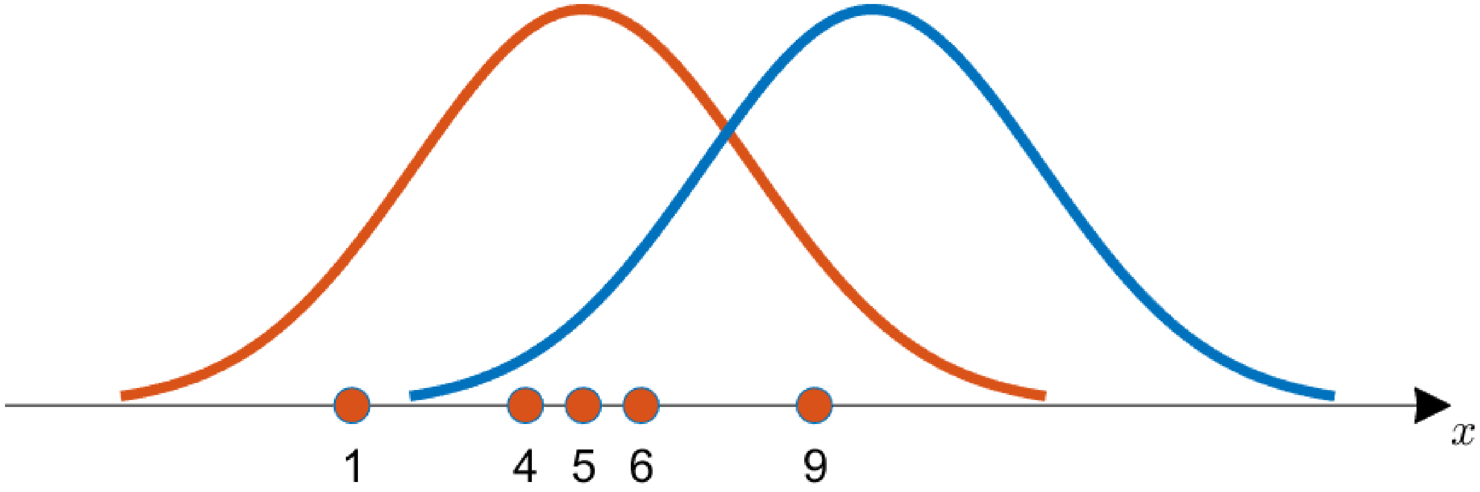
Maximum Likelihood Estimation (MLE)

20201181 JIHWAN OH

What is Likelihood?

