In [2]:
```
import numpy as np
import pandas as pd
import math
```

In [67]:
```
"""
Question 1.2 EM Algorithm
"""
points = np.array([-67, -48, 6, 8, 14, 16, 23, 24])
K = 2

means = np.array([-67, 24])
variances = np.array([100, 100])
weights = np.array([0.5, 0.5])

def gaussian_pdf(x, mean, covar):
    return (1 / (math.sqrt(2 * math.pi * covar))) * math.exp(-((x - mean) **

def calculate_responsibility(x, k):
    k0 = weights[0] * gaussian_pdf(x, means[0], variances[0])
    k1 = weights[1] * gaussian_pdf(x, means[1], variances[1])
    print(f"Gaussian Density (x={x}, k=0) = {k0}, (x={x}, k=1) = {k1}")
    return (weights[k] * gaussian_pdf(x, means[k], variances[k])) / (k0 + k1

def update_weights(res, k):
    return res[:, k].mean()

def update_means(res, k):
    return np.sum(res[:, k] * points) / np.sum(res[:, k])

def update_variances(res, k):
    return np.sum(res[:, k] * (points - means[k]) ** 2) / np.sum(res[:, k])

def expectation_step():
    res = np.array([
        [
            calculate_responsibility(x, 0),
            calculate_responsibility(x, 1)
        ] for x in points
    ])
    print()
    for i, x in enumerate(points):
        for k in range(K):
            print(f"Responsibility (x={x}, k={k}) = {res[i, k]}")
    return res

def maximization_step(res):
    for k in range(K):
        weights[k] = update_weights(res, k)
        means[k] = update_means(res, k)
        variances[k] = update_variances(res, k)
        print(f"Weight (k={k}) = {weights[k]}")
        print(f"Mean (k={k}) = {means[k]}")
        print(f"Variance (k={k}) = {variances[k]}")
```

```
def em_algorithm():
    res = expectation_step()
    print()
    maximization_step(res)
```

In [68]:
```
em_algorithm()
```

```
Gaussian Density (x=-67, k=0) = 0.019947114020071634, (x=-67, k=1) = 2.07929
94895575653e-20
Gaussian Density (x=-67, k=0) = 0.019947114020071634, (x=-67, k=1) = 2.07929
94895575653e-20
Gaussian Density (x=-48, k=0) = 0.0032807907387338298, (x=-48, k=1) = 1.1039
949815685695e-13
Gaussian Density (x=-48, k=0) = 0.0032807907387338298, (x=-48, k=1) = 1.1039
949815685695e-13
Gaussian Density (x=6, k=0) = 5.3469189357708194e-14, (x=6, k=1) = 0.0039475
07915044707
Gaussian Density (x=6, k=0) = 5.3469189357708194e-14, (x=6, k=1) = 0.0039475
07915044707
Gaussian Density (x=8, k=0) = 1.2171602665145047e-14, (x=8, k=1) = 0.0055460
41733972778
Gaussian Density (x=8, k=0) = 1.2171602665145047e-14, (x=8, k=1) = 0.0055460
41733972778
Gaussian Density (x=14, k=0) = 1.1294047015771514e-16, (x=14, k=1) = 0.01209
8536225957168
Gaussian Density (x=14, k=0) = 1.1294047015771514e-16, (x=14, k=1) = 0.01209
8536225957168
Gaussian Density (x=16, k=0) = 2.1908197177546787e-17, (x=16, k=1) = 0.01448
4577638074137
Gaussian Density (x=16, k=0) = 2.1908197177546787e-17, (x=16, k=1) = 0.01448
4577638074137
Gaussian Density (x=23, k=0) = 5.139886785834457e-20, (x=23, k=1) = 0.019847
62737385059
Gaussian Density (x=23, k=0) = 5.139886785834457e-20, (x=23, k=1) = 0.019847
62737385059
Gaussian Density (x=24, k=0) = 2.0792994895575653e-20, (x=24, k=1) = 0.01994
7114020071634
Gaussian Density (x=24, k=0) = 2.0792994895575653e-20, (x=24, k=1) = 0.01994
7114020071634

Responsibility (x=-67, k=0) = 1.0
Responsibility (x=-67, k=1) = 1.04240617839016e-18
Responsibility (x=-48, k=0) = 0.9999999999663497
Responsibility (x=-48, k=1) = 3.3650271213502925e-11
Responsibility (x=6, k=0) = 1.354504930900902e-11
Responsibility (x=6, k=1) = 0.999999999986455
Responsibility (x=8, k=0) = 2.1946467857535377e-12
Responsibility (x=8, k=1) = 0.9999999999978053
Responsibility (x=14, k=0) = 9.335052443402414e-15
Responsibility (x=14, k=1) = 0.9999999999999907
Responsibility (x=16, k=0) = 1.5125188821494458e-15
Responsibility (x=16, k=1) = 0.9999999999999984
Responsibility (x=23, k=0) = 2.58967315791423e-18
Responsibility (x=23, k=1) = 1.0
Responsibility (x=24, k=0) = 1.04240617839016e-18
Responsibility (x=24, k=1) = 1.0

Weight (k=0) = 0.2499999999977625
Mean (k=0) = -57
Variance (k=0) = 90
Weight (k=1) = 0.7500000000022375
Mean (k=1) = 15
Variance (k=1) = 46
```

In [ ]: