Maximum Likelihood Estimation

- To choose a set of parameters θ in a way that maximizes the likelihood function *L(θ)* : $L(\theta) = f(x_1, x_2, \cdots, x_n|\theta) = \prod_{i=1}^{n} f(x_i|\theta)$
  - where $x_1, x_2, \ldots, x_n$ is a set of random samples from the distribution of a random variable ***X*** with density *f* and associated parameter θ.
- The ML estimation $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k)$ is the set of estimated values that maximizes *L(θ),* or values that satisfies the simultaneous equations $\frac{\partial L(\theta)}{\partial \theta_i} = 0, \ i = 1, \cdots, k$
- Properties:
  - Maximum-likelihood estimates are (1) consistent, and (2) asymptotically efficient.
  - Let $\hat{\theta}_{ML}$ be an MLE of θ , then $g(\hat{\theta}_{ML})$ is an MLE of g(θ) , i.e., $[\hat{g}(\theta)]_{ML} = g(\hat{\theta}_{ML})$ if g(.) is a monotonic function.