NC STATE UNIVERSITY

Estimate Variances of Model Parameters Using Perturbed SSE Curve Fitting (PSCF) Method

Client: Prof. Robert Hayes
Department of Nuclear Engineering, North Carolina State University

Team: Chien-Lan Hsueh
Department of Statistics, North Carolina State University

Outline

- Introduction
  - Background
  - Research Questions
  - Research Methods
- Perturbed SSE Curve Fitting (PSCF) Method
  - Proposed Algorithm and Its Applications
  - Implementation
  - Fitting Strategies
  - Results
- Statistical Methods
  - Non-Linear Regression Model
  - Bootstrap Confidence Intervals
  - Delta Method Normality
- Comparison of Estimation
- Conclusion

Introduction

- To model a physical process, we often need to estimate the model parameters using the measurement data.
  - We are interested in not only the point estimates but also the uncertainties associated with them.
- Our client, Prof. Robert Hayes (Department of Nuclear Engineering, NCSU), has developed an algorithm to estimate such uncertainties.
- He would like to know if this method is valid, how well does it estimate and if it can be backed up by statistics theory.

Background

- The research of retrospective dosimetry involves measurements of radiation dosages and inferences of the actual radiation exposures
- A direct radiation measurement on the subjects provides an accurate depth profile measurement
  - But it is costly and time consuming. Sometimes, it can be dangerous to do.
- Therefore, Prof. Hayes has developed a new method:
  1. Analyze forensic luminescence data that is high correlated to the actual radiation dosage
  2. Fit the dose deposition profile using Monte Carlo n-particle method (MCNP)
  3. With the obtained best fitted model parameter as a reference, a series of "perturbed" model fittings is then used in a normal curve fitting to estimate the variance of the parameter

Research Questions

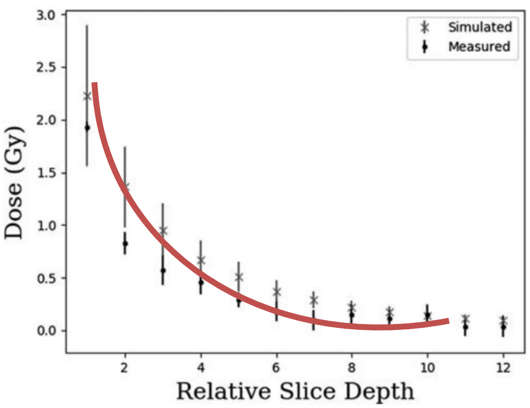
1. How good the algorithm is in term of determining the variances of physical model parameters?
2. If there is a solid statistics ground to support and backup the validity of this method?
3. If yes in (2), can it be improved and further generalized to any physical models?
4. If no in (2), how well does it estimate as an approximation approach?

Research Methods

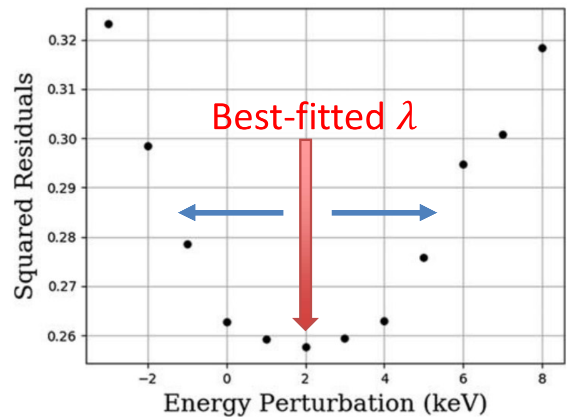
- In this study, we implement the algorithm and investigate how well dose it estimate a model parameter's standard deviation.
- We also compare it with other statistical methods:
  1. Nonlinear Regression Model
  2. Bootstrap Confidence Intervals
     - Parametric bootstrap
     - Non-parametric bootstrap
  3. Delta Method Normality

Perturbed SSE Curve Fitting Method

1. Measurements of forensic luminescence

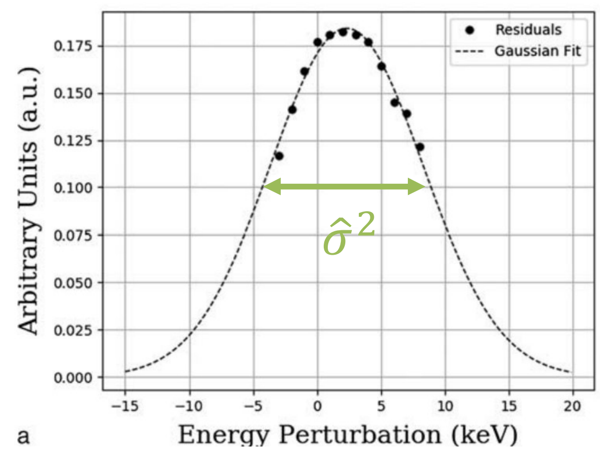


2. MCNP fitting to obtain the best-fit parameter λ



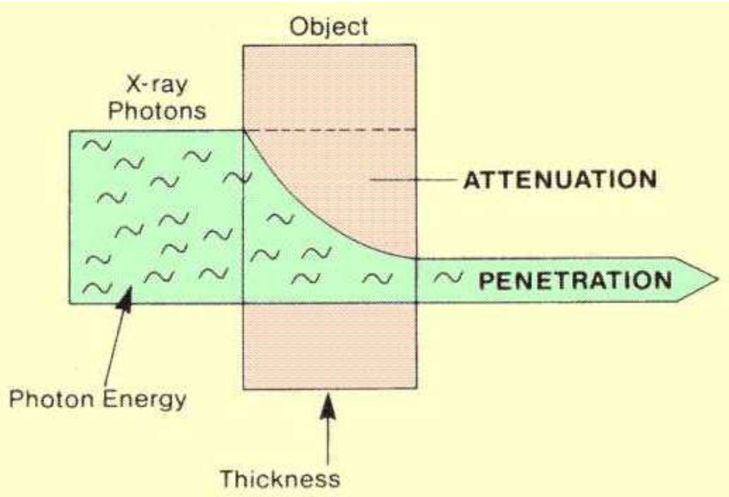
3. "Perturbed" the best-fit parameter λ for SSE curve

