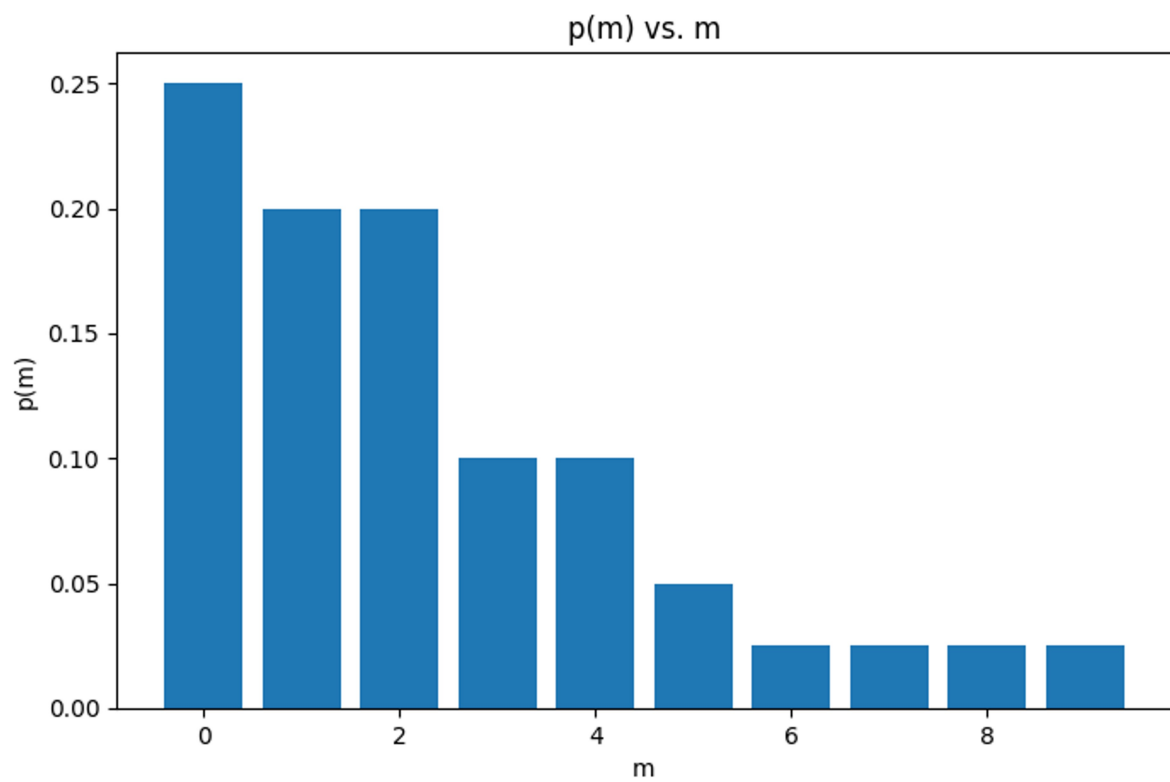


1.1

```
def prior(m):  
    dic = {  
        0: 0.25,  
        1: 0.2,  
        2: 0.2,  
        3: 0.1,  
        4: 0.1,  
        5: 0.05,  
        6: 0.025,  
        7: 0.025,  
        8: 0.025,  
        9: 0.025,  
    }  
    return dic[m]
```

2.1

```
m 0 prior(m) 0.25
m 1 prior(m) 0.2
m 2 prior(m) 0.2
m 3 prior(m) 0.1
m 4 prior(m) 0.1
m 5 prior(m) 0.05
m 6 prior(m) 0.025
m 7 prior(m) 0.025
m 8 prior(m) 0.025
m 9 prior(m) 0.025
```



3.1

```
def likelihood_single(x,y,m):  
    if m > 0:  
        f_mx = x**m  
    else:  
        f_mx = 0  
    return sp.stats.norm.pdf(y, f_mx, 0.1**2)
```

