ps5_1

November 25, 2019

[278]:
```
import numpy as np
import pandas as pd
```

[279]:
```
dataset = pd.read_excel('ascending_dat.xlsx')
```

[280]:
```
dataset.head()
```

[280]:

	nbid	max_price
0	3	57.8926
1	3	37.7828
2	3	52.7520
3	3	57.4114
4	3	45.6665

[281]:
```
k = np.linspace(0,1,1000)
#print(k)
```

[282]:
```
sorted_dataset = dataset.sort_values(by=['nbid','max_price'])
sorted_dataset = sorted_dataset.reset_index()

sorted_dataset_3 = sorted_dataset[sorted_dataset['nbid'] == 3]
sorted_dataset_4 = sorted_dataset[sorted_dataset['nbid'] == 4]
sorted_dataset_5 = sorted_dataset[sorted_dataset['nbid'] == 5]
```

[283]:
```
listofindex = sorted_dataset_3.index.tolist()
probofindex = []
for x in listofindex:
    probofindex.append(x/len(listofindex))
print(len(probofindex))
sorted_dataset_3 = sorted_dataset_3.assign(prob=probofindex)
```

```
200
```

[284]:
```
listofindex = []
listofindex = sorted_dataset_4.index.tolist()
probofindex = []
for x in listofindex:
    probofindex.append((x-200)/len(listofindex))
print(len(probofindex))
```

```
sorted_dataset_4 = sorted_dataset_4.assign(prob=probofindex)
```

200

[285]:
```
listofindex = []
listofindex = sorted_dataset_5.index.tolist()
probofindex = []
for x in listofindex:
    probofindex.append((x-400)/len(listofindex))
print(len(probofindex))
sorted_dataset_5 = sorted_dataset_5.assign(prob=probofindex)
```

200

[286]:
```
def cdfcalc(t):
    for value in k:
        #print(3*(value**2)-2*(value**3))
        if abs(3*(value**2)-2*(value**3)- t) < 0.001 :
            return value
def cdfcalc2(t):
    for value in k:
        #print(4*(value**3)-3*(value**4)-t)
        if abs(4*(value**3)-3*(value**4)- t) < 0.001 :
            return value

def cdfcalc3(t):
    for value in k:
        if abs(5*(value**4)-4*(value**5)- t) < 0.001 :
            return value
cdfcalc3(0.2)
```

[286]: 0.5095095095095095

[287]:
```
sorted_dataset_3["cdfofvaluations"] = np.nan
for i in range(len(sorted_dataset_3['prob'])):
    prob = sorted_dataset_3.iloc[i]['prob']
    #print(cdfcalc(prob))
    #print(cdfcalc(prob))
    sorted_dataset_3['cdfofvaluations'][i] = cdfcalc(prob)
```