4. Inverse the SSE curve



5. Estimate $\hat{\sigma}^2$ by fitting a Gaussian curve

Case Study – Radiation Decay

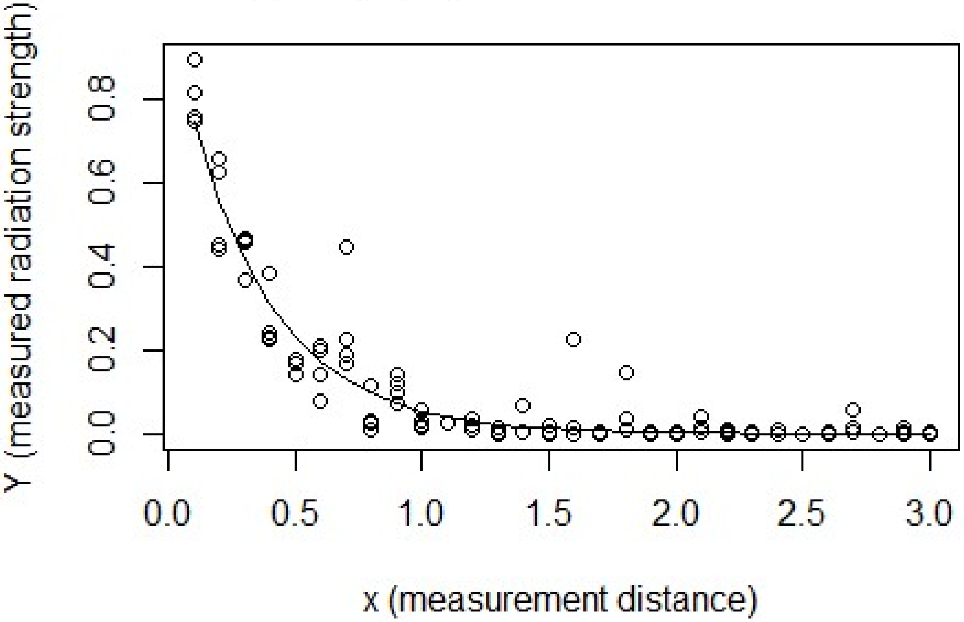


http://www.sprawls.org/ppmi2/RADPEN/

$$Y = e^{-\lambda X}$$
$$\lambda_i \overset{iid}{\sim} N(\mu, \sigma^2)$$

Physical model

Model parameter λ



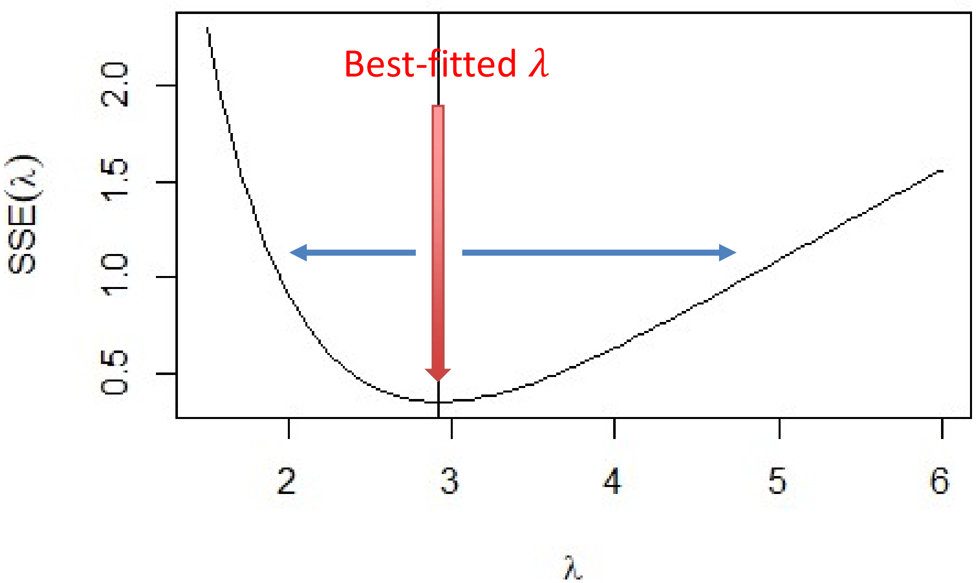
Gaussian Curve Fit

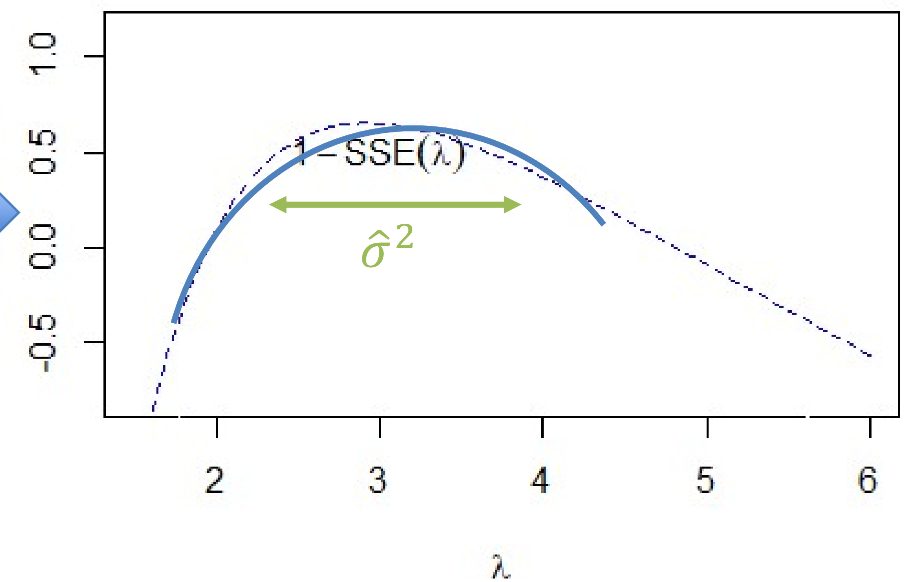
$Y = e^{-\lambda X}$ — Physical model

$\lambda_i \overset{iid}{\sim} N(\mu, \sigma^2)$ — Model parameter λ

Inverse SSE curve

Fit a Gaussian curve

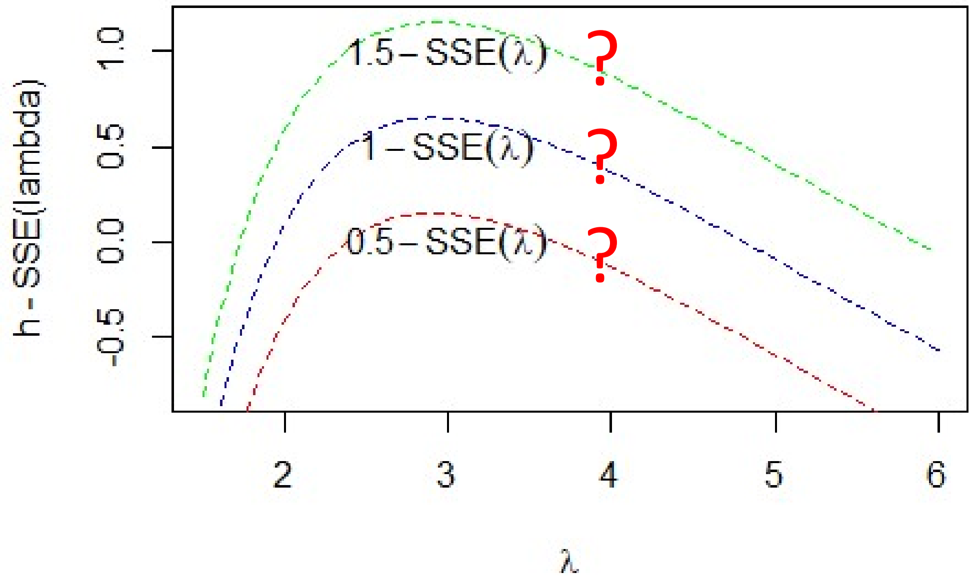




