μ : "population mean"
$\overline{X}_n$: "sample mean"
As $n \to \infty$, $\overline{X}_n \to \mu$ (Law of Large Numbers)
Central Limit Theorem:
For large n, $\overline{X}_n \approx N(\mu, \sigma^2/n)$

Maximum Likelihood Estimators

Parameter estimation: Given independent samples $x_1, x_2, \cdots, x_n$ from a distribution $f(x|\theta)$, estimate θ.

Ex: Given samples HHTHH of independent flips of a coin, estimate $\theta = P(\text{heads})$.

$P(x|\theta)$: if x is unknown and θ is known, this is a conditional probability.
In parameter estimation, x is known and θ is unknown: $P(x|\theta)$ is called a likelihood and often written $L(x|\theta)$.

Maximum Likelihood estimation:
What value of θ maximizes $L(x_1, x_2, \cdots, x_n|\theta)$
$$= \prod_{i=1}^{n} f(x_i|\theta) \quad ?$$

Approach: $\frac{\partial}{\partial\theta} L(\vec{x}|\theta) = 0$ and solve for $\theta = \hat{\theta}$.
(and check we have a max rather than a min.)
Or: $\frac{\partial}{\partial\theta} \ln L(\vec{x}|\theta) = 0$: ln is an increasing function

Ex: n independent flips $x_1, x_2, \ldots, x_n$ of a coin that has probability θ of heads, yielding n_0 tails and n_1 heads, with $n_0 + n_1 = n$. Find MLE $\hat{\theta}$ of θ.

$$L(x_1, x_2, \ldots, x_n \mid \theta) = (1-\theta)^{n_0} \theta^{n_1} \qquad \left[\binom{n}{n_1} ?\right]$$

$$\ln L(x_1, x_2, \ldots, x_n \mid \theta) = n_0 \ln(1-\theta) + n_1 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln L(x_1, x_2, \ldots, x_n \mid \theta) = -\frac{n_0}{1-\theta} + \frac{n_1}{\theta} = 0$$

$$-n_0 \hat{\theta} + n_1 (1 - \hat{\theta}) = 0$$

$$n_1 = n_0 \hat{\theta} + n_1 \hat{\theta}$$

$$\hat{\theta} = \frac{n_1}{n_0 + n_1} = \boxed{\frac{n_1}{n}}$$

Is $\hat{\theta}$ a maximum?

$$\frac{\partial^2}{\partial \theta^2} \ln L(x_1, x_2, \ldots, x_n \mid \theta) = -\frac{n_0}{(1-\theta)^2} - \frac{n_1}{\theta^2} < 0$$

so $\ln L(x_1, x_2, \ldots, x_n \mid \theta)$ is concave downward everywhere, so $\hat{\theta}$ is a maximum.

Ex: $x_1, x_2, \ldots, x_n$ from $N(\theta_1, \theta_2)$.