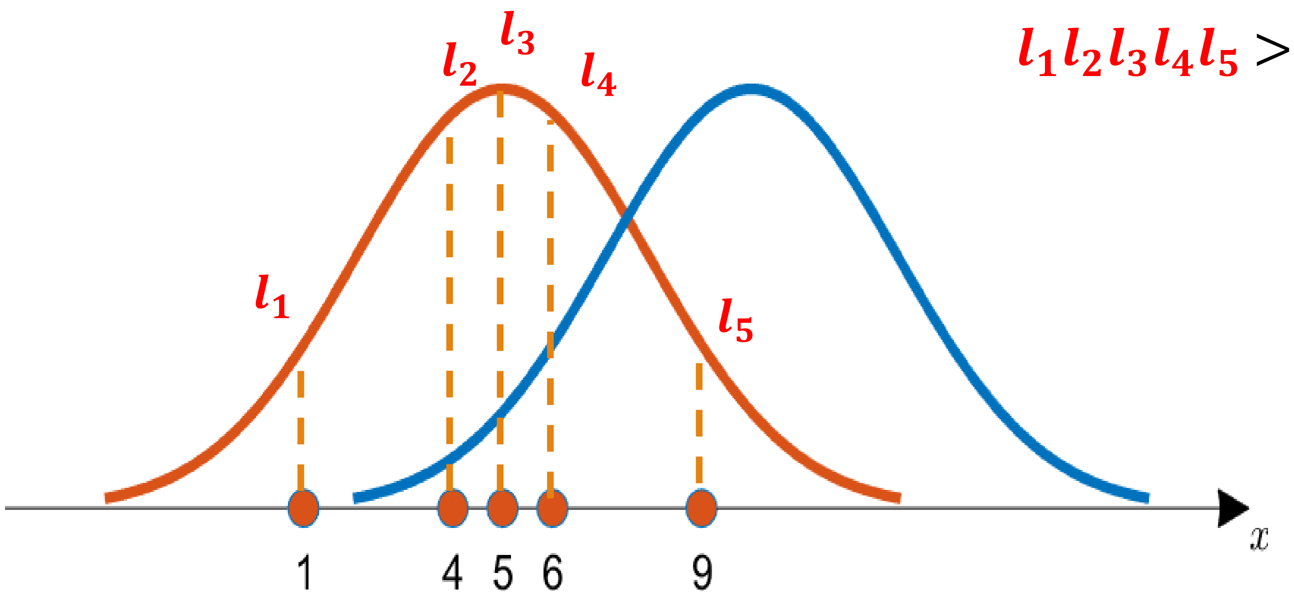
Given data x = { 1, 4, 5, 6, 9 }

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)}{2\sigma_1}} \qquad f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{\frac{(x-\mu_2)}{2\sigma_2}}$$

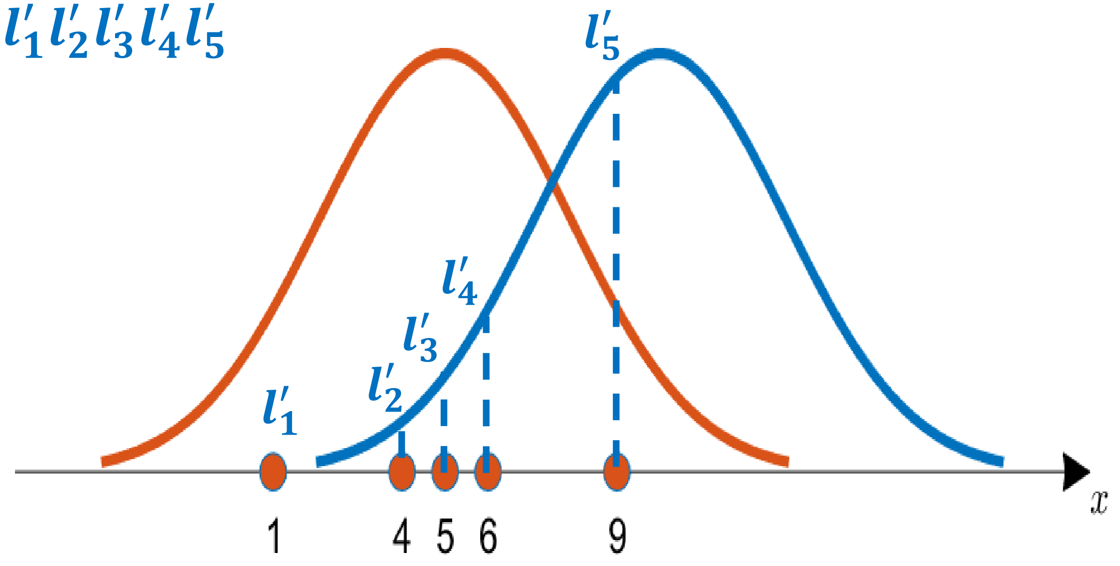


What is Likelihood?

Given data x = { 1, 4, 5, 6, 9 }



$$l_1 l_2 l_3 l_4 l_5 > l'_1 l'_2 l'_3 l'_4 l'_5$$

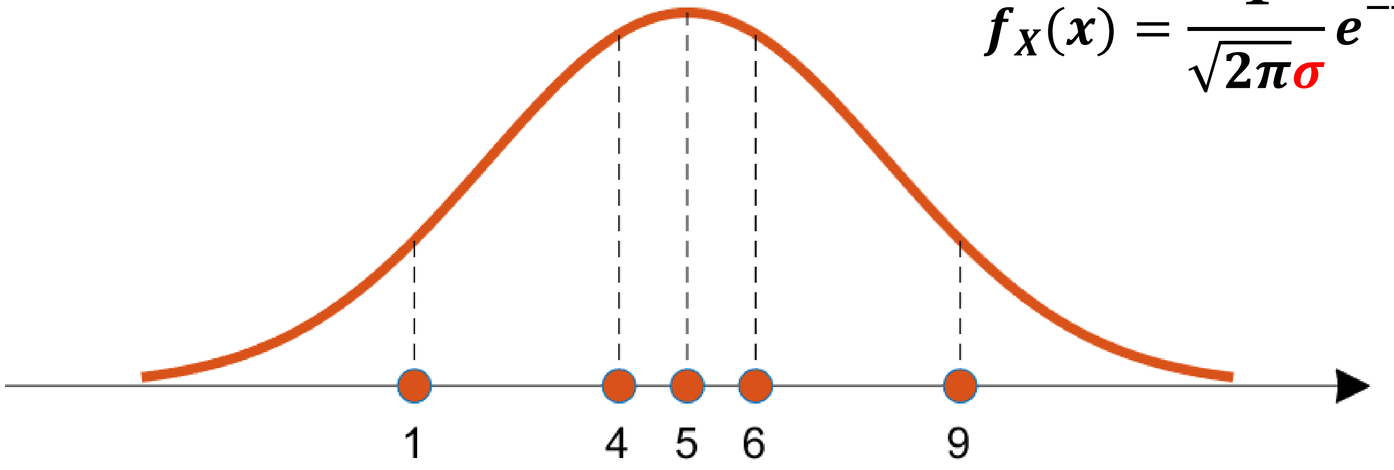


What is Likelihood?

Given data x = { 1, 4, 5, 6, 9 }

Q. What is best parameter $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ that explain the given data well ?

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)}{2\sigma}}$$



What is Likelihood?

Definition: $L(\theta) = p(X|\theta)$ $(Likelihood\ function)$

$$= \prod_{n=1}^{N} p(x_n|\theta)$$ x_n are independent

In Gaussian distribution, $\theta = (\mu, \sigma)$

$$\mathrm{p}(\mathrm{x_n} \mid \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x_n - \mu)^2}{2\sigma^2}\}$$

Maximum Likelihood Estimation

We want to maximize $$L(\theta) = p(X|\theta) = \prod_{n=1}^{N} p(x_n|\theta) \quad (Likelihood)$$

Objective: Minimize $$E(\theta) = -\ln L(\theta) = -\sum_{n=1}^{N} \ln p(x_n|\theta) \quad (-log\ Likelihood)$$

Find parameter θ that minimize –Log Likelihood

$$\boxed{\frac{\partial}{\partial\theta}E(\theta) = -\frac{\partial}{\partial\theta}\sum_{n=1}^{N} \ln p(x_n|\theta) = -\sum_{n=1}^{N} \frac{\frac{\partial}{\partial\theta}p(x_n|\theta)}{p(x_n|\theta)} \stackrel{!}{=} 0}$$

Maximum Likelihood Estimation

In Gaussian distribution : $p(x_n|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{||x_n-\mu||^2}{2\sigma^2}}$

- $$\begin{aligned} \frac{\partial}{\partial\mu} E(\mu,\sigma) &= -\sum_{n=1}^{N} \frac{\frac{\partial}{\partial\mu} p(x_n|\mu,\sigma)}{p(x_n|\mu,\sigma)} \\ &= -\sum_{n=1}^{N} -\frac{2(x_n-\mu)}{2\sigma^2} \\ &= \frac{1}{\sigma^2}\sum_{n=1}^{N}(x_n-\mu) \\ &= \frac{1}{\sigma^2}\left(\sum_{n=1}^{N} x_n - N\mu\right) \end{aligned}$$

Estimated parameters

$$\boxed{\hat{\mu} = \frac{1}{N}\sum_{n=1}^{N} x_n}$$

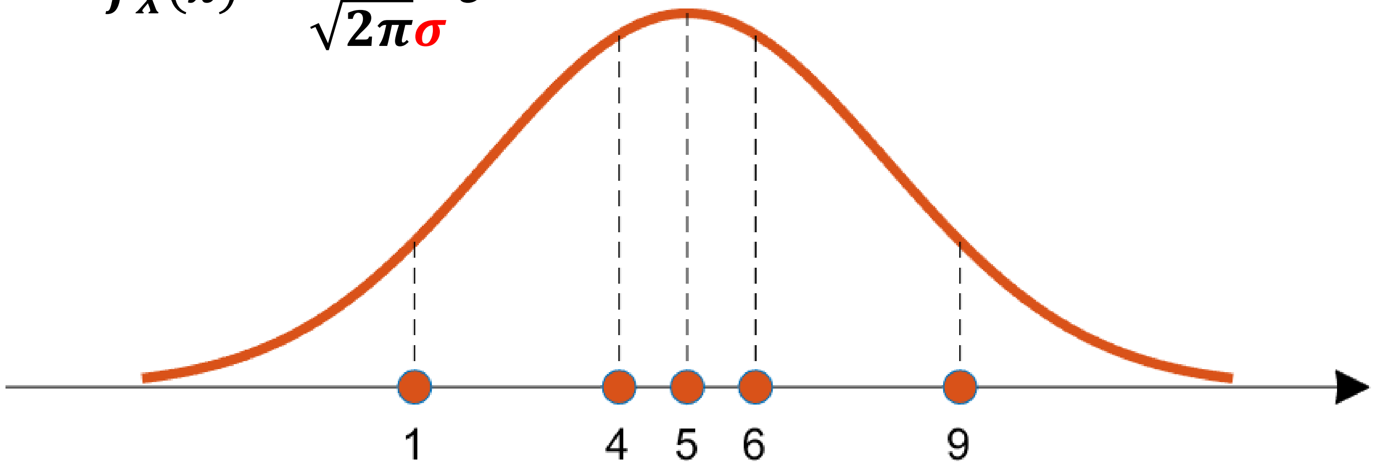
$$\boxed{\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n-\hat{\mu})^2}$$

Maximum Likelihood Estimation

Given data x = { 1, 4, 5, 6, 9 }

Q. What is best parameter $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma})$ that explain the given data well ?

$$\boldsymbol{f_X(x)} = \frac{\mathbf{1}}{\sqrt{\mathbf{2\pi}\boldsymbol{\sigma}}} \boldsymbol{e}^{-\frac{(\boldsymbol{x}-\boldsymbol{\mu})}{\mathbf{2}\boldsymbol{\sigma}}}$$



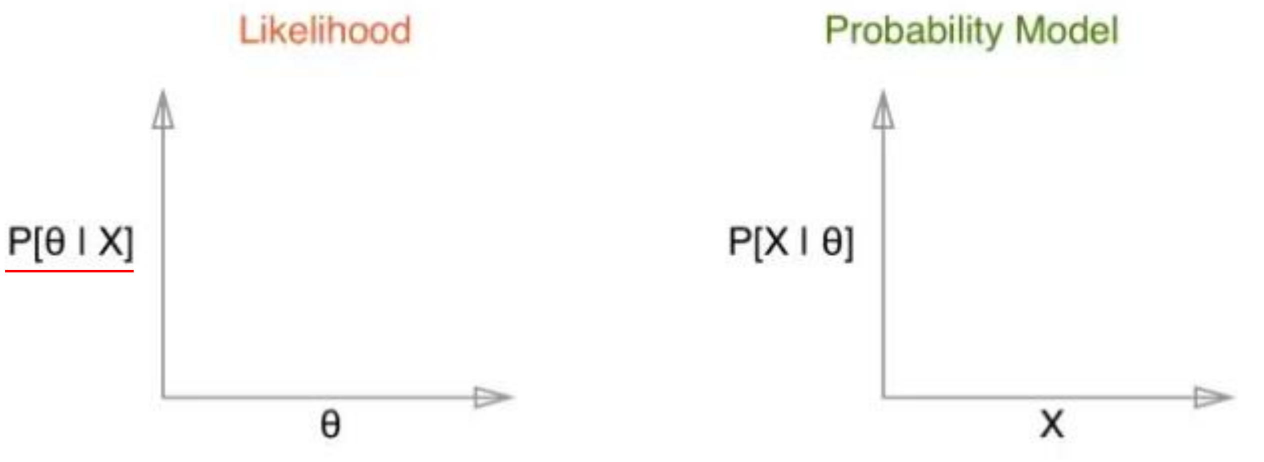
Answer:

$$\hat{\mu} = \frac{1}{N}\sum_{n=1}^{N} x_n = \frac{1+4+5+6+9}{5} = \mathbf{5}$$

$$\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \hat{\mu})^2$$

$$= \frac{(-4)^2+(-1)^2+0+1^2+4^2}{5} = \mathbf{6.8}$$

Likelihood ≠ Probability



$$= \prod_{n=1}^{N} p(x_n|\theta)$$

(Product of Probabilities)

Not probability!