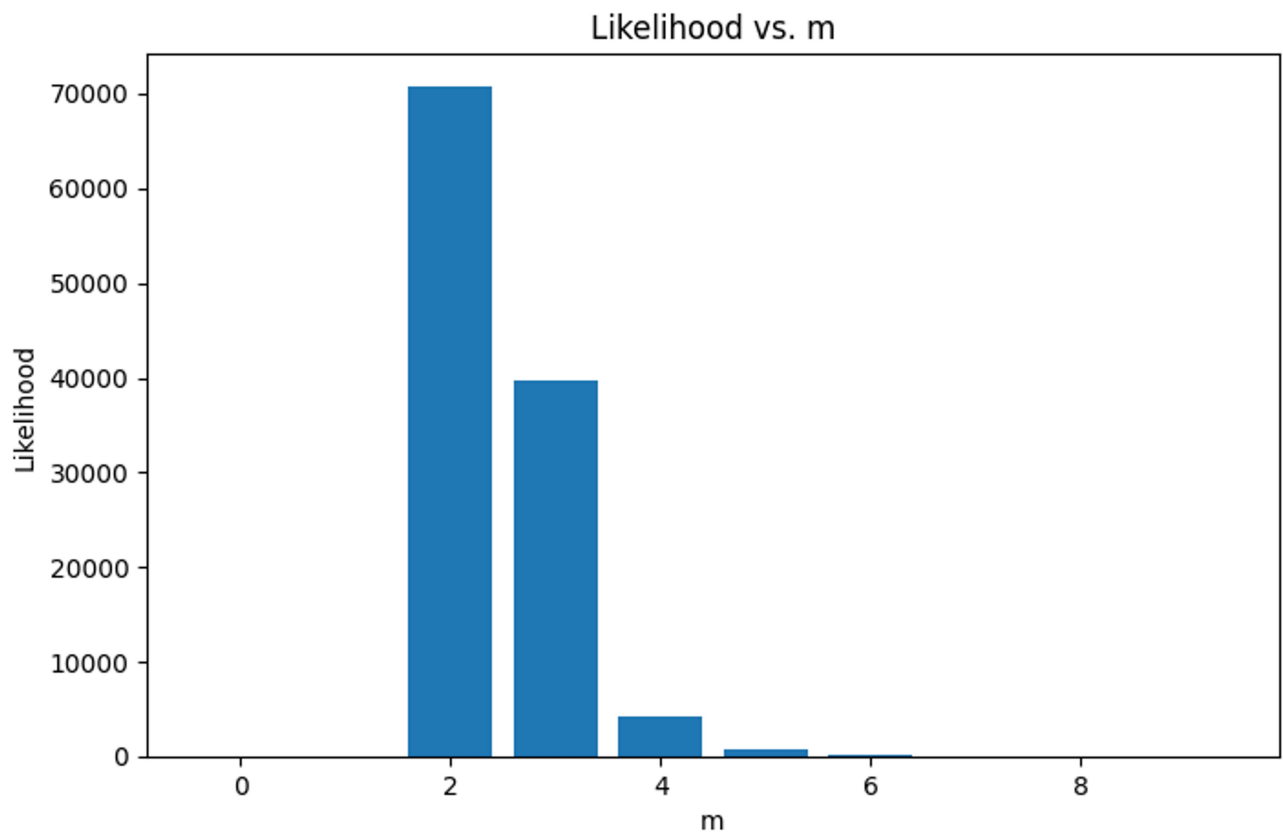
4.1

Thursday, December 1, 2022 4:40 PM

```
def likelihood(X,Y,m):  
    p = likelihood_single(X[0], Y[0], m)  
    for n in range(1, len(X)):  
        p = p * likelihood_single(X[n], Y[n], m)  
    return p
```

5.1

Thursday, December 1, 2022 5:51 PM



$$P(M=m | \text{Data}) = \frac{1}{P(\text{Data})} \times P(M=m) \times P(\text{Data} | M=m)$$

$$\sum_m^M P(M=m | \text{Data}) = 1 \quad M = \{0, 1, 2, \dots, m\}$$

$$\text{Let } \frac{1}{P(\text{Data})} = C, \quad X_m = P(M=m) \times P(\text{Data} | M=m)$$

$$\text{So, } P(M=m | \text{Data}) = C \times X_m$$

$$\therefore 1 = \sum_m^M C \times X_m \quad \text{or} \quad 1 = C \times \sum_m^M X_m$$

