

Homework1

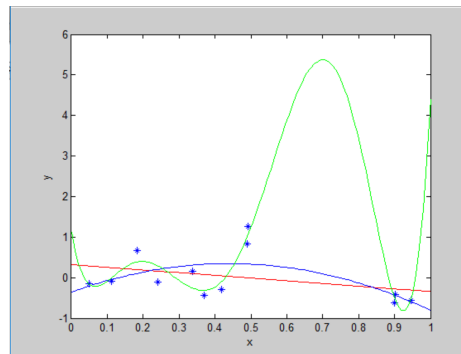
Yuchen Zhou yz2mf

1.2.

There are three representative data samples with mse and graph. The red line is degree1, blue is degree2, green is degree 6.

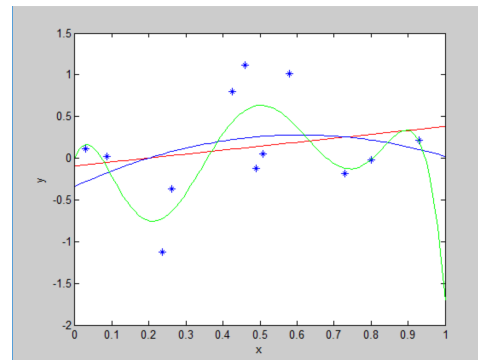
Data simple1

```
xy =  
  
Columns 1 through 7  
0.1839 0.2400 0.4173 0.0497 0.9027 0.9448 0.4909  
0.6757 -0.1124 -0.2945 -0.1469 -0.4240 -0.5601 1.2630  
  
Columns 8 through 12  
0.4893 0.3377 0.9001 0.3692 0.1112  
0.8277 0.1538 -0.6286 -0.4327 -0.0883  
  
>> mse  
  
mse =  
  
1.0000 2.0000 6.0000  
0.2908 0.2148 0.0494
```



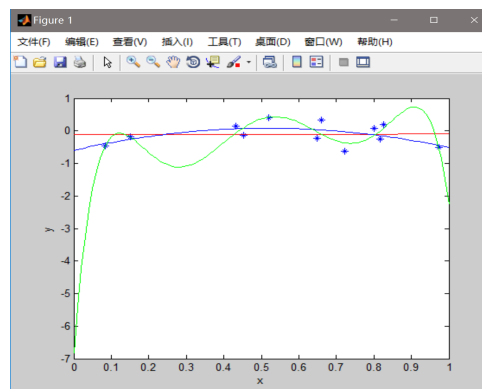
Data Sample2:

```
xy =  
  
Columns 1 through 7  
0.4242 0.5079 0.0855 0.2625 0.8010 0.0292 0.9289  
0.8020 0.0492 0.0207 -0.3671 -0.0154 0.1162 0.2132  
  
Columns 8 through 12  
0.7303 0.4886 0.5785 0.2373 0.4588  
-0.1864 -0.1182 1.0118 -1.1292 1.1147  
  
>> mse  
  
mse =  
  
1.0000 2.0000 6.0000  
0.3377 0.3222 0.1577
```



Data sample3:

```
xy =  
  
Columns 1 through 7  
0.8175 0.7224 0.1499 0.6596 0.5186 0.9730 0.6490  
-0.2523 -0.6353 -0.1913 0.3243 0.4129 -0.5075 -0.2355  
  
Columns 8 through 12  
0.8003 0.4538 0.4324 0.8253 0.0835  
0.0685 -0.1459 0.1509 0.2000 -0.4650  
  
>> mse  
  
mse =  
  
1.0000 2.0000 6.0000  
0.1051 0.0740 0.0439
```



From the graphs and data, it is easy to figure out that the more the degree is, the more mse is. Visual comparison of the three estimated models suggests that the linear model is closer to the

true function which will lead to better prediction. Therefore, small fitting error indicates small prediction error.

1.3.

