Likelihoods and Curvature

Jo Hardin

10/13/2020

Fisher Information

Fisher Information describes the curvature of a likelihood function, that is, the negative of the expected value of the second derivative. Why the second derivative? Let's look at some examples.

Important all the plots below / derivatives / intuition is in thinking about the likelihood as a function of θ!! (Not a function of the data.)

Exponential Distribution

$$f(x|\theta) = (1/\theta)e^{-x/\theta}$$

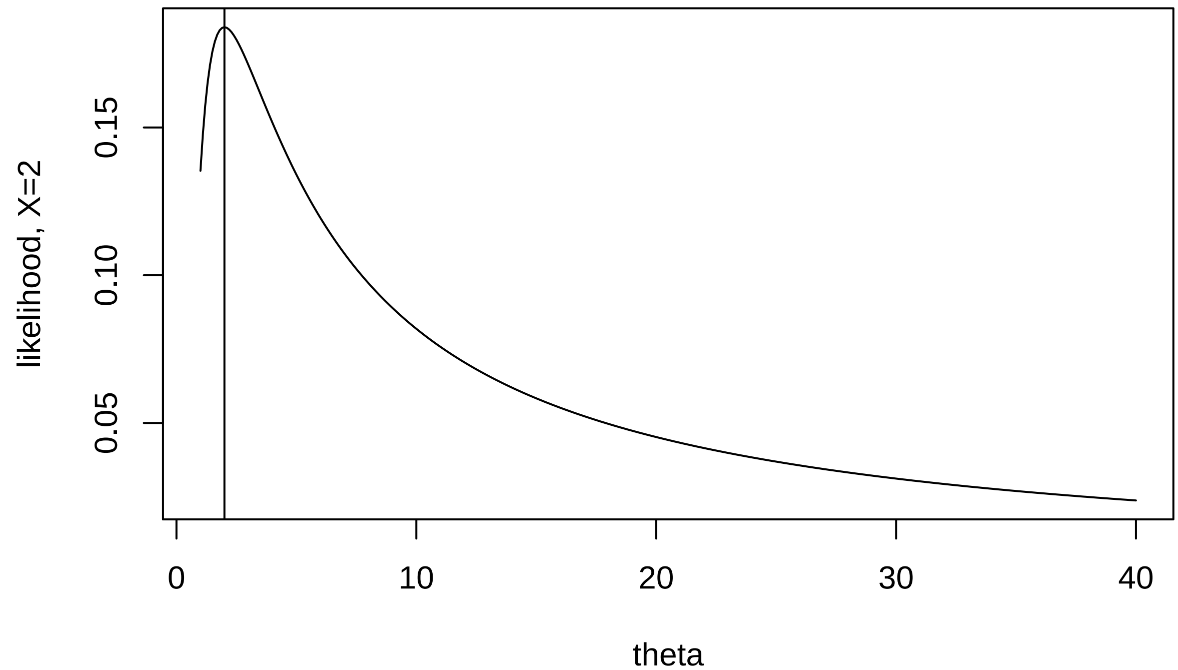
Consider a simplified example with only a single observation from an exponential distribution with mean θ.

Recall that the maximum likelihood estimator is the value of the parameter the maximizes the likelihood. That is, in plotting the likelihood, find the value of θ (on the x-axis) that gives the highest likelihood (on the y-axis).

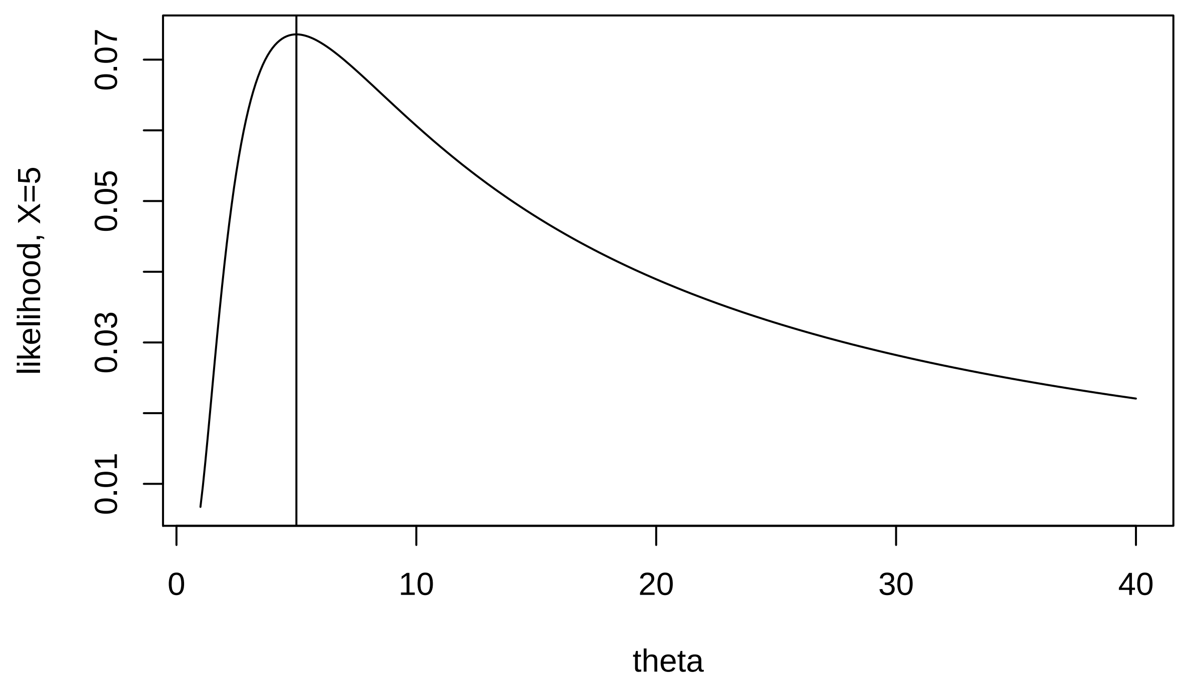
$$\text{MLE} = \frac{\sum_i X_i}{n} \quad \text{if } n = 1, \text{MLE} = X$$

The task for today is to consider how certain we are about the estimate. When $X = 2$ the likelihood is extremely peaked, and the maximum value appears somewhat obvious. When $X = 10$ it is still possible to maximize the function, but the process seems somewhat less certain to give the "best" value of θ.

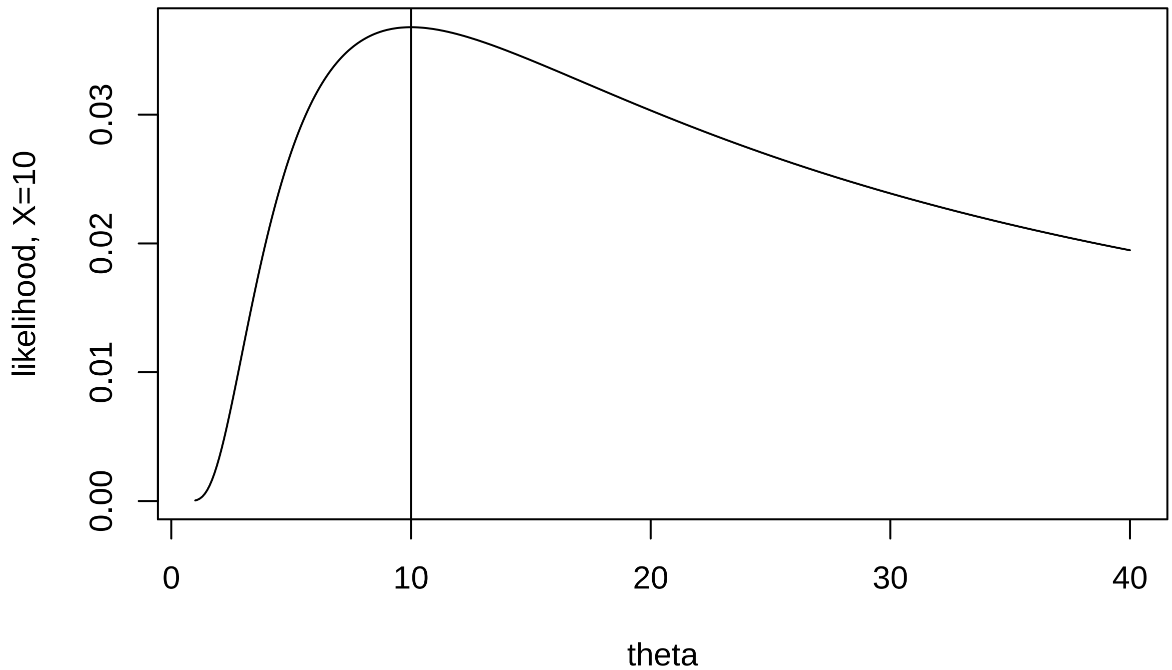
```
theta = seq(1,40, by = 0.1)

ex = 2
plot(theta, (1/theta)*exp(-ex/theta), type = "l", ylab="likelihood, X=2")
abline(v=ex)
```




```
ex = 5
plot(theta, (1/theta)*exp(-ex/theta), type = "l", ylab="likelihood, X=5")
abline(v=ex)
```




```
ex = 10
plot(theta, (1/theta)*exp(-ex/theta), type = "l", ylab="likelihood, X=10")
abline(v=ex)
```



The information itself is given by the expected value of the second derivative. For each value of X, there is a particular shape to the second derivative (some more peaked, some less peaked). The expected value will average over the values of the second derivative, giving weights as defined by the pdf (because that's what expected value does).