Limit Theorems

↳ The most powerful and useful theorems in statistics / Probability.

IID Random Sample.

$$X_1, X_2, \ldots, X_n$$

↳ independent, identically distributed.

"X_i are samples of some unknown RV X"

$$\mathbb{E}X_i = \mu \quad , \quad \mathrm{Var}(X_i) = \sigma^2$$

↳ Randomly sampling from a distribution.

↳ or Taking a bunch of independent measurements.

<u>Def</u> (Sample mean): Given iid random sample $X_1, X_2, \ldots, X_n$, the sample mean is.

$$\overline{X}_n = \overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

- The sample mean $\bar{X}$ is another RV.

- $$\mathbb{E}\bar{X} = \mathbb{E}\left[\frac{X_1 + X_2 + \dots + X_n}{n}\right]$$ linearity of $\mathbb{E}$.

$$= \frac{1}{n}\left[\underbrace{\mathbb{E}X_1}_{=\mu} + \underbrace{\mathbb{E}X_2}_{=\mu} + \dots + \underbrace{\mathbb{E}X_n}_{=\mu}\right] = \frac{\cancel{n}\mu}{\cancel{n}}$$

$$= \mu .$$

- $$\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$ $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$

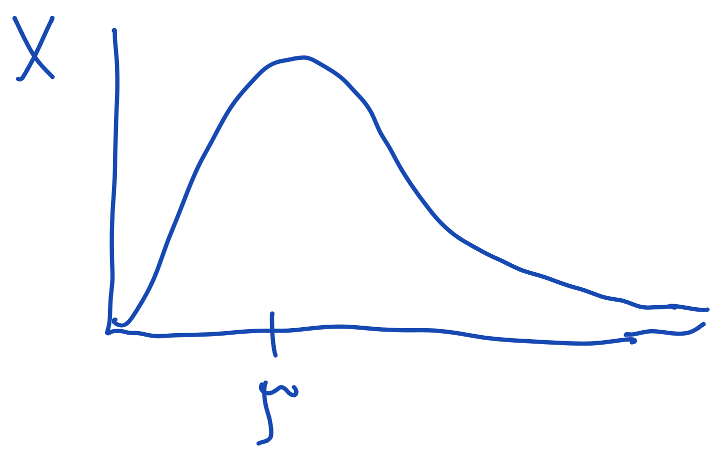
$$= \frac{1}{n^2}\mathrm{Var}(X_1 + X_2 + \dots + X_n)$$ independence

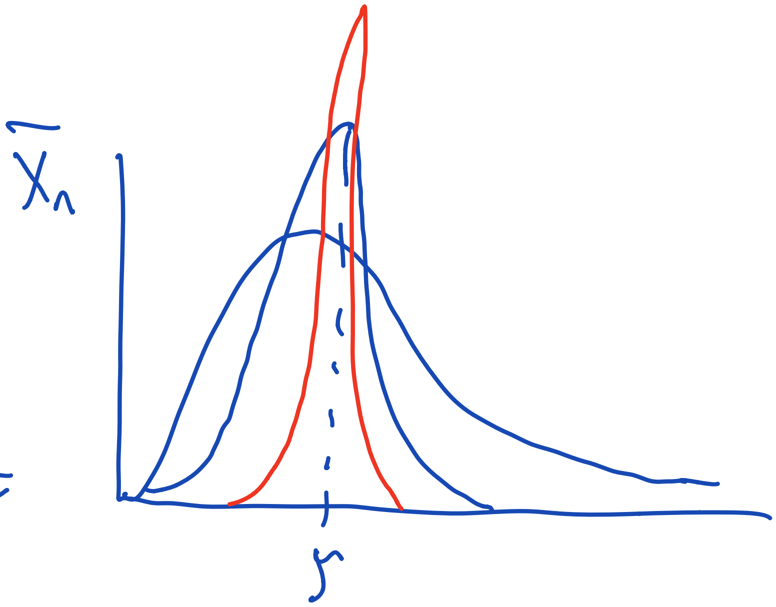
$$= \frac{1}{n^2}\left[\underbrace{\mathrm{Var}(X_1)}_{=\sigma^2} + \underbrace{\mathrm{Var}(X_2)}_{=\sigma^2} + \dots + \underbrace{\mathrm{Var}(X_n)}_{=\sigma^2}\right]$$

$$= \frac{\cancel{n}\sigma^2}{n^{\cancel{2}}} = \frac{\sigma^2}{n}$$

$$\boxed{\mathrm{Var}(\bar{X}_n) = \frac{\sigma^2}{n}}$$

Pictures





The weak law of large numbers (WLLN)