Challenges of PSCF Method

1. How to inverse SSE(λ)?
2. Fitting range to consider?
    - How wide?
    - Location of fitting range?

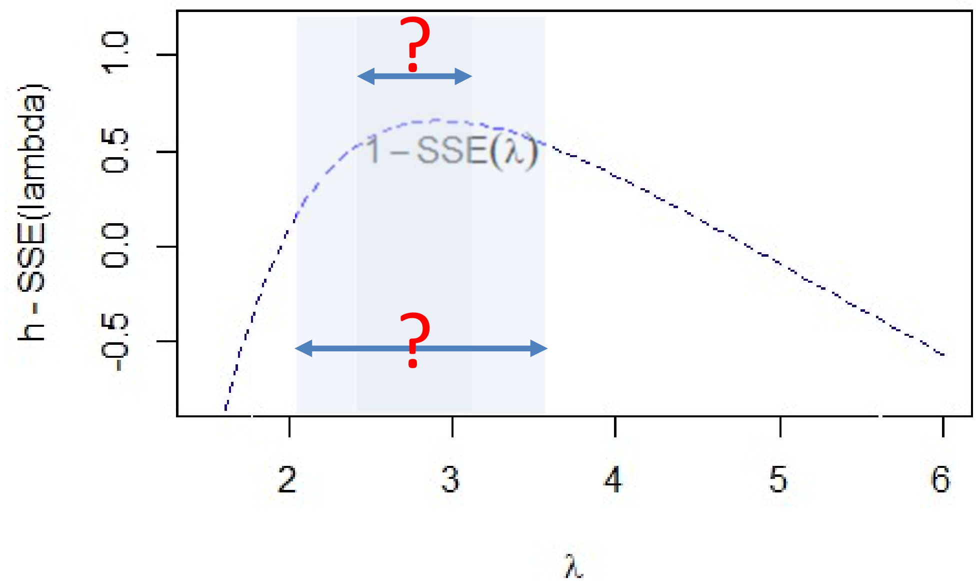
Inverse SSE curve
Fit a Gaussian curve



Challenges of PSCF Method

1. How to inverse SSE(λ)?
2. Fitting range to consider?
   - How wide?
   - Location of fitting range?

Inverse SSE curve
Fit a Gaussian curve



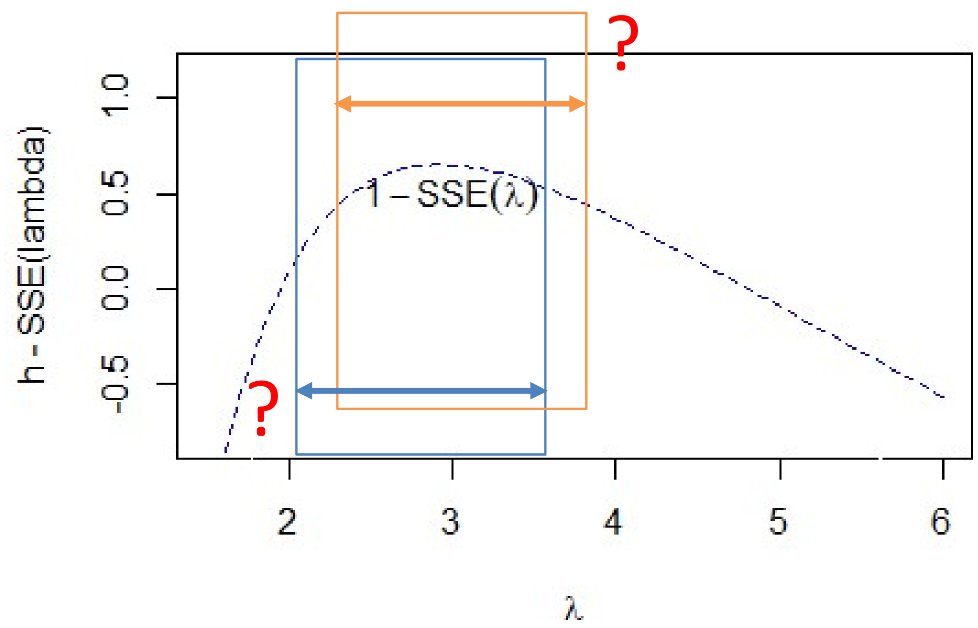
Challenges of PSCF Method

1. How to inverse SSE(λ)?

2. Fitting range to consider?
   - How wide?
   - Location of fitting range?

Inverse SSE curve

