

Homework 1.6

October 30, 2020

1 hw1.6

```
[1]: import numpy as np
import pandas as pd
```

2 (a)

```
[2]: data = np.random.uniform(size = 1000)
data = pd.Series(data)
data.plot.hist()
None
```

It shows fairly uniform by the histogram.

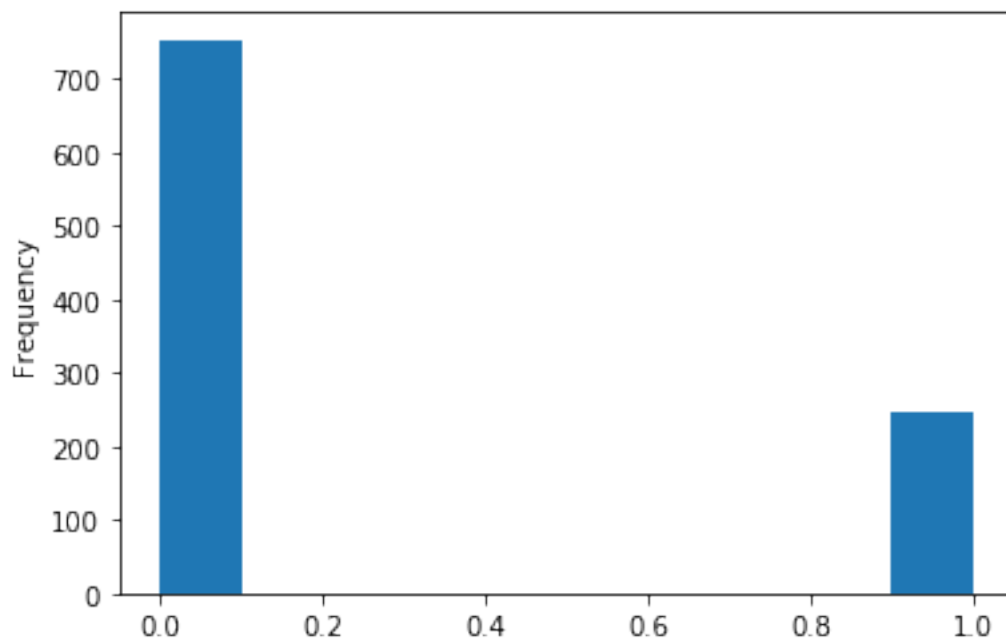
2.1 (b)

$$P(y_i = 1) = P(x_i \leq p) = p$$
$$P(y_i = 0) = P(x_i > p) = 1 - p$$

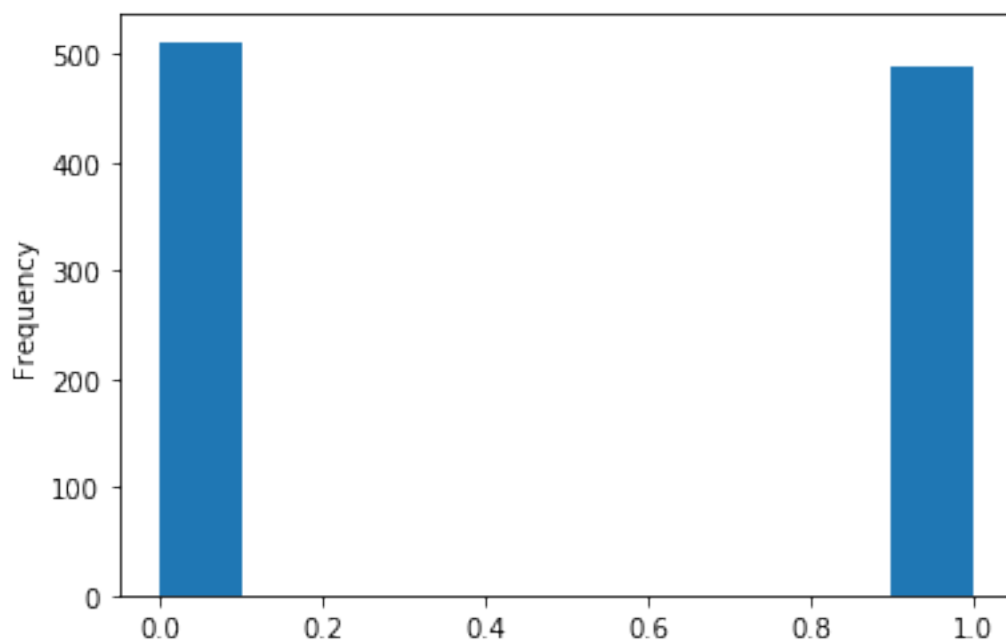
y_i is Bernoulli distribution.

