Different parameter settings

1. DATA_N
   - origin

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 7/10
```

DATA_N = 50

   - DATA_N increases

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 9/10
By summing:  Model1 data, correct selection 9/10
             Model2 data, correct selection 9/10
```

DATA_N = 100

由大數法則可知，隨著試驗次數的增加，事件發生的頻率會趨於一個穩定值。所以 DATA_N 的增加能讓模型參數的估計更加穩定，使得正確區分模型的能力提高。

2. sampleRepeatNum
   - origin

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 7/10
```

sampleRepeatNum = 2000000

   - sampleRepeatNum decreases

```
By sampling: Model1 data, correct selection 6/10
             Model2 data, correct selection 4/10
By summing:  Model1 data, correct selection 8/10
             Model2 data, correct selection 5/10
```

sampleRepeatNum = 200

3. mu_prior_params – standard deviation
   - origin

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 7/10
```

SD = 4

   - SD increases

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 7/10
By summing:  Model1 data, correct selection 8/10
             Model2 data, correct selection 8/10
```

SD = 8

- SD decreases

```
By sampling: Model1 data, correct selection 7/10
             Model2 data, correct selection 4/10
By summing:  Model1 data, correct selection 5/10
             Model2 data, correct selection 5/10
```

SD = 2

我認為過於嚴格(標準差小)有可能會限制模型，造成表現變差；反之如果標準差較大的話，有助於模型靈活調整參數，提升推論效果。

4. sigma_prior_param_a
   - origin

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 7/10
```

a = 0.5

   - a increases

```
By sampling: Model1 data, correct selection 10/10
             Model2 data, correct selection 8/10
By summing:  Model1 data, correct selection 8/10
             Model2 data, correct selection 8/10
```

a = 1

   - a decreases

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 5/10
```

a = 0.4

5. sigma_prior_param_b
   - origin

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 5/10
By summing:  Model1 data, correct selection 7/10
             Model2 data, correct selection 7/10
```

b = 2

   - b increases

```
By sampling: Model1 data, correct selection 9/10
             Model2 data, correct selection 7/10
By summing:  Model1 data, correct selection 8/10
             Model2 data, correct selection 8/10
```

b = 8

- b decreases

```
By sampling: Model1 data, correct selection 7/10
             Model2 data, correct selection 4/10
By summing:  Model1 data, correct selection 5/10
             Model2 data, correct selection 5/10
```

b = 0.5

Gamma 分布的概率密度函數是：

$$f(x; a, b) = \frac{1}{\Gamma(a) b^a} x^{a-1} e^{-x/b} \quad \text{for } x > 0$$

其中：

- a 是形狀參數 (shape)
- b 是尺度參數 (scale)
- $\Gamma(a)$ 是 Gamma 函數

平均值：$\mathbb{E}[X] = a \cdot b$

變異數：$\mathrm{Var}(X) = a \cdot b^2$

所以當 a, b 變大時，變異數也會變大，模型選擇的效果會變好；反之 a, b 變小，效果變差。

Extra Credit

根據作業提供的相關公式來計算 2-component 的 Maximum likelihood，並且使用 Expectation-Maximization Algorithm 來實作

$$\mathrm{P}[x_i | M_2, \mu_2, \sigma_2] \;=\; (1-b)\, \mathcal{N}(x_i; \mu_{2a}, \sigma_{2a}) + (b)\, \mathcal{N}(x_i; \mu_{2b}, \sigma_{2b})$$

參考:[https://chih-sheng-huang821.medium.com/機器學習-em-演算法-expectation-maximization-algorithm-em-高斯混合模型-gaussian-mixture-model-gmm-和 gmm-em 詳細推導-c6f634410483](https://chih-sheng-huang821.medium.com/機器學習-em-演算法-expectation-maximization-algorithm-em-高斯混合模型-gaussian-mixture-model-gmm-和gmm-em詳細推導-c6f634410483)

```
Data generated with two components
generating data with:  m; (μ1,σ1); (μ2,σ2) =  0.078; (-1.11,1.63); (-0.83,3.15)
1-component maximum likelihood: 3.29161e-70
2-component mixture maximum likelihood: 8.63218e-55
Integrals by sampling= (6.588e-58,9.03338e-58)  by summing: (3.07579e-58,8.55344e-58)
```