Theorem $X_1, X_2, \ldots, X_n$ iid. with finite

$$\mu = \mathbb{E} X_i \quad , \quad \sigma^2 = \mathrm{Var}(X_i)$$

then for each $\varepsilon > 0$

Convergence in probability

$$\lim_{n \to \infty} P\left(|\bar{X}_n - \mu| > \varepsilon \right) = 0.$$

the probability that $\bar{X}_n$ deviates by more than ε from the mean μ.

" $\bar{X}_n \longrightarrow \mu$ in probability $n \to \infty$. "

random deterministic.

Proof Chebyshev. "$P(|X-\mu_X| > b) \le \frac{Var(X)}{b^2}$"

$$P(|\bar{X}_n - \mu| > \varepsilon) \le \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

$$\to 0 \text{ as } n \to \infty.$$

Example

$$X \sim \text{Binomial}(n, p)$$

$$\frac{X}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \quad , \quad X_i \sim \text{Bernoulli}(p).$$

$$P\left(\left|\frac{X}{n} - p\right| > \varepsilon\right) \le \frac{pq}{n\varepsilon^2}$$

$$\frac{X}{n} \to p \text{ in probability.}$$

or

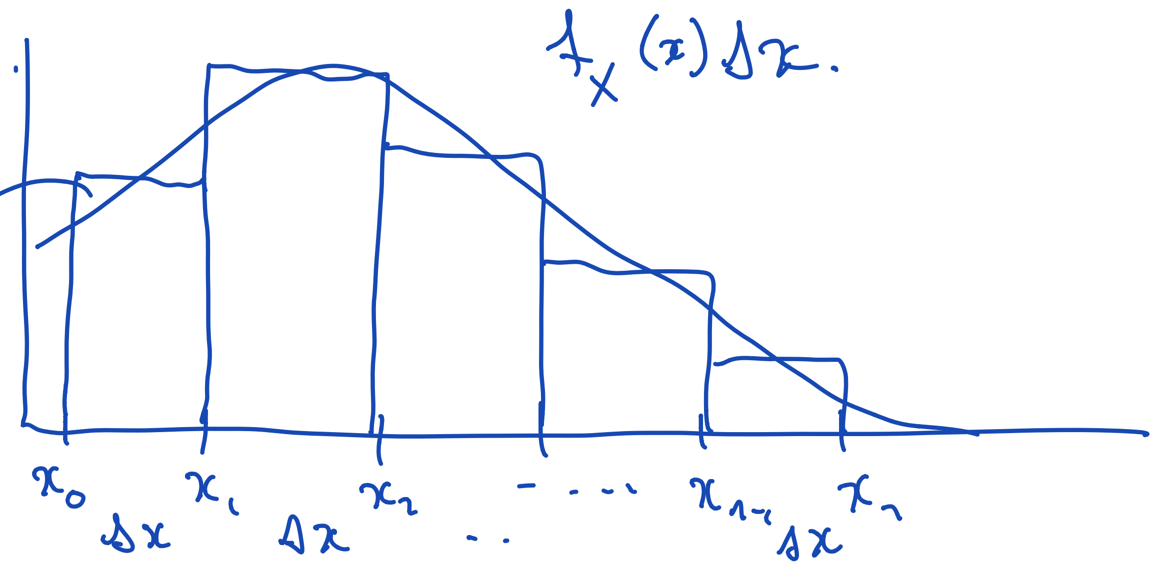
true bias.

$$\frac{\#\text{ of heads}}{\#\text{ trials}} \approx p$$

Ex Histogram.

Sample from a PDF f_X.

$X_1, X_2, \ldots, X_n$



$$\frac{\#\ X_i \text{ that lie between } [x_0, x_1]}{n}$$

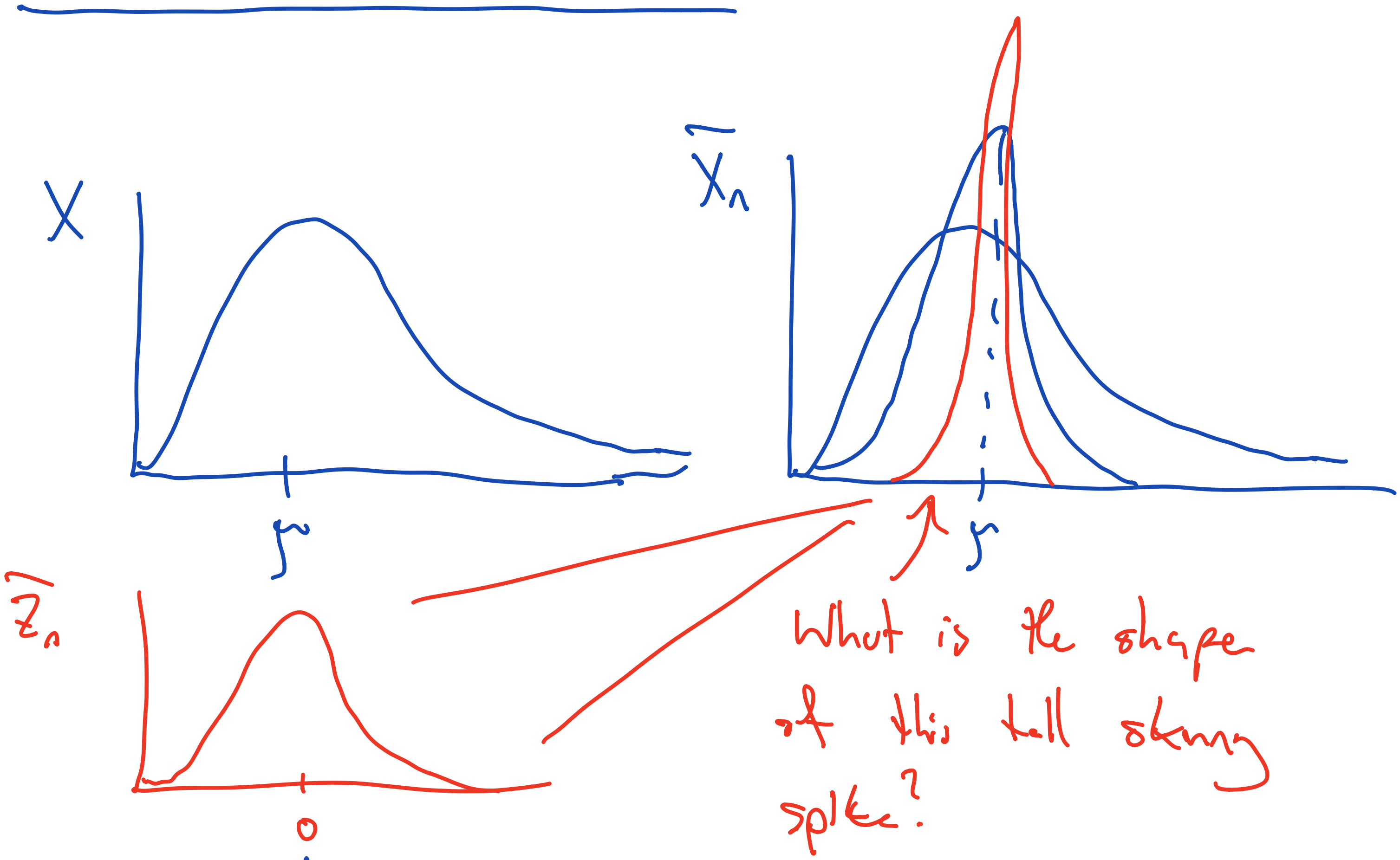
$$= \frac{1}{n}\sum_{i=1}^{n} Y_i \quad , \quad Y_i = \begin{cases} 1 & X_i \in [x_0, x_1] \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E} Y_i = P(x_0 \le X \le x_1)$$

By WLLN $\frac{1}{n}\sum_{i=1}^{n} Y_i \longrightarrow P(x_0 \le X \le x_1)$

in probability.

Central Limit Theorem



Lets standardize (normalize $\bar{X}_n$).

$$\bar{Z}_n = \frac{\bar{X}_n - \mathbb{E}\bar{X}_n}{\sqrt{Var(\bar{X}_n)}} = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}$$

Theorem (CLT)

$$Z_n \xrightarrow{d} Z \sim N(0,1)$$

Convergence in distribution.

$\Updownarrow$

Convergence of CDFs.

i.e. $$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z).$$

for all $z \in \mathbb{R}$.

This is independent of the starting distribution.