$\frac{d}{d\theta}$

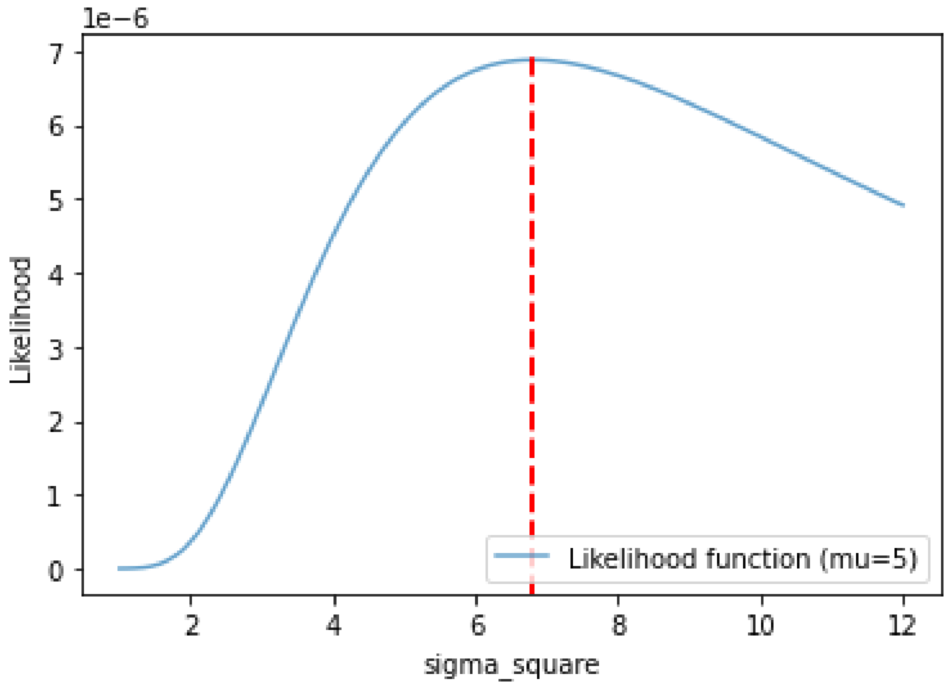
argmax of f with respect to θ

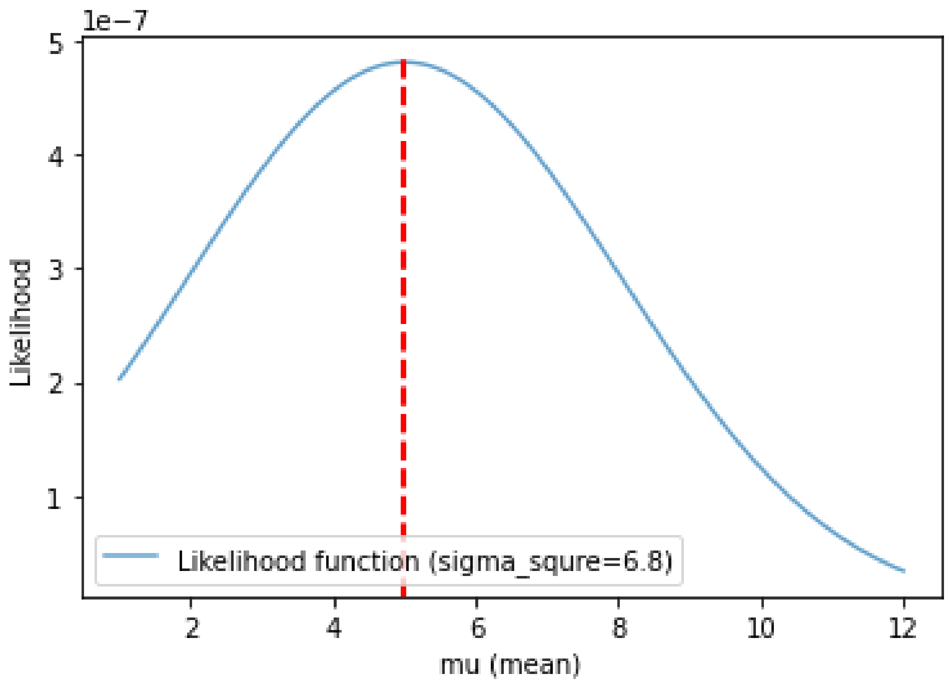
$\frac{d}{dx}$

argmax of f with respect to x

MLE in Python

Given data x = { 1, 4, 5, 6, 9 }





Conclusion: We can estimate the parameters of distribution using MLE.