2.2 (c)

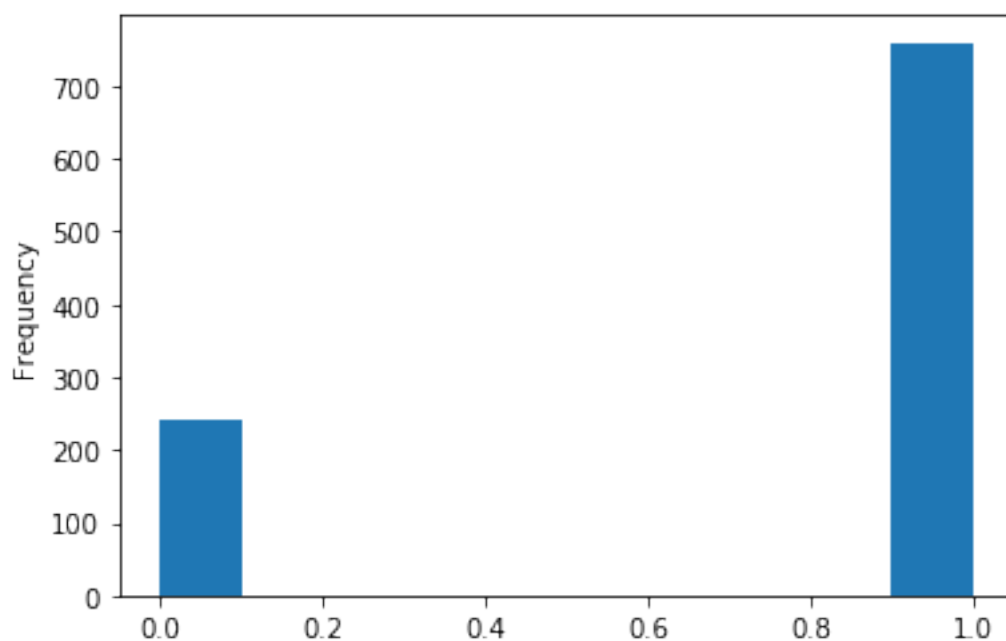
```
[3]: y1 = data.apply(lambda x: 1 if x <= 1/4 else 0)
y1.plot.hist()
None
```



```
[4]: y2 = data.apply(lambda x: 1 if x <= 1/2 else 0)
      y2.plot.hist()
      None
```



```
[5]: y3 = data.apply(lambda x: 1 if x <= 3/4 else 0)
y3.plot.hist()
None
```



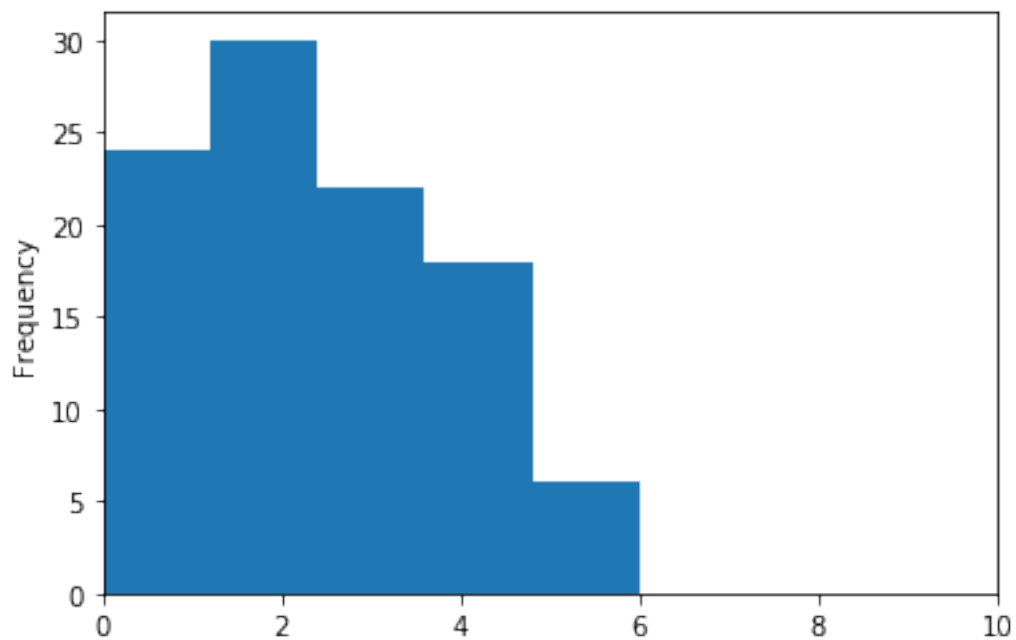
These histograms match with the conclusion from (b).

