Newton Raphson for Optimization, Attempt 3

Evan L. Ray

March 2, 2018

Exponential Model Example

Model: $X_1, \ldots, X_n \sim \mathrm{Exp}(\lambda)$

Example: X_i is a waiting time in minutes for individual i who goes to the emergency room. λ is the number of patients they see per minute, on average.

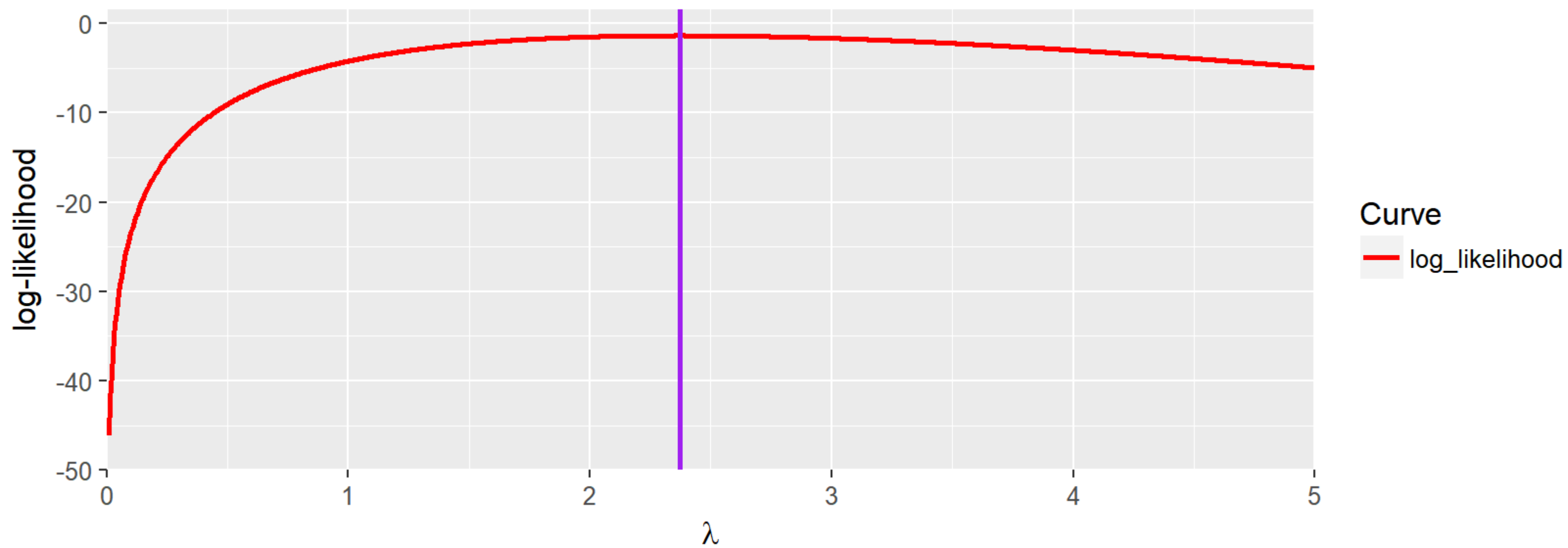
$$f(x_i|\lambda) = \lambda e^{-\lambda x_i}$$

$$L(\lambda|x_1, \ldots, x_n) = \cdots = n\log(\lambda) - \lambda \sum_{i=1}^{n} x_i$$

$$\frac{d}{d\lambda} L(\lambda|x_1, \ldots, x_n) = \frac{n}{\lambda} - \sum_{i=1}^{n} x_i$$

$$\frac{d^2}{d\lambda^2} L(\lambda|x_1, \ldots, x_n) = \frac{-n}{\lambda^2}$$

Log-likelihood function, maximum likelihood estimate



The purple line is at the MLE.

…But what if we couldn't solve for the MLE directly?

Taylor Series Approximation to L

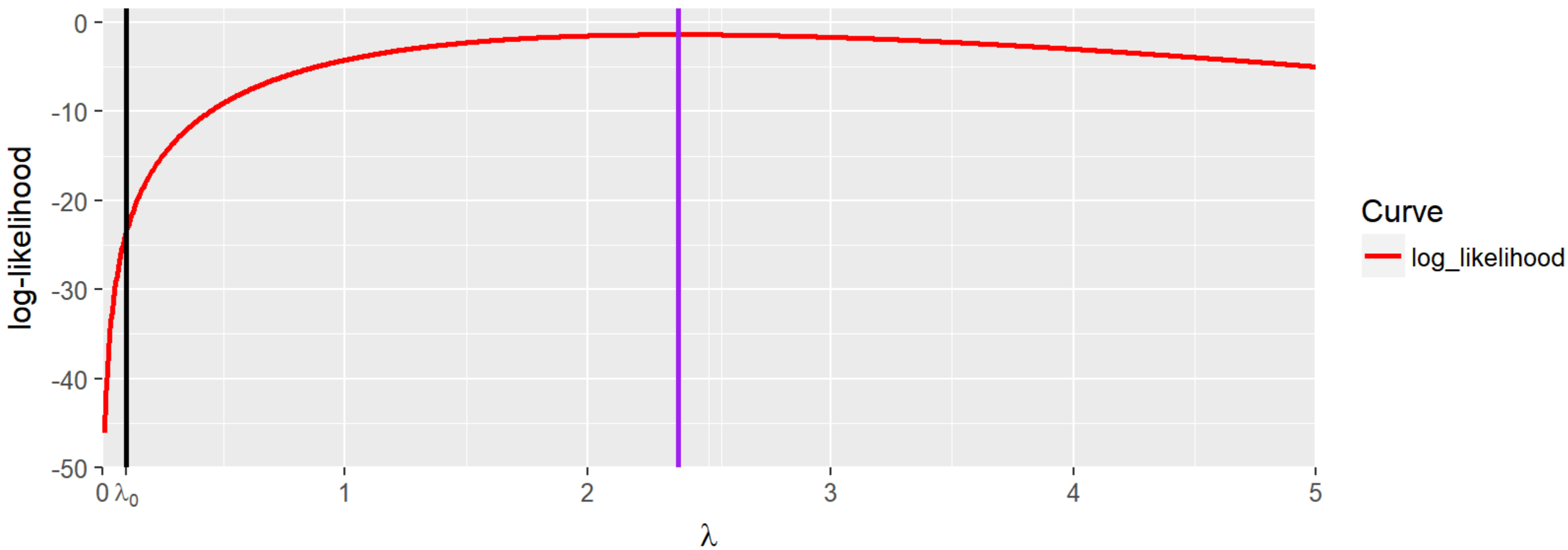
Pick a value λ_0. The second-order Taylor Series approximation to $L(\lambda|x_1, \ldots, x_n)$ around λ_0 is

$$\begin{aligned} P_2(\lambda) &= L(\lambda_0|x_1, \ldots, x_n) + \frac{d}{d\lambda}L(\lambda_0|x_1, \ldots, x_n)(\lambda - \lambda_0) \\ &+ \frac{1}{2}\frac{d^2}{d\lambda^2}L(\lambda_0|x_1, \ldots, x_n)(\lambda - \lambda_0)^2 \end{aligned}$$

The maximum of $P_2(\lambda)$ is at $\lambda_1 = \lambda_0 - \frac{\frac{d}{d\lambda}L(\lambda_0|x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2}L(\lambda_0|x_1, \ldots, x_n)}$

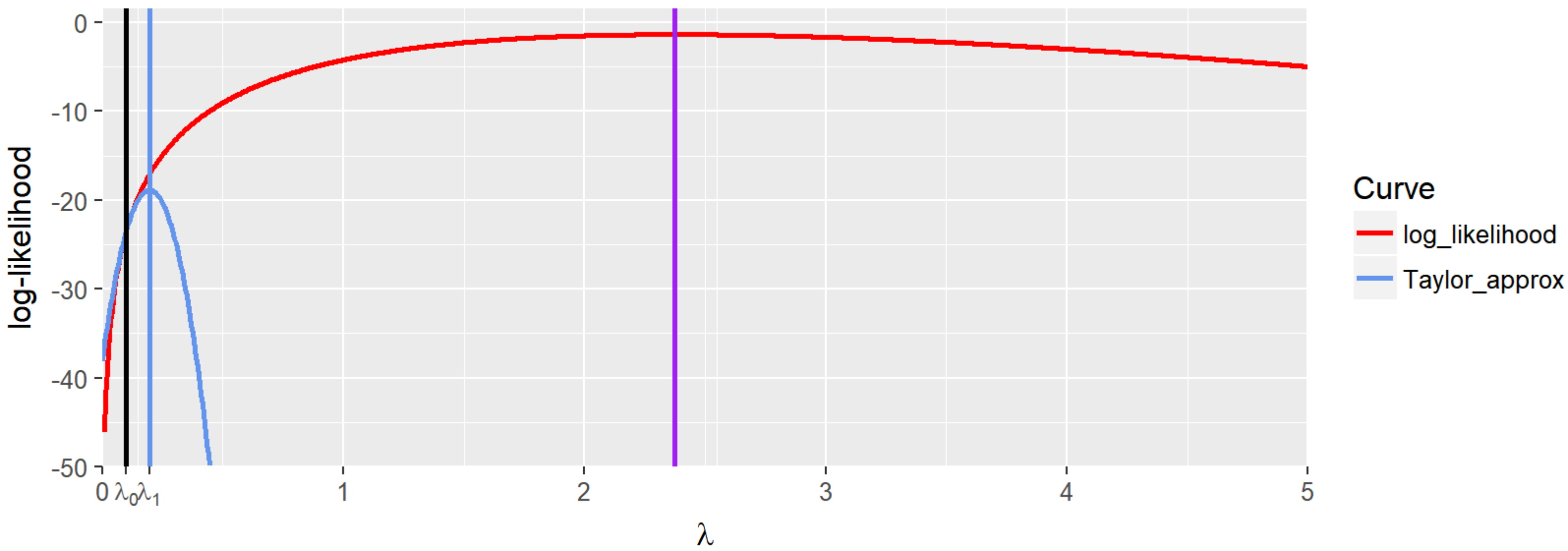
Now repeat, but centering the Taylor Series approximation at λ_1.

Pick λ_0



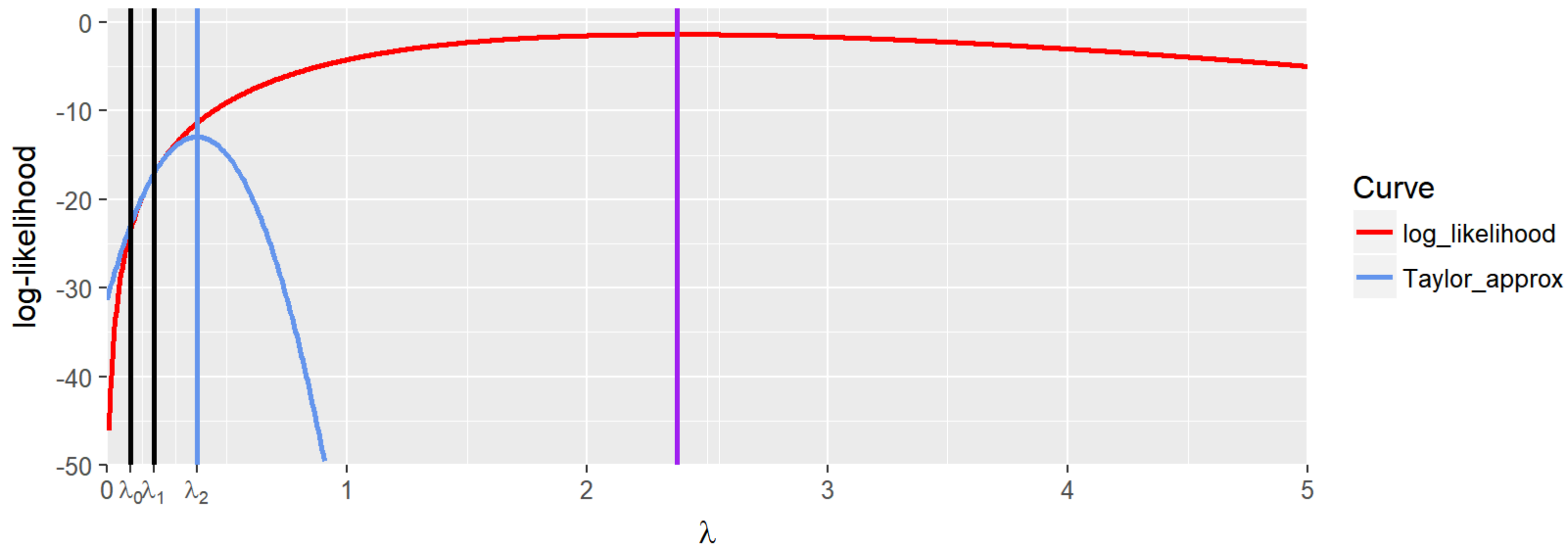
$$\lambda_0 = 0.1$$

Approximate L around λ_0, get λ_1



$$\lambda_0 = 0.1, \lambda_1 = \lambda_0 - \frac{\frac{d}{d\lambda} L(\lambda_0 | x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2} L(\lambda_0 | x_1, \ldots, x_n)} = 0.196$$