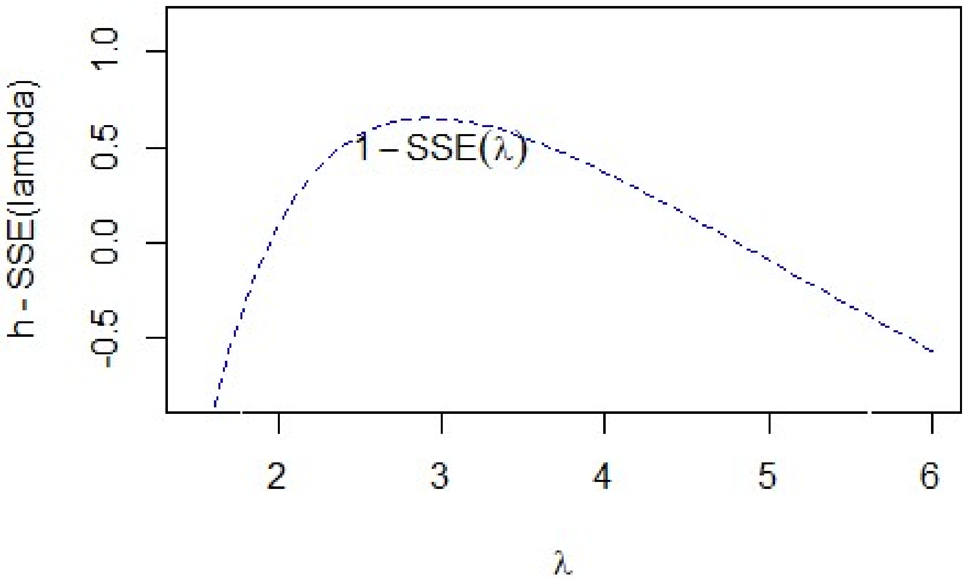
Fit a Gaussian curve



Objective Function

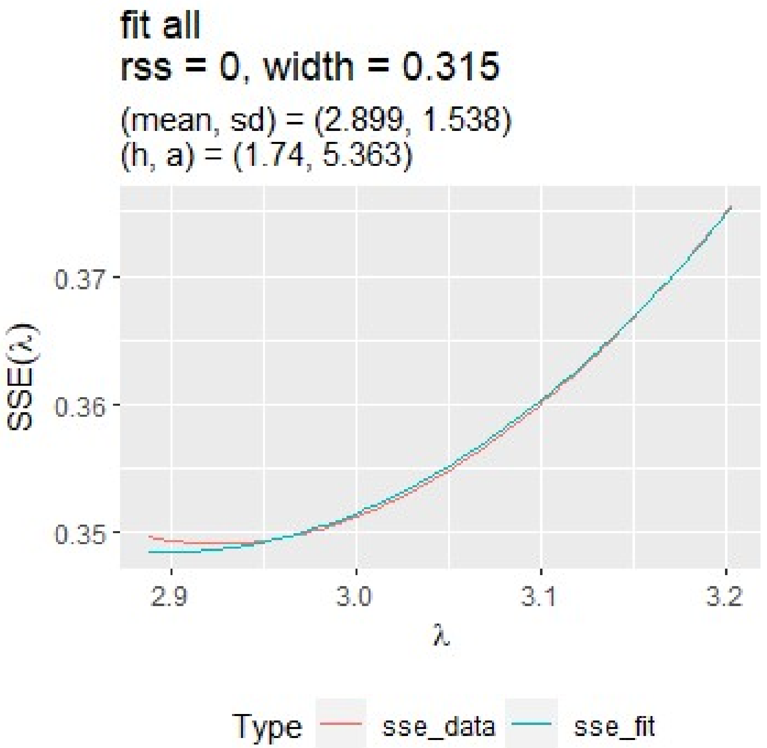
1. How to inverse SSE(λ)?
2. Fitting range to consider?
   - How wide?
   - Location of fitting range?

$$\widehat{SSE}(\lambda = x) \simeq h - a \cdot e^{\frac{-1}{2\sigma^2}(x-\mu)^2}$$



- h, a: how to inverse and scale SSE(λ)
- μ, σ^2: Gaussian curve
- range
  - Lower bound
  - Width

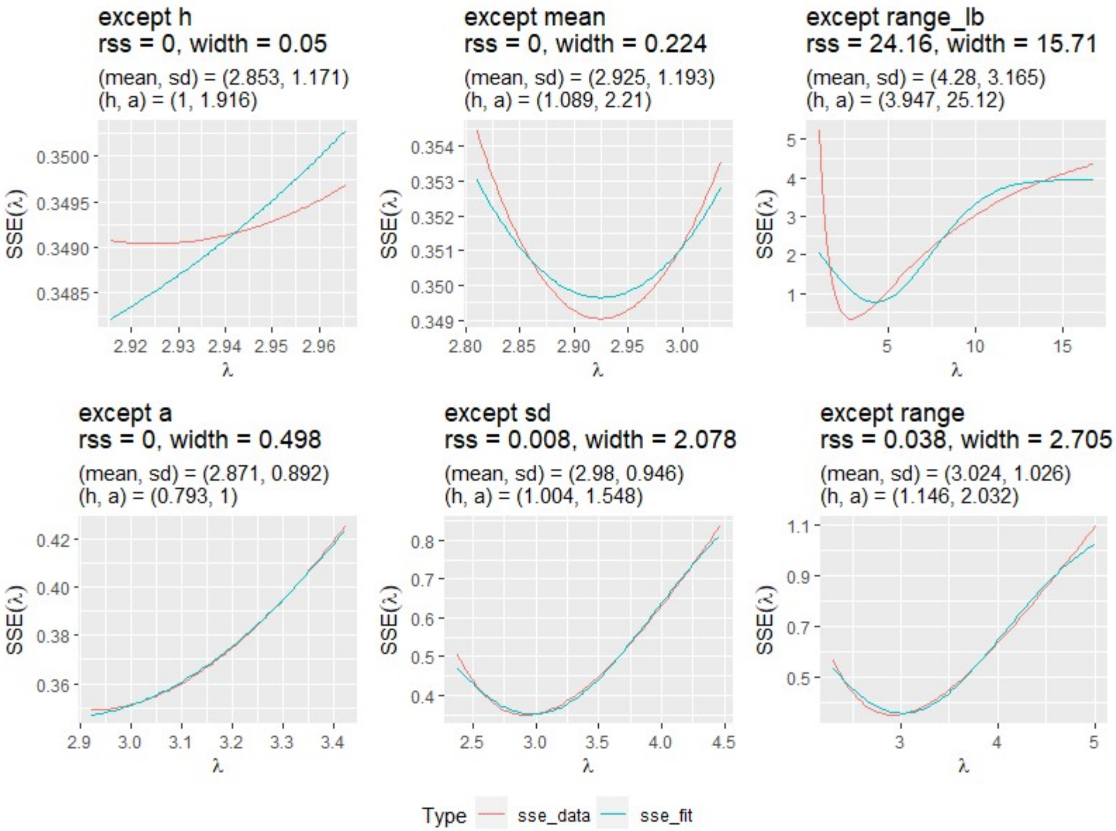
Fit All 6 Parameters



- The smallest fitting error (RSS) we can get
- But a very narrow fitting range is used to fit
- Off-centered fitting range

