{X_1 + \cdots + X_n}{n}$. Then $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]$ is a 95%-CI for μ.
(Take a look at normal curve.)