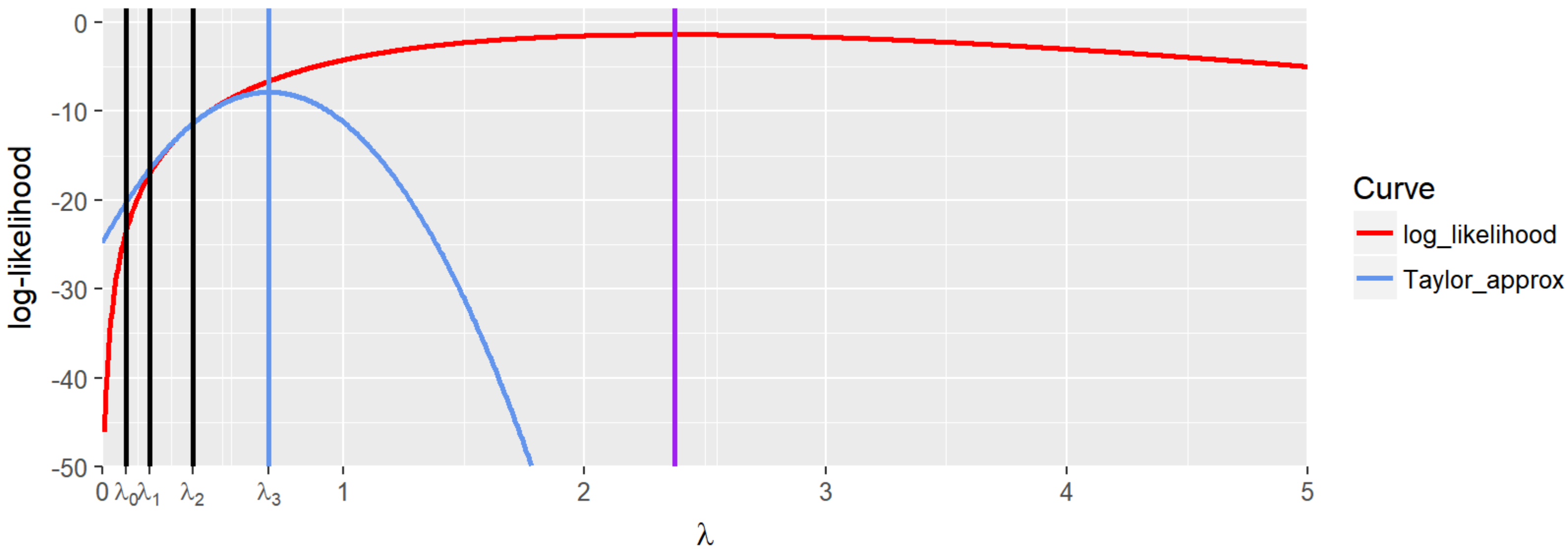
Approximate L around λ_1, get λ_2



$$\lambda_0 = 0.1,\ \lambda_1 = 0.196,$$

$$\lambda_2 = \lambda_1 - \frac{\frac{d}{d\lambda} L(\lambda_1 | x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2} L(\lambda_1 | x_1, \ldots, x_n)} = 0.375$$