Fit 6-1=5 Parameters

Fix 1 parameter and keep the others free to vary



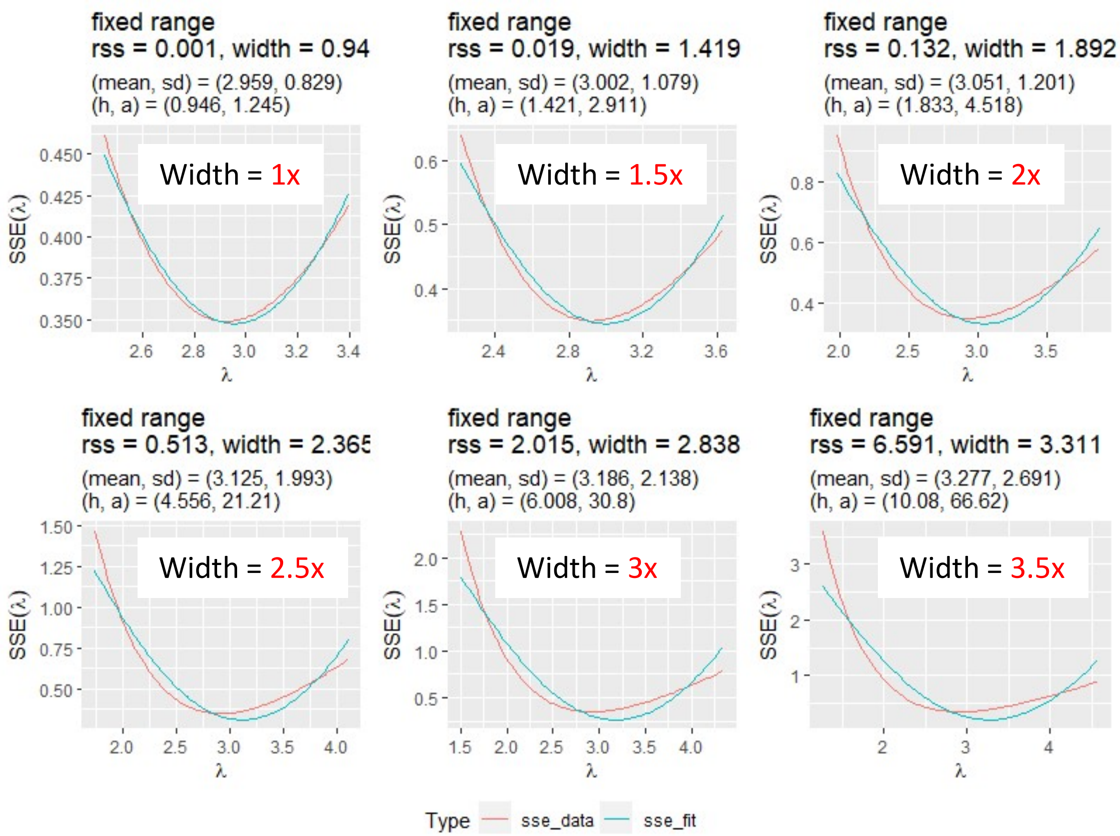
- The fittings with small RSS use narrow fitting range
- Many of them are off-centered

Learnings:

- Need to specify the fitting range (width)

Fit 6-2=4 Parameters

Fix range (width) and force it centered (location)



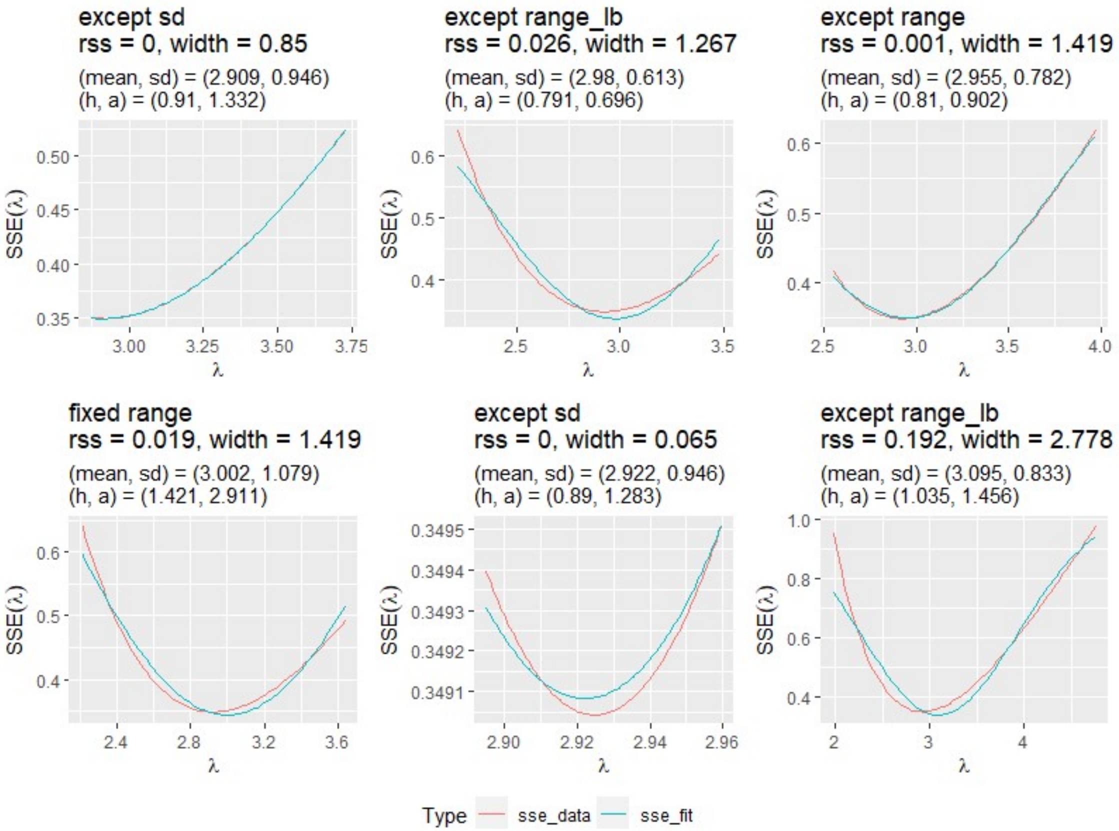
- Fix the fitting range to be 1x, 1.5x, 2x, 2.5x, 3x and 3.5x wide
- Force the fitting range centered at the minimum of SSE(λ)

Learnings:

- The obtained $\hat{\sigma}$ vary from 0.829 to 2.691
- The estimate using this method is not stable

Fit 6-2=4 Parameters

Fix the 2 range parameters (lower bound and width)



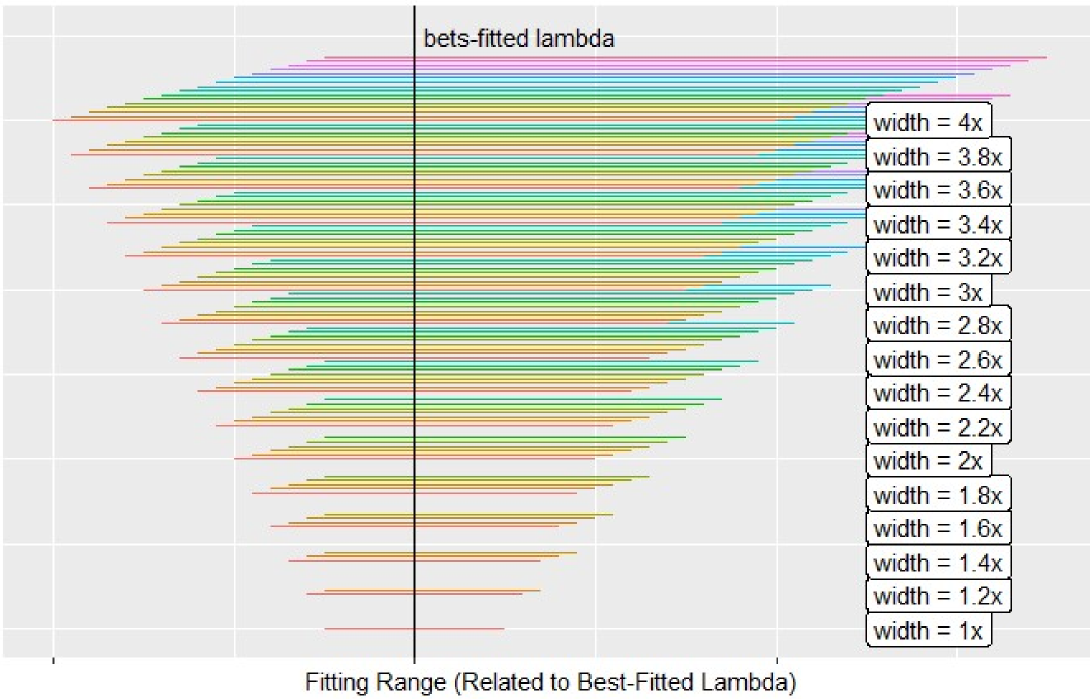
- The fittings with small RSS use narrow fitting range
- Many of them are off-centered

Learnings:

- Need to specify the fitting range (width)
- Need to specify the center

Fit 6-2=4 Parameters – Grid Search

Specify fitting range (width) and lower bound (location)

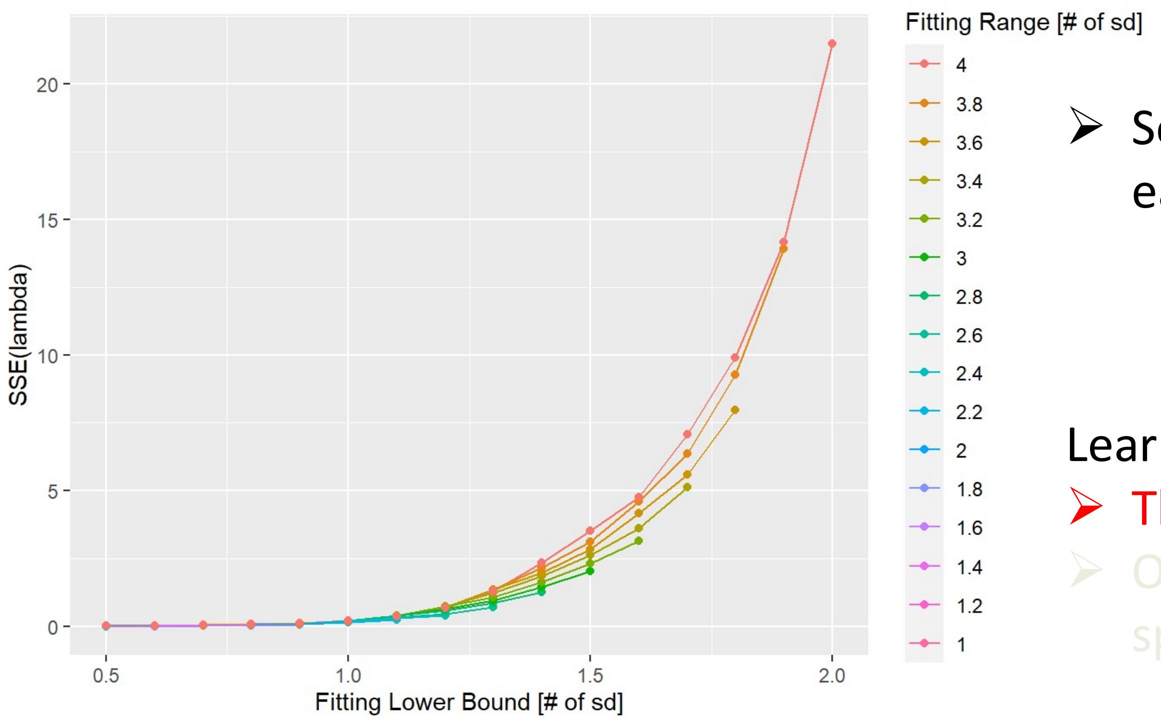


- Set up a grid of the fitting ranges for each Gaussian curve fitting
  - Fitting range (width): 1x, 1.2x,…, 4x
  - Lower bound (location)

  - Totally 136 scenarios

Fit 6-2=4 Parameters – Grid Search

Fix range (width) and force it centered (location)



- Set up a grid of the fitting ranges for each Gaussian curve fitting
  - Fitting range (width)
  - Lower bound (location)

Learnings:

- There is no global minimum for SSE(λ)
- Obtained $\hat{\sigma}$ vary depending on how we specify the fitting range