Code

[1]

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import erf
```

```python
mu = 5
list1 = []
for i in np.linspace(1, 12, 100):
    sigma2 = i
    Product = 1
    for x in [1,4,5,6,9]:

        y = (1 / np.sqrt(2 * np.pi * sigma2)) * np.exp(-(x-mu)**2 / (2 * sigma2))

        Product = Product * y

    list1.append(Product)
    print(Product)

```

[43]

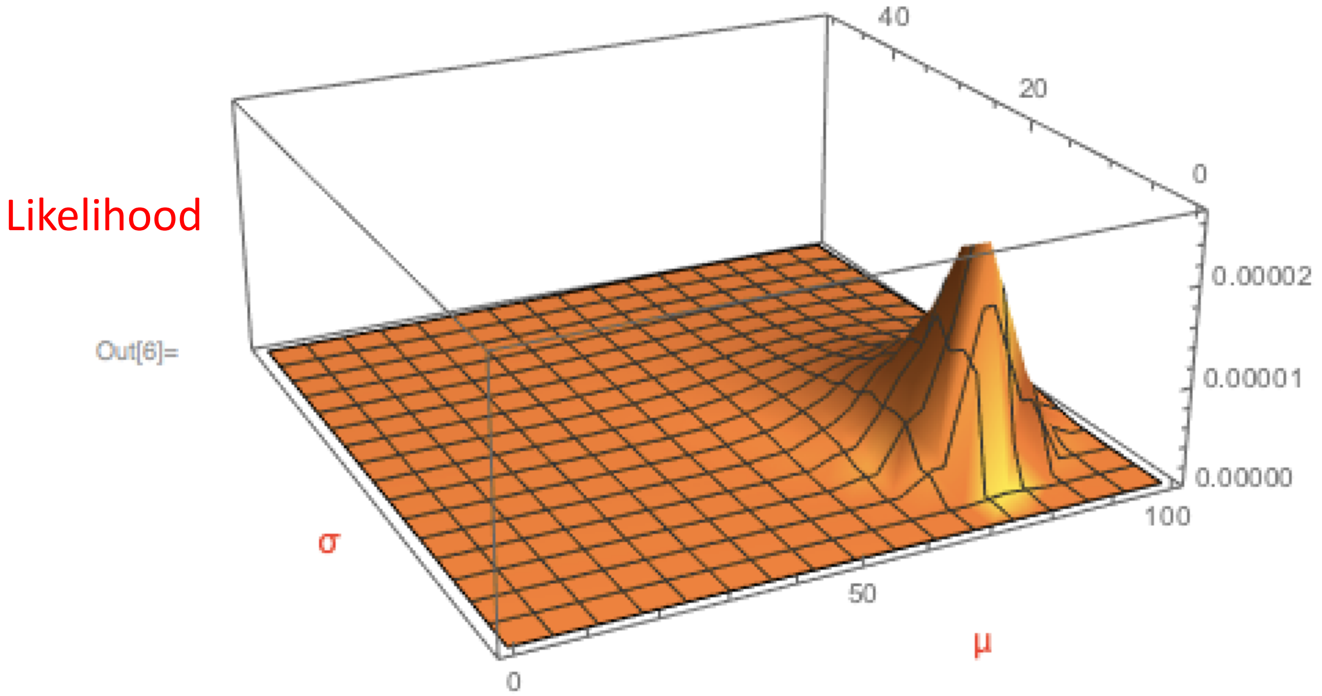
```python
plt.plot(np.linspace(1, 12, 100), list1, alpha=0.7, label='Likelihood function (mu=5)')
plt.xlabel('sigma_square')
plt.ylabel('Likelihood')
plt.legend(loc='lower right')
plt.axvline(6.8, 0, 0.97, color='red', linestyle='--', linewidth=2)
plt.show()
```

```python
sigma2 = 6.8
list2 = []
for i in np.linspace(1, 12, 100):
    mu = i
    Product = 1
    for x in [1,4,5,6,9]:

        y = (1 / np.sqrt(2 * np.pi * sigma2**2)) * np.exp(-(x-mu)**2 / (2 * sigma2**2))

        Product = Product * y

    list2.append(Product)
    print(Product)
```

[44]

```python
plt.plot(np.linspace(1, 12, 100), list2, alpha=0.7, label='Likelihood function (sigma_squre=6.8)')
plt.xlabel('mu (mean)')
plt.ylabel('Likelihood')
plt.legend(loc='lower left')
plt.axvline(5.0, 0, 0.97, color='red', linestyle='--', linewidth=2)
plt.show()
```

3D Likelihood function (Gaussian Dist.)



Reference: https://towardsdatascience.com/bayes-classifier-with-maximum-likelihood-estimation-4b754b641488

Thank you