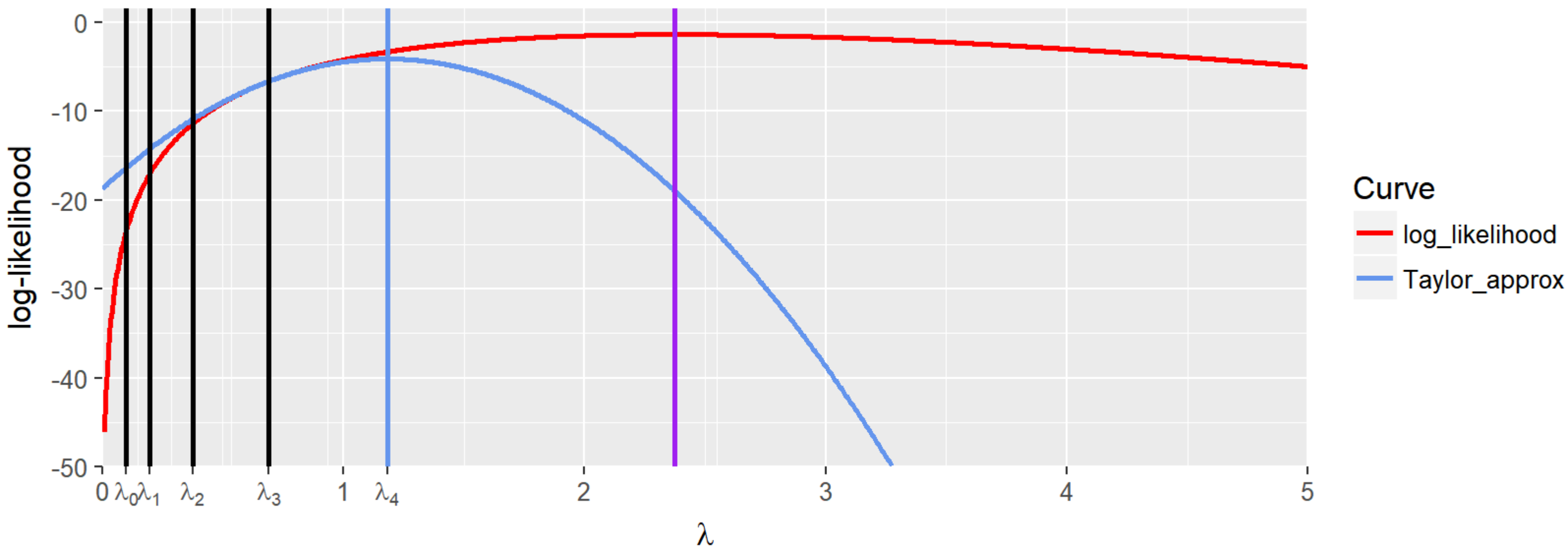
Approximate L around λ_2, get λ_3



$$\lambda_0 = 0.1,\ \lambda_1 = 0.196, \lambda_2 = 0.375$$

$$\lambda_3 = \lambda_2 - \frac{\frac{d}{d\lambda}L(\lambda_2|x_1,\ldots,x_n)}{\frac{d^2}{d\lambda^2}L(\lambda_2|x_1,\ldots,x_n)} = 0.691$$

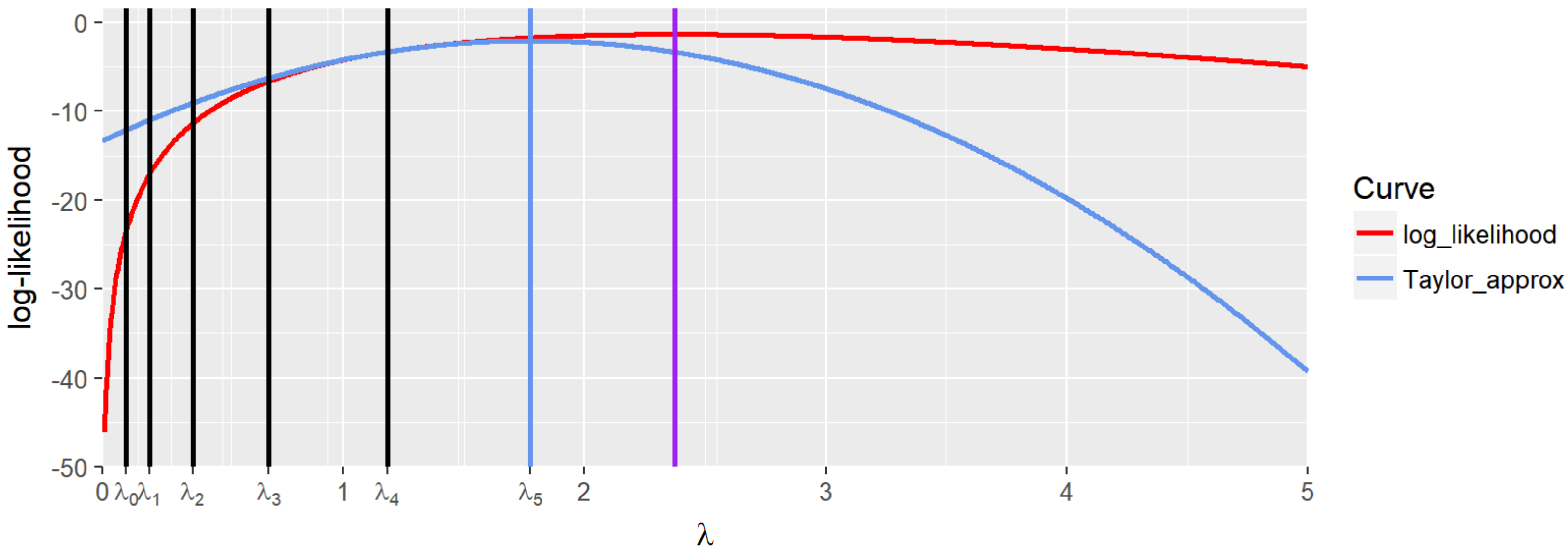
Approximate L around λ_3, get λ_4



$$\lambda_0 = 0.1,\ \lambda_1 = 0.196, \lambda_2 = 0.375, \lambda_3 = 0.691$$

$$\lambda_4 = \lambda_3 - \frac{\frac{d}{d\lambda} L(\lambda_3 | x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2} L(\lambda_3 | x_1, \ldots, x_n)} = 1.181$$

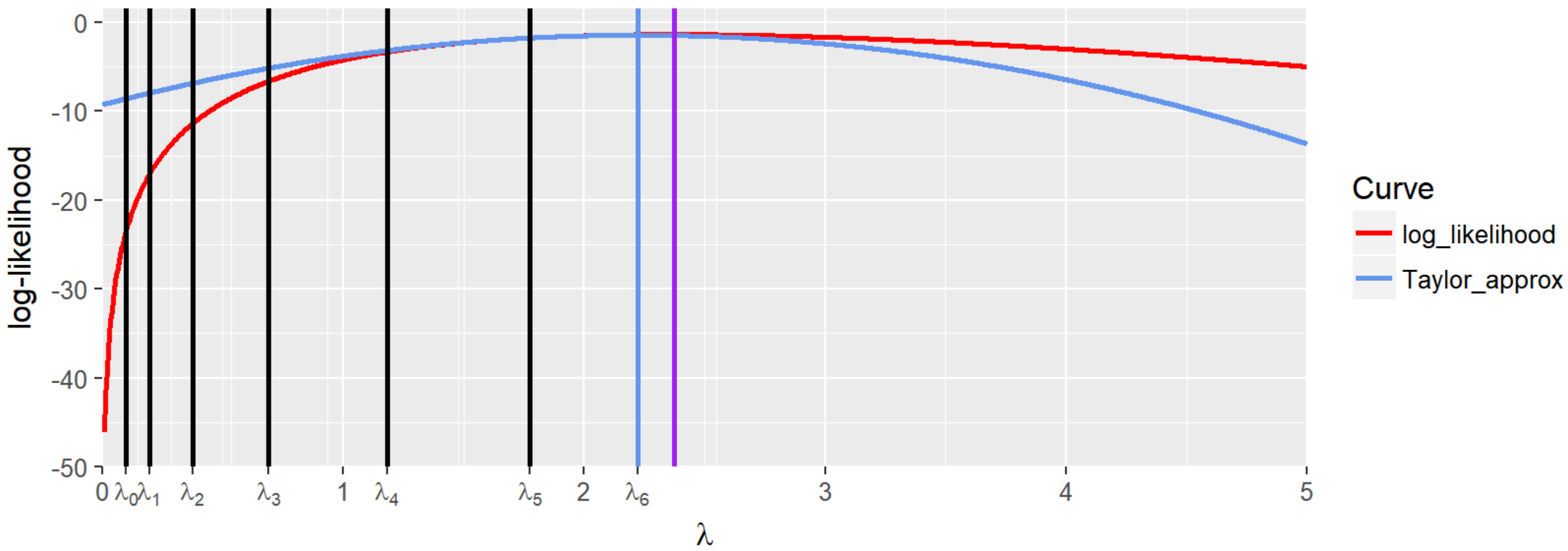
Approximate L around λ_4, get λ_5



$$\lambda_0 = 0.1,\ \lambda_1 = 0.196, \lambda_2 = 0.375, \lambda_3 = 0.691, \lambda_4 = 1.181$$

$$\lambda_5 = \lambda_4 - \frac{\frac{d}{d\lambda} L(\lambda_4 | x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2} L(\lambda_4 | x_1, \ldots, x_n)} = 1.775$$

Approximate L around λ_5, get λ_6



$$\lambda_0 = 0.1,\ \lambda_1 = 0.196, \lambda_2 = 0.375, \lambda_3 = 0.691, \lambda_4 = 1.181$$

$$\lambda_5 = 1.775,\ \lambda_6 = \lambda_5 - \frac{\frac{d}{d\lambda} L(\lambda_5 | x_1, \ldots, x_n)}{\frac{d^2}{d\lambda^2} L(\lambda_5 | x_1, \ldots, x_n)} = 2.223$$