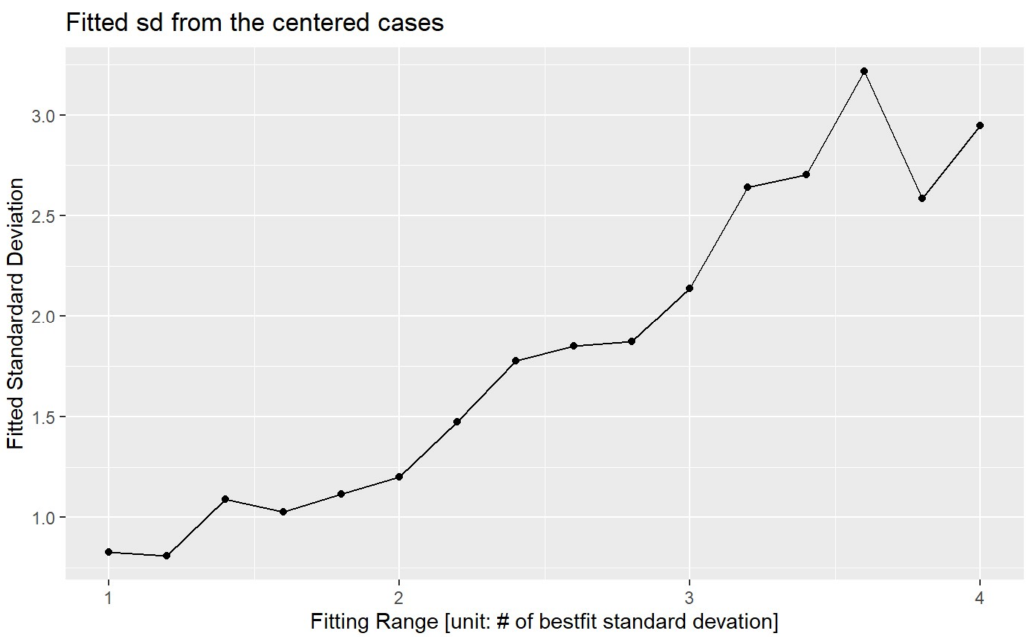
PSCF Method

method	fitting_initial_setup	lambda	std.error
<chr>	<S3: glue>	<dbl>	<dbl>
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -2, range=4)	3.349	0.29470
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.9, range=3.8)	3.324	0.25860
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.8, range=3.6)	3.274	0.32150
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.7, range=3.4)	3.247	0.27020
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.6, range=3.2)	3.230	0.26390
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.5, range=3)	3.186	0.21380
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.4, range=2.8)	3.153	0.18730
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.3, range=2.6)	3.119	0.18510
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.2, range=2.4)	3.102	0.17760
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1.1, range=2.2)	3.081	0.14730
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -1, range=2)	3.051	0.12010
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -0.9, range=1.8)	3.038	0.11160
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -0.8, range=1.6)	3.011	0.10280
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -0.7, range=1.4)	2.990	0.10900
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -0.6, range=1.2)	2.973	0.08096
PSCF - Centered	init. (h, a, lb, range)= (1, 1, -0.5, range=1)	2.959	0.08292

Range: 0.08096 - 0.3215

Fit 6-2=4 Parameters – Grid Search

Fix range (width) and force it centered (location)



- Set up a grid of the fitting ranges for each Gaussian curve fitting
  - Fitting range (width)
  - Lower bound (location)

Learnings:

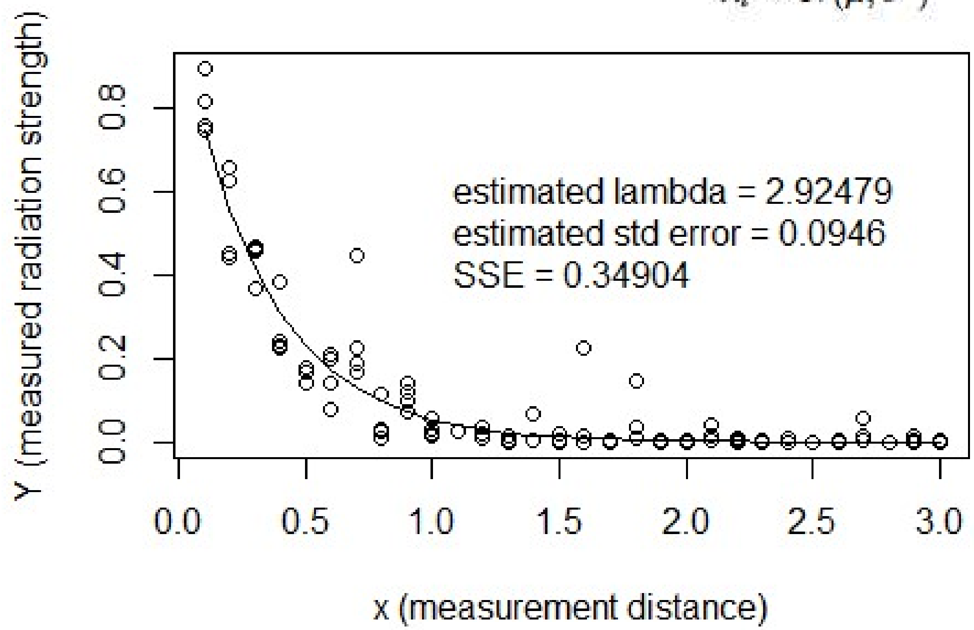
- There is no global minimum for SSE(λ)
- Obtained $\hat{\sigma}$ vary depending on how we specify the fitting range

Non-Linear Regression

$$Y = e^{-\lambda X}$$
$$\lambda_i \overset{iid}{\sim} N(\mu, \sigma^2)$$

```
Nonlinear regression model
  model: y ~ fmod(x, lambda)
   data: df
lambda
 2.925
 residual sum-of-squares: 0.349

Number of iterations to convergence: 2
Achieved convergence tolerance: 7.308e-06
```

