Chebyshev's Inequality

Recall Markov's inequality: if $f \geq 0$, $\varepsilon, p > 0$, then

$$\mu\{f \geq \varepsilon\} \leq \frac{1}{\varepsilon^p} \int f^p \, d\mu$$

Suppose μ is a probabilty measure, $X \in L^2$. Set $p = 2$, and apply Markov's inequality to $f = |\mathring{X}| = |X - \mathbb{E}[X]|$

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq \varepsilon) = \mathbb{P}(|\mathring{X}| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \mathbb{E}[|\mathring{X}|^2] = \frac{1}{\varepsilon^2} \operatorname{Var}(X)$$

This is Chebyshev's inequality.

Alternative form: let $\sigma(X) = \sqrt{\operatorname{Var}(X)}$ standard deviation

set $\varepsilon = k\sigma(X)$.

$$\therefore \mathbb{P}(|X - \mathbb{E}[X]| \geq k\sigma(X)) \leq \frac{\operatorname{Var}(X)}{(k\sigma(X))^2} = \frac{1}{k^2}.$$

What is $\mathbb{E}$, Really?

Eg. Toss a fair coin n times.

$$\Omega = \{\omega = (\omega_1, \cdots, \omega_n),\ \omega_j \in \{0,1\}\}$$

$$\mathcal{F} = 2^{\Omega} \qquad \mathbb{P}\{\omega\} = 2^{-n} \text{ for all } \omega \in \Omega.$$

$$X_j(\omega) = \omega_j \qquad \mathbb{E}[X_j] = \sum_{\omega \in \Omega} \omega_j 2^{-n} \underbrace{\begin{cases} \omega_j = 0 & \text{in } 2^{n-1} \text{ cases} \\ \omega_j = 1 & \text{in } 2^{n-1} \text{ cases} \end{cases}} = \frac{1}{2}.$$

How many Heads come up?

$$S_n = X_1 + X_2 + \cdots + X_n \qquad \mathbb{E}[S_n] = \sum_{j=1}^{n} \mathbb{E}[X_j] = \frac{n}{2}$$

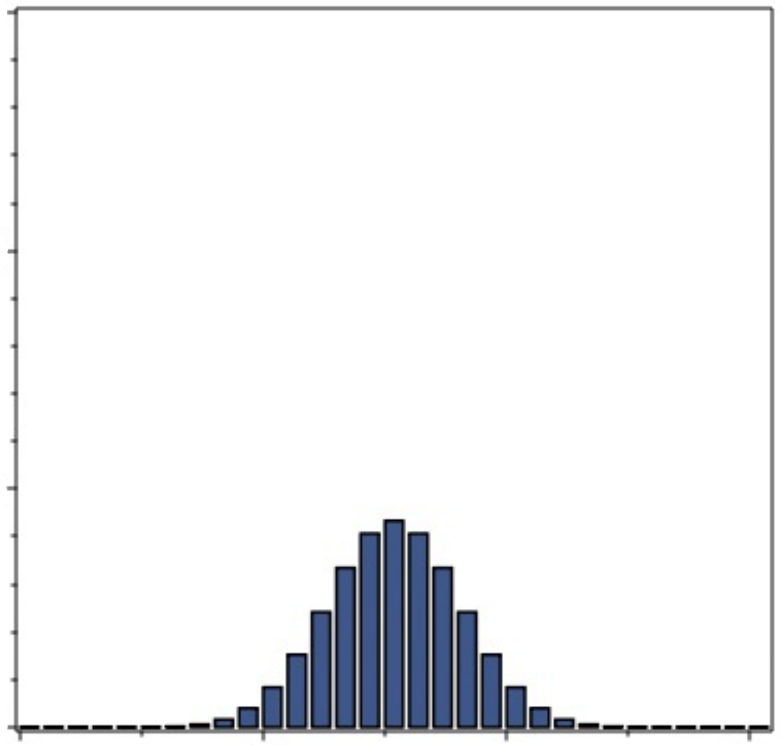
i.e. $S_n(\omega) = \sum_{j=1}^{n} \omega_j \in \{0, 1, 2, \cdots, n\}$

$$\left.\mathbb{P}(S_n = k) = \frac{1}{2^n} \#\{\omega : \omega_1 + \cdots + \omega_n = k\} = \frac{1}{2^n}\binom{n}{k}\right\} \quad S_n \overset{d}{=} \mathrm{Binom}(n, \tfrac{1}{2})$$

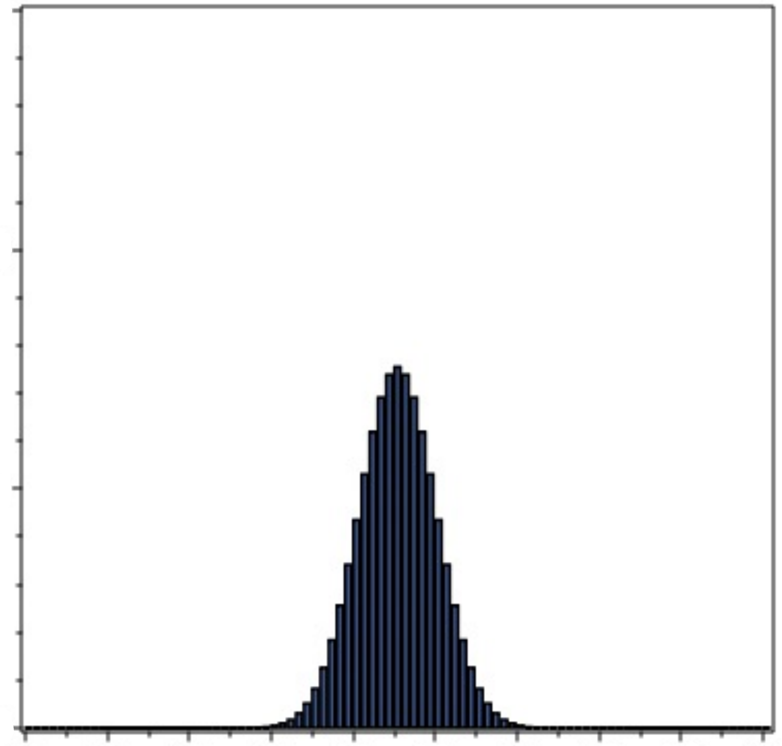
What about the proportion of Heads?

$$\mathbb{E}\left(\frac{S_n}{n}\right) = \frac{1}{2} \qquad \mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \leq \frac{\mathrm{Var}\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{\mathrm{Var}(S_n)}{n^2 \varepsilon^2} \overset{\uparrow}{=} \frac{1}{4n\varepsilon^2} = O\left(\frac{1}{n}\right)$$

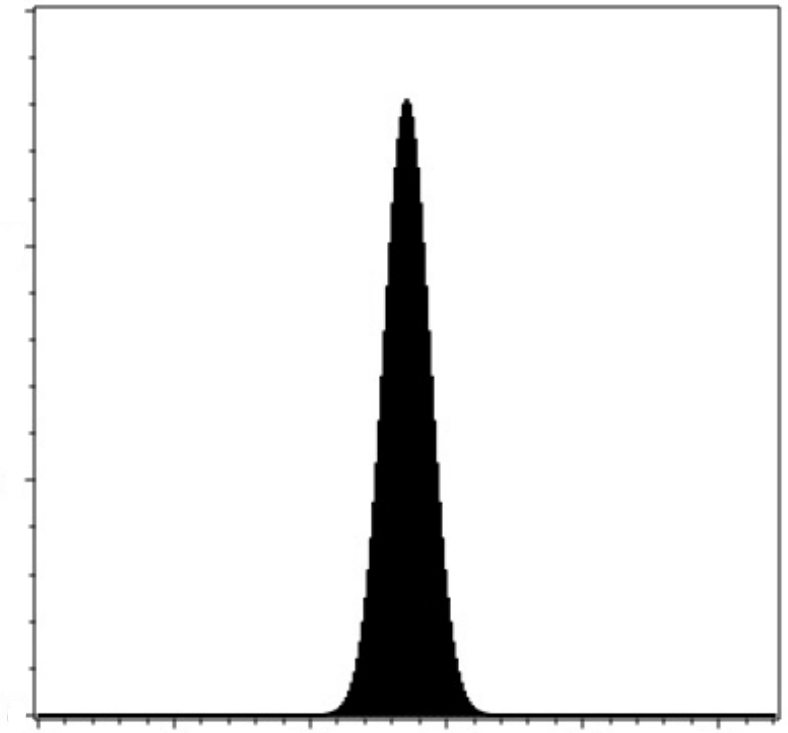
$$\mathrm{Var}(S_n) = \sum_{k=0}^{n} k^2 \frac{1}{2^n}\binom{n}{k} - \left(\frac{n}{2}\right)^2 = \frac{n^2+n}{4} - \left(\frac{n}{2}\right)^2 = \frac{n}{4}.$$



$\frac{S_{30}}{30}$



$\frac{S_{90}}{90}$



$\frac{S_{270}}{270}$

This concentration isn't really specific to $\text{Binom}(\frac{1}{2}, n)$ distributions. It's because of the normalization S_n / n and, crucially,

$$S_n = \underbrace{X_1 + X_2 + \cdots + X_n}_{\text{uncorrelated.}}$$

The Weak Law of Large Numbers

Theorem. Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of L^2 random variables on a probability space, that are pairwise uncorrelated:

$$\mathrm{Cov}(X_n, X_m) = 0 \quad \text{if } n \neq m$$

and all with the same mean and variance:

$$\mathbb{E}[X_n] = \alpha \,, \quad \mathrm{Var}[X_n] = t \quad \forall n.$$

Let $S_n = X_1 + \cdots + X_n$. Then for any $\varepsilon > 0$,

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \alpha \right| \geq \varepsilon \right) \leq \frac{\mathrm{Var}(S_n/n)}{\varepsilon^2} = \frac{\mathrm{Var}(S_n)}{n^2 \varepsilon^2} = \frac{n \cdot t}{n^2 \varepsilon^2} = \frac{t}{n\varepsilon^2} = O\left(\frac{1}{n}\right)$$

Pf. By assump. $\mathbb{E}[X_n] = \alpha$ $\therefore$ $\mathbb{E}[S_n] = n \cdot \alpha$. $\therefore \mathbb{E}\left[\frac{S_n}{n}\right] = \alpha$.

$$\mathrm{Var}(S_n) = \underbrace{\mathrm{Var}(X_1 + \cdots + X_n)}_{\text{uncor.}} = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) = n\,t.$$

///

What kind of limit is this?

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \alpha\right| \geq \varepsilon\right) = 0, \quad \forall \varepsilon > 0.$$

It means $\frac{S_n}{n}$ is asymptotically concentrated at α.

But does it mean $\frac{S_n}{n} \to \alpha$ a.s.?

Eg. Suppose $X_1, X_2, X_3, \dots$ are coin tosses, with $\mathbb{P}(X_n = 1) = \frac{1}{n}$

$\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$.

$$\mathbb{P}(|X_n| \geq \varepsilon) = \mathbb{P}(X_n \geq \varepsilon \text{ or } \cancel{X_n \leq -\varepsilon})$$
$$= \mathbb{P}(X_n = 1) = \frac{1}{n} \longrightarrow 0.$$

Does $X_n \to 0$ a.s.?

Borel-Cantelli (II): $\mathbb{P}(X_n = 1 \text{ i.o.}) = 1$.