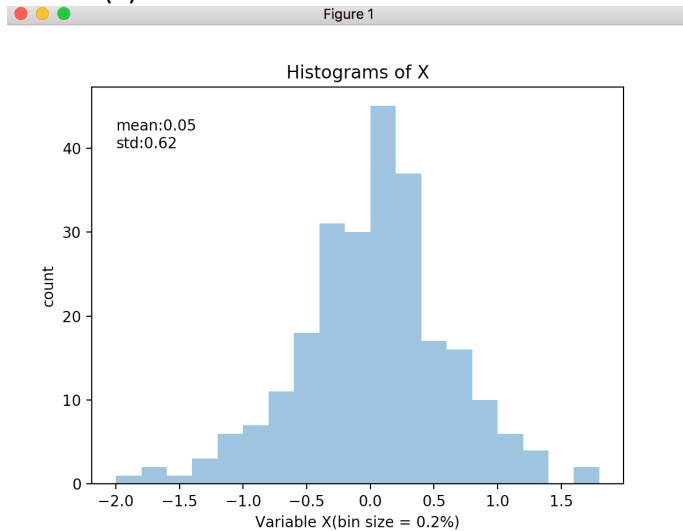
(a) 10^{10}

(b) $A(10,10) = 10!$

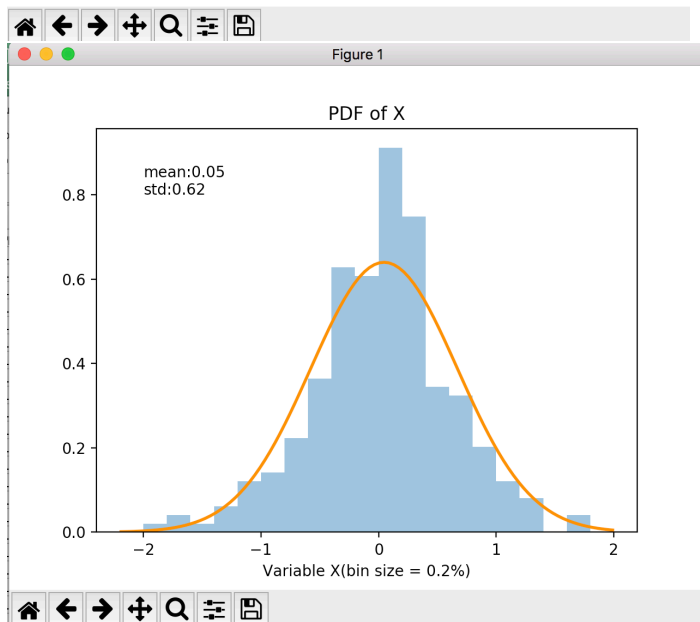
1.4 $P = 1 - (1 - 0.004)^{250} = 1 - 0.996^{250} \approx 0.633$

1.7 I use python to make the figures and fit the PDF.

(a)



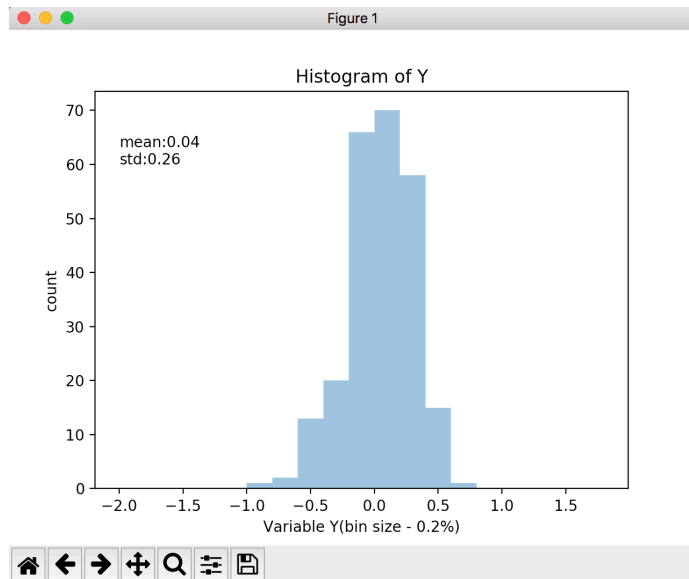
As it shows in the picture, the mean of the sample of X is 0.05 and the standard deviation is 0.62



It is the P.D.F. of X distribution

(b) As it shows in the picture, the mean of the sample of Y is 0.04 and the standard deviation is 0.26

Y are more concentrated then X.



1.13

I think these theories represent empirical knowledge. From the textbook, there are a few points contributing to empirical knowledge.

The first one is that the principle is derived from large set of data or experience. The Maya calendar, for example, is a complex system of calendars which comes from long-time observation, calculation and analyzation of various parameters.

The second one is useful and actionable. As the foundational aspects of Maya culture, there is no doubt the Maya calendar has been used for a long time. It is amazing that it can calculate the length of a year precisely.

To be honest, there is imperfection of these theories. Even in some way it acts in opposite with our modern science. However, the experience knowledge allows to make non-deterministic prediction while first-principal knowledge not.

And these theories are not pseudoscience since all of the imperfection should due to the lack of scientific equipment and the limitation of the times.