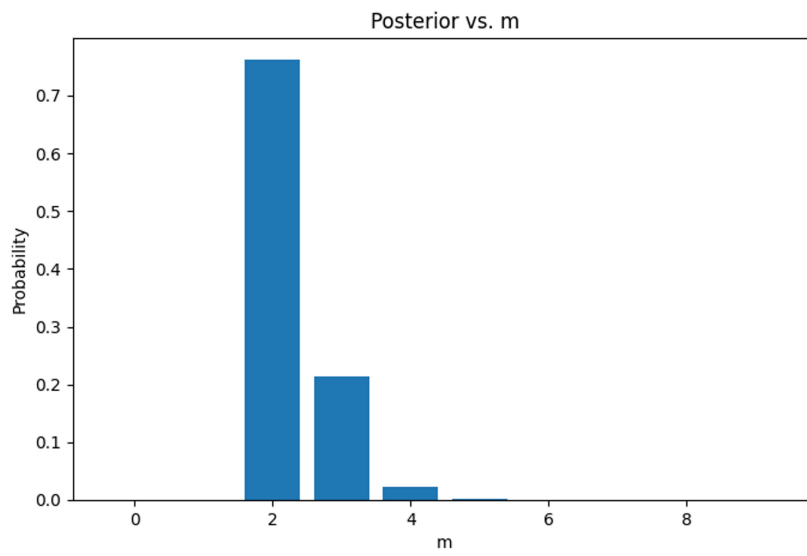
$$\frac{1}{C} = \sum_m^M X_m \quad \text{or} \quad \frac{1}{P(\text{Data})} = \sum_m^M P(M=m) \times P(\text{Data} | M=m)$$

$$\therefore P(M=m | \text{Data}) = \frac{P(M=m) \times P(\text{Data} | M=m)}{\sum_x^M P(\text{Data} | M=x) \cdot P(M=x)}$$

7.1

```
def posterior(X,Y,m):  
    p_data = 0  
    for i in range(10):  
        p_data += likelihood(X, Y, i) * prior(i)  
    return (prior(m)*likelihood(X, Y, m))/p_data
```

8.1



9.1

```
def MAP(X,Y):  
    ap = [posterior(X, Y, i) for i in range(10)]  
    return ap.index(max(ap))
```

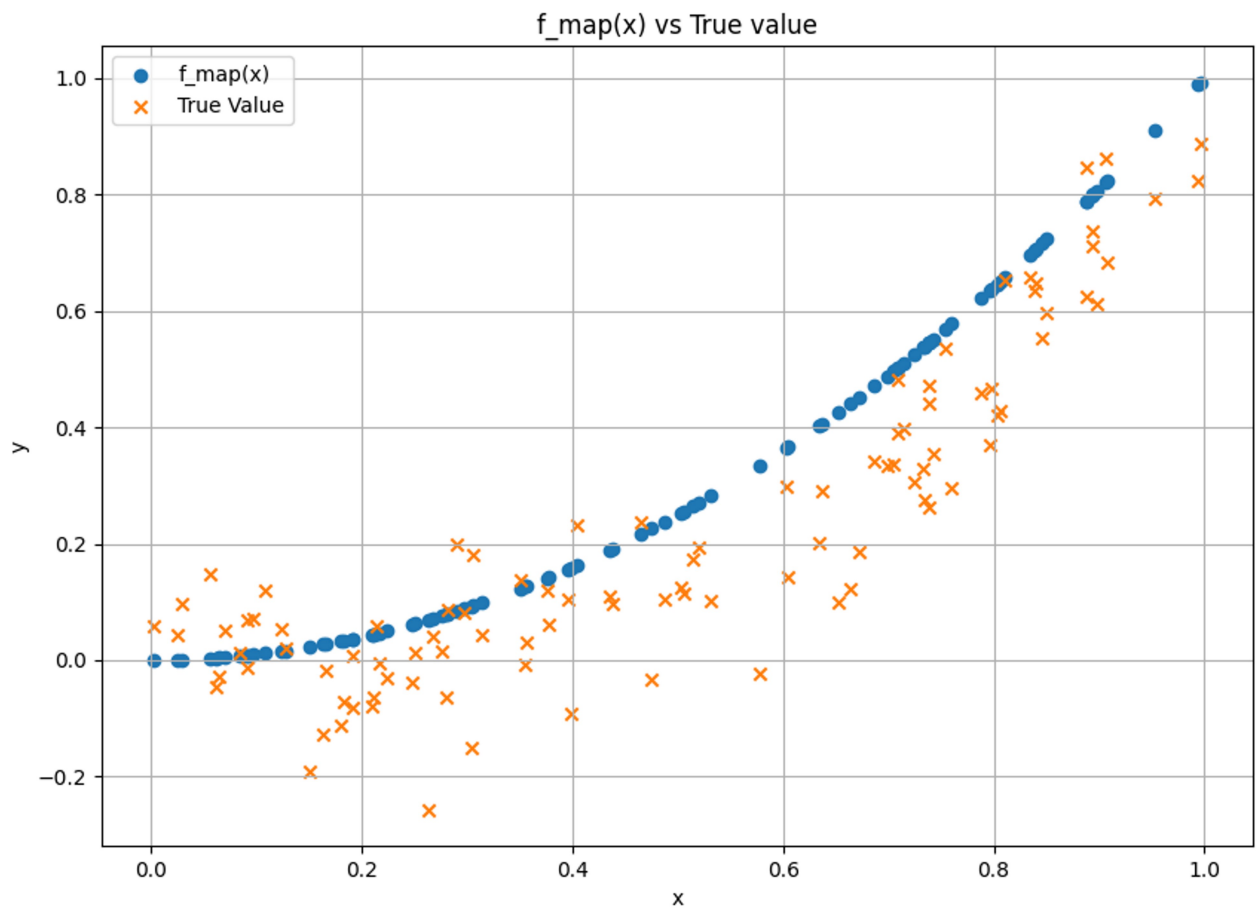
10.1

I got $m = 2$ and a posterior of 0.7613023972080507

11.1

```
def predict_MAP(x,X,Y):  
    if MAP(X, Y) == 0:  
        return 0  
    return x**(MAP(X, Y))
```

12.1



13.1

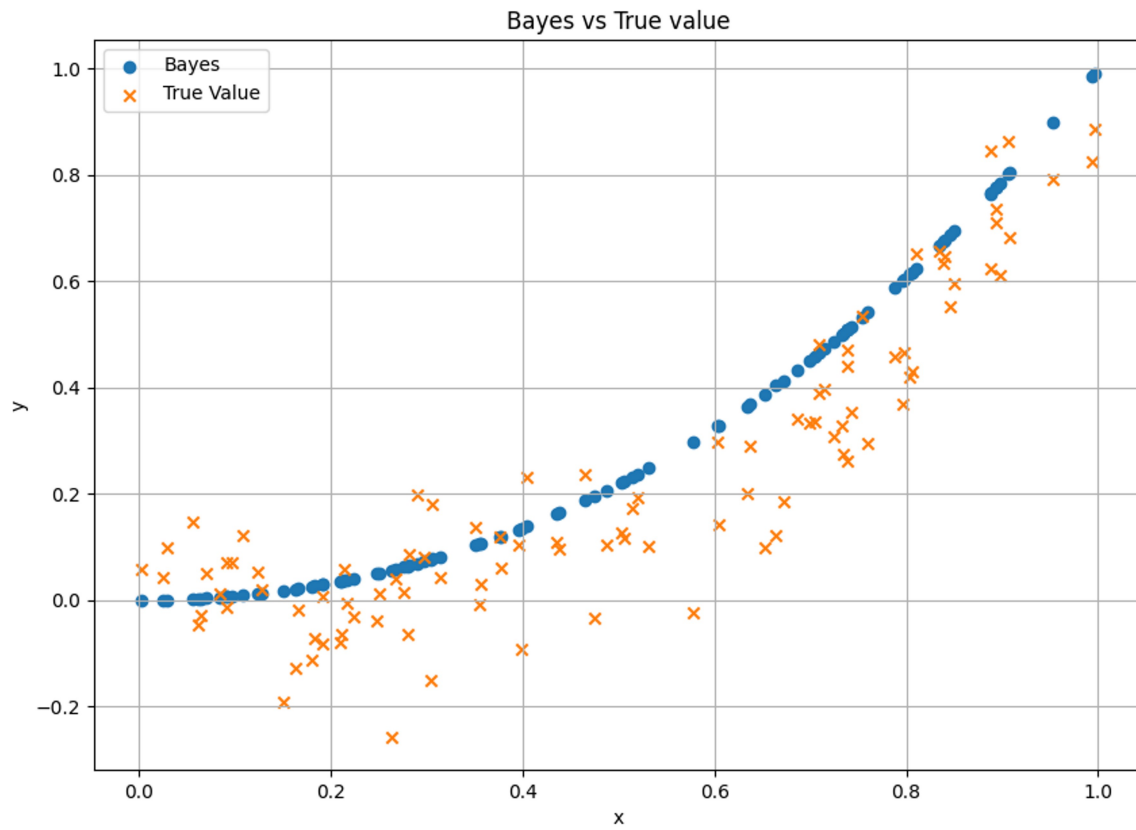
MSE = 0.02146290417007275

m = 2

14.2

```
def predict_Bayes(x,X,Y):  
    p_bayes = 0.0  
    for m in range(10):  
        if m == 0:  
            p_bayes += posterior(X, Y, m)*0  
        else:  
            p_bayes += posterior(X, Y, m)*(x**m)  
    return p_bayes
```

15.1



16.1

0.016372313860773324

17.1

Bayes error is lower. This makes sense because Bayes takes into account every model while MAP only chooses one model then predicts. MAP limits itself so it loses information thus making a worse prediction.