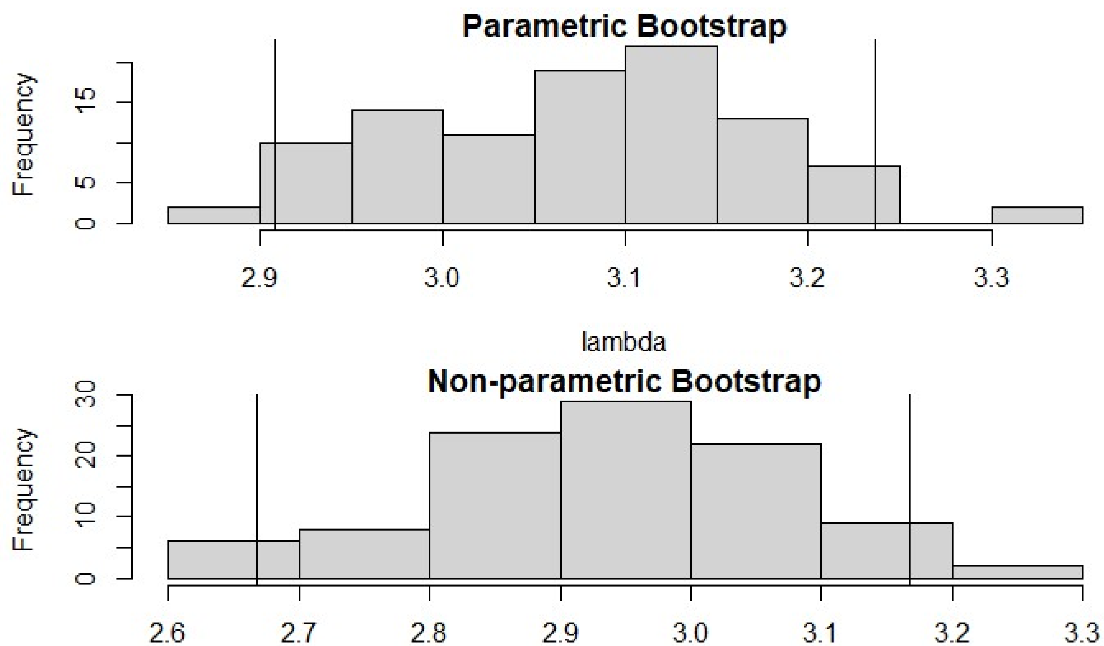


method <chr>	lambda <dbl>	std.error <dbl>	lower_0.95 <dbl>	upper_0.95 <dbl>	width <dbl>
Nonlinear Regression	2.924786	0.0946027	2.812288	3.187712	0.3754246

Bootstrap Confidence Interval

$$\hat{SE}(\hat{\underset{\sim}{\alpha}}) = \sqrt{\frac{1}{B}\sum_{i=1}^{B}\left(\hat{\alpha}_j^* - \bar{\hat{\alpha}}^*\right)^2}$$

$$= \sqrt{\begin{array}{c}\text{Biased sample variance}\\ \text{of bootstrapped estimates}\end{array}}$$

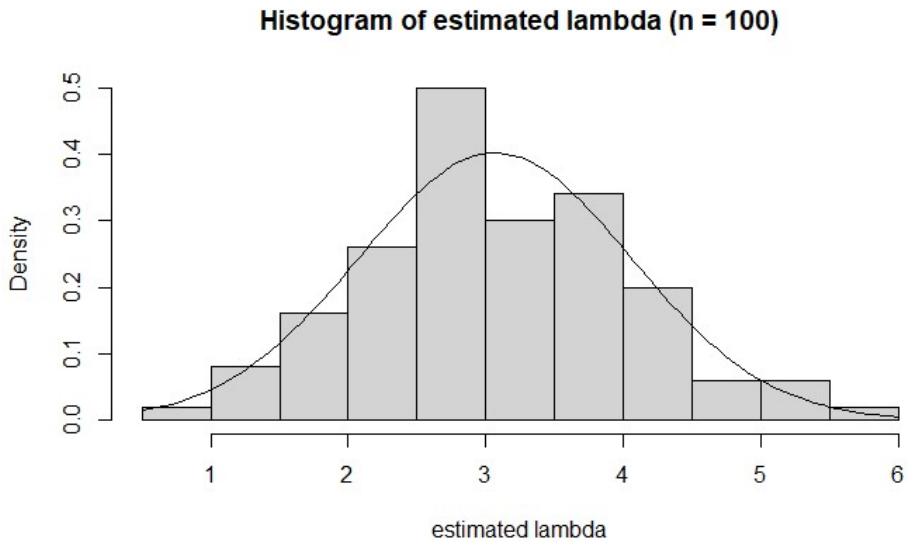


method <chr>	lambda <dbl>	std.error <lgl>	lower_0.95 <dbl>	upper_0.95 <dbl>	width <dbl>
Parametric Bootstrap	3.075166	NA	2.908155	3.236575	0.3284194
Non-parametric Bootstrap	2.931697	NA	2.668177	3.167169	0.4989925

Delta Method Normality

$$\hat{\lambda}(Y) = -\frac{1}{x}\ln Y$$

$$\Rightarrow g(Y) \overset{\bullet}{\sim} N\left(g(\mu),\, [g'(\mu)]^2\sigma^2\right)$$
$$g(Y_n) \overset{\bullet}{\sim} N\left(g(\theta_0),\, [g'(\theta_0)]^2\sigma^2/n\right)$$



method <chr>	lambda <dbl>	std.error <dbl>	lower_0.95 <dbl>	upper_0.95 <dbl>	width <dbl>
Delta Method Normality	3.067518	0.09930124	2.870483	3.264554	0.3940704

Comparison

method <chr>	<chr>	lambda <dbl>	std.error <dbl>
PSCF - Centered	max	3.349	0.32150
	min	2.959	0.08096

method <chr>	lambda <dbl>	std.error <dbl>	lower_0.95 <dbl>	upper_0.95 <dbl>	width <dbl>
Nonlinear Regression	2.924786	0.09460270	2.812288	3.187712	0.3754246
Parametric Bootstrap	3.075166	NA	2.908155	3.236575	0.3284194
Non-parametric Bootstrap	2.931697	NA	2.668177	3.167169	0.4989925
Delta Method Normality	3.067518	0.09930124	2.870483	3.264554	0.3940704
Truth (All sample sizes = 100)	3.000000	0.10000000	2.801578	3.198422	0.3968434

Conclusion

Findings Regarding the Research Question

1. How good the algorithm is in term of determining the variances of physical model parameters? The obtained estimation greatly depends on initial conditions of the Gaussian curve fitting.
2. If there is a solid statistics ground to support and backup the validity of this method? No, because there is no global minimum in the Gaussian curve fitting.
3. If yes in (2), can it be improved and further generalized to any physical models?
4. If no in (2), how well does it estimate as an approximation approach? It gives a wide range of estimation on the standard deviation.

Although the proposed method is not a good way to estimate the variance of a model parameter, we can still rely on the other statistical methods like Non-linear regression, Bootstrap CI and Delta method.