Fundamental limit theorem

Markov Inequality:

Suppose R.V. $X$ takes on a non-negative value.
Then: $P(X > a) \leq \frac{E(X)}{a}$

Proof:
Define $Y_a = \begin{cases} a & X \leq a \\ X & X > a \end{cases}$

Diagram:

\[ 0 \quad \frac{a}{a} \quad X \]

\[ 0 \quad Y_a \quad a \]
\[ y_a \leq x \implies E[y_a] \leq E(x) \]

Compute:
\[ E(y_a) = a \cdot P[y_a = a] + 0 \cdot P[y_a = 0] \]
\[ = a \cdot P[y_a = a] = a \cdot P(x \geq a) \]
\[ \implies P(x \geq a) \leq \frac{E(x)}{a} \]

\text{Markov}

\text{Ex. R.V. X, uniform \([0, 4]\) \& \(E(X) = 2\)}

\text{Markov Inequality:} \quad a = 2
\[ P(X \geq 2) \leq \frac{2}{2} = 1 \quad \implies \text{useless} \]
\[
\begin{align*}
&\alpha = 3 \
&\Rightarrow \
&\begin{cases} 
\alpha = 4 \\
\alpha = 0 
\end{cases} 
\end{align*}
\]

\[p(x \geq 3) = \frac{3}{2} \leq \frac{3}{2} \]

\[p(x \geq 4) \leq \frac{3}{2}\]

\[p(x \geq 2) = \frac{1}{2} \]

\[p(x > 3) = \frac{1}{4}\]

\[p(x > 4) = 0\]

Proof.

Chernoff Inequality
$X$ is R.V. mean $\mu$, variance $\sigma^2$

$$P(\lvert X - \mu \rvert > C) \leq \frac{\sigma^2}{C^2} \quad \forall C > 0$$

**Proof:** Consider $(X - \mu)^2$ as a positive R.V.

Apply Markov inequality with $a = C^2$

$$P((X - \mu)^2 > C^2) \leq \frac{E[(X - \mu)^2]}{C^2} = \frac{\sigma^2}{C^2}$$

**Note:** event $(X - \mu)^2 > C^2$ is identical to $\lvert X - \mu \rvert > C$
\[ P \left( |X - \mu| \geq C \right) \leq \frac{\sigma^2}{C^2} \quad \forall C > 0 \]

**Chebyshev Inequality.**

Let \( C = k \sigma \) \( k \) positive

\[ P \left( |X - \mu| \geq k \sigma \right) \leq \frac{\sigma^2}{k^2 \sigma^2} = \frac{1}{k^2} \]

Words: prob that a R.V. taken is more than \( k \) standard deviation away from mean is at most \( \frac{1}{k^2} \)
Ex R.V. uniform $[0, 4]$  \[ E(x) = 2, \quad \sigma^2 = \frac{4}{3} \]

\[ P(|x - 2| \geq 1) \leq \frac{4}{3} \quad \rightarrow \text{useless.} \]

Ex \hspace{1cm} \text{X in exponential at} \hspace{1cm} \text{KU.} \hspace{1cm} \lambda = 1

$E(X) = Var(X) = 1$. \hspace{1cm} \text{c} \gg 1

\[ P(X > C) = P(X - 1 > C - 1) \leq \frac{6^2}{(C-1)^2 (C-3)^3} \]

\[ P(X > C) \leq \frac{1}{Cc(C-1)^2} \quad \leftarrow \text{Chebyshev.} \]

\[ P(X > c) = e^{-c} \quad \leftarrow \text{loose bound} \]

\[ \text{using exponential} \]

\[ \text{P}(X > c) = e^{-c} \]
Weak Law of Large Numbers

Consider $X_1, X_2, \ldots$ i.i.d. RVs, mean $\mu$, var $\sigma^2$

Define $M_n = \frac{X_1 + X_2 + \ldots + X_n}{n}$

$\text{Sample mean}$

$E[M_n] = \frac{E[X_1] + E[X_2] + \ldots + E[X_n]}{n} = \frac{n\mu}{n} = \mu$

$\text{Var}[M_n] = \frac{n\sigma^2}{n} = \frac{\sigma^2}{n}$

Apply Chebyshev's To $M_n$: 

\[ P(\left| M_n - \mu \right| \geq \varepsilon) \leq \frac{\delta^2}{n\varepsilon^2} \quad \forall \varepsilon > 0 \]

\text{lim as } n \to \infty \text{ of both sides.}

\[ P(\left| M_n - \mu \right| \geq \varepsilon) \to 0 \quad \text{as } n \to \infty \]

\text{Weak law of large numbers.}

**EX** Event A defined in terms of probabilistic experiment.

\[ P(A) = \begin{cases} \rho & \text{if } \text{independent repetition of the experiment, } \text{ith repetition of the experiment } X_i \text{ event } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases} \]
\[ M_n = \frac{X_1 + X_2 + \ldots + X_n}{n} \quad \text{measured quantity} \]

\[ E[X_i] = 1 \cdot P(A) + 0 \cdot P(\bar{A}) = P(A) = p. \]

Apply Weak law of large numbers:

\[ P\left(\left|M_n - p\right| \geq \epsilon\right) \to 0 \quad \text{as} \quad n \to \infty \quad \forall \epsilon > 0 \]

Empirical frequency is a good way to estimate \( p \).
Ex polling: p = probability of supporting Kerry. 
interview n randomly selected person 

Record Mn.

Reply of the n persons is an independent Bernoulli R.V. \( X_i \) with success prob.

of \( P \): ????

Apply Chebyshev:

\[
P(\left| M_n - P \right| \geq \varepsilon) \leq \frac{P(1-P)}{\varepsilon^2} \]

note: \( P(1-P) \leq \frac{1}{4} \)
\[ P \left( \left| M_n - \mu \right| > \varepsilon \right) \leq \frac{1}{4n\varepsilon^2} \]

\[ \varepsilon = 0.1, \quad n = 100 \]

\[ P \left( \left| M_{100} - \mu \right| > 0.1 \right) \leq \frac{1}{4} \]

It is shown that our \( M_{100} \) estimate of \( \mu \) is off by more than 10\% is less than \( \frac{1}{4} \).
3. our estimate needs to be within 1% of \( \mu \) with prob of at least 95%.

\[
P\left( |M_n - \mu| > 0.01 \right) \leq 5%
\]

\[
P\left( |M_n - \mu| > 0.01 \right) \leq \frac{1}{4n(0.01)^2} \leq 5%
\]

\[
\iff n \geq 50,000
\]