Lecture 30, Mar 27, 2023

Log-Likelihood

- With distributions such as the normal, exponential, or Poisson distributions, we can make them easier to work with if we instead maximize the log of the likelihood function, since $\frac{\mathrm{d}}{\mathrm{d}x}\ln x > 0$
- Example: normal distribution, find μ, σ^2
  - $$\begin{aligned} L(x_1, \cdots, x_n) &= \prod_{i=1}^{n} n(x_i; \mu, \sigma) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2} \\ &= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2} \end{aligned}$$
  - $$\begin{aligned} \ln L &= \ln\left(\frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} e^{-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2}\right) \\ &= -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^2 \end{aligned}$$
  - $$\begin{aligned} &\frac{\partial L}{\partial \mu} = 0 \\ \implies &\frac{1}{2}\sum_{i=1}^{n} 2\left(\frac{x_i - \mu}{\sigma}\right)^2 = 0 \\ \implies &\sum_{i=1}^{n} x_i - \mu = 0 \\ \implies &\left(\sum_{i=1}^{n} x_i\right) - n\mu = 0 \\ \implies &\mu = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \end{aligned}$$
    - We just ended up getting $\bar{x}$, which makes sense
  - $$\begin{aligned} &\frac{\partial L}{\partial \sigma^2} = 0 \\ \implies &-\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 = 0 \\ \implies &\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \end{aligned}$$
    - This is an unbiased estimator if we do know the true mean μ, but if we use the optimal $\mu = \bar{x}$, this would be a biased estimator
    - In the limit as $n \to \infty$ the sample mean approaches the true mean, and the estimator becomes unbiased

Important

When we use maximum likelihood estimation for each parameter, it is assumed that we know the exact values of the other parameters; this means that in the case of multiple parameters, MLE can give a biased estimate

Note

There is often a tradeoff between bias and variance of an estimator. Sometimes it might be worth it to have an estimator that is a little biased if it significantly reduces the variance of the estimator, making it more efficient.