Maximum Likelihood Estimation

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

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

$P(x|\theta)$

Viewed as a function of x (θ fixed), it's a probability.
Viewed as a function of θ (x fixed), it's called a likelihood, and often written $L(x|\theta)$.

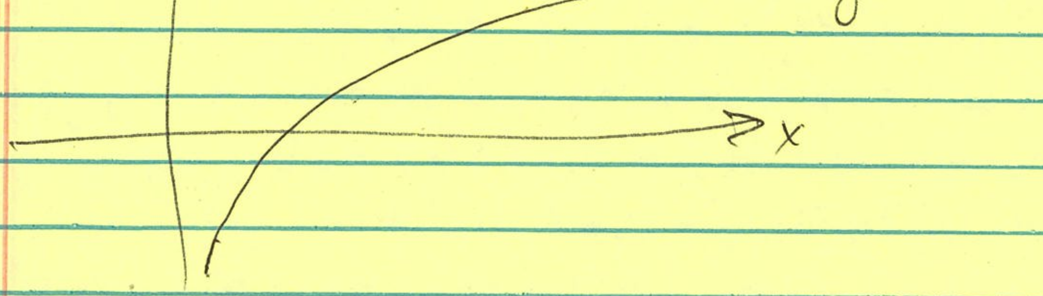
Ex: What value of θ maximizes $L(x_1, x_2, \dots, x_n|\theta)$

$$= \prod_{i=1}^{n} f(x_i|\theta) \,?$$

Approach: $\frac{\partial}{\partial\theta} L(x_1, \dots, x_n|\theta) = 0$ and solve for θ.

To avoid the mess of $\frac{\partial}{\partial\theta}$ of a product:

$$\frac{\partial}{\partial\theta} \ln L(x_1, \dots, x_n|\theta) = 0 \text{ and solve for } \theta.$$



the value of θ that maximizes L also maximizes $\ln L$, because $\ln$ is monotonically increasing.

Ex: $\theta = P(\text{heads})$, given n independent flips $x_1, x_2, \ldots, x_n$, yielding n_0 tails and n_1 heads, with $n_0 + n_1 = n$.

$$L(x_1, x_2, \ldots, x_n | \theta) = (1-\theta)^{n_0}\,\theta^{n_1} \qquad (\text{mult. by } \binom{n}{n_0})?$$

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

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

Let $\hat{\theta}$ be the value of θ that satisfies this equation.

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

$$n_1 = \hat{\theta}(n_0 + n_1) = \hat{\theta}\,n$$

$$\hat{\theta} = \frac{n_1}{n}$$

To show that $\hat{\theta}$ maximizes L:

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

So $\ln L$ is concave downward everywhere, and $\hat{\theta}$ is a maximum.

Since $\hat{\theta} = \frac{n_1}{n}$, the fraction of heads in the sample is a good estimate of the probability of heads.