2.3 (d)

$$z_k = \sum_{i=1}^n y_{ki} \sim \text{Binomial}(n, p)$$

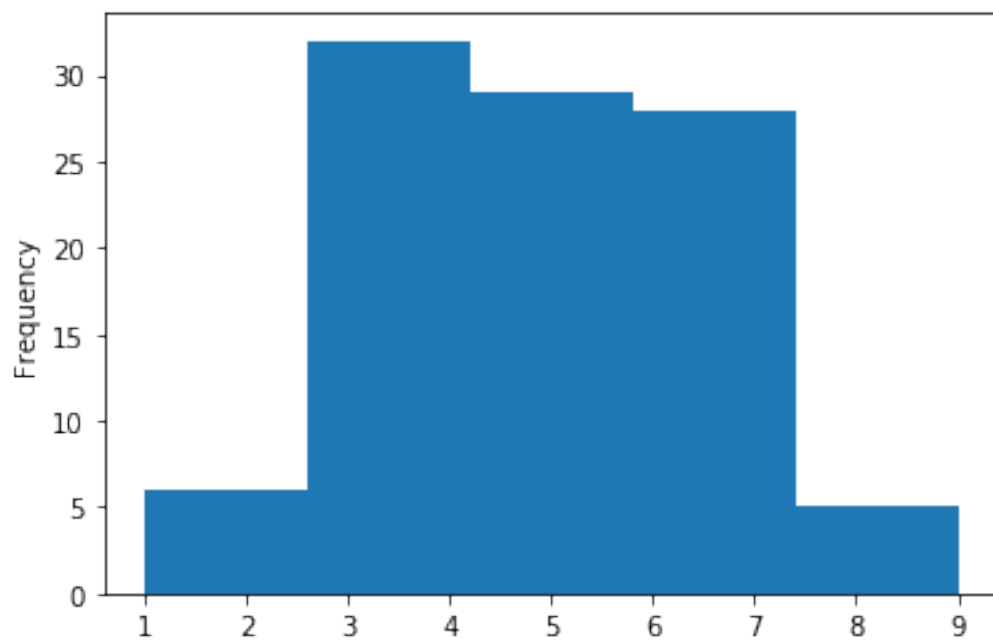
2.4 (e)

```
[11]: z1 = np.array(y1).reshape(-1,10).sum(1)
ax = pd.Series(z1).plot.hist(bins = 5)
ax.set_xlim([0,10])
None
```



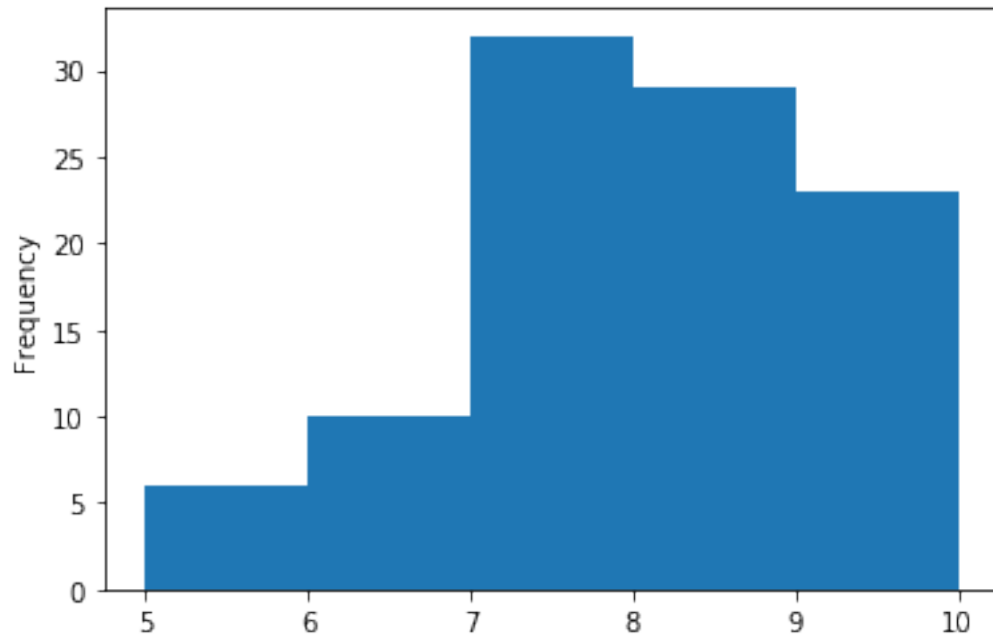
```
[12]: z2 = np.array(y2).reshape(-1,10).sum(1)
      pd.Series(z2).plot.hist(bins = 5)
      ax.set_xlim([0,10])
```

[12]: (0, 10)



```
[13]: z3 = np.array(y3).reshape(-1,10).sum(1)
pd.Series(z3).plot.hist(bins = 5)
ax.set_xlim([0,10])
```

[13]: (0, 10)



These histograms match the conclusion in (d). They are binomial distribution, with larger p , the data become more left skewed.

[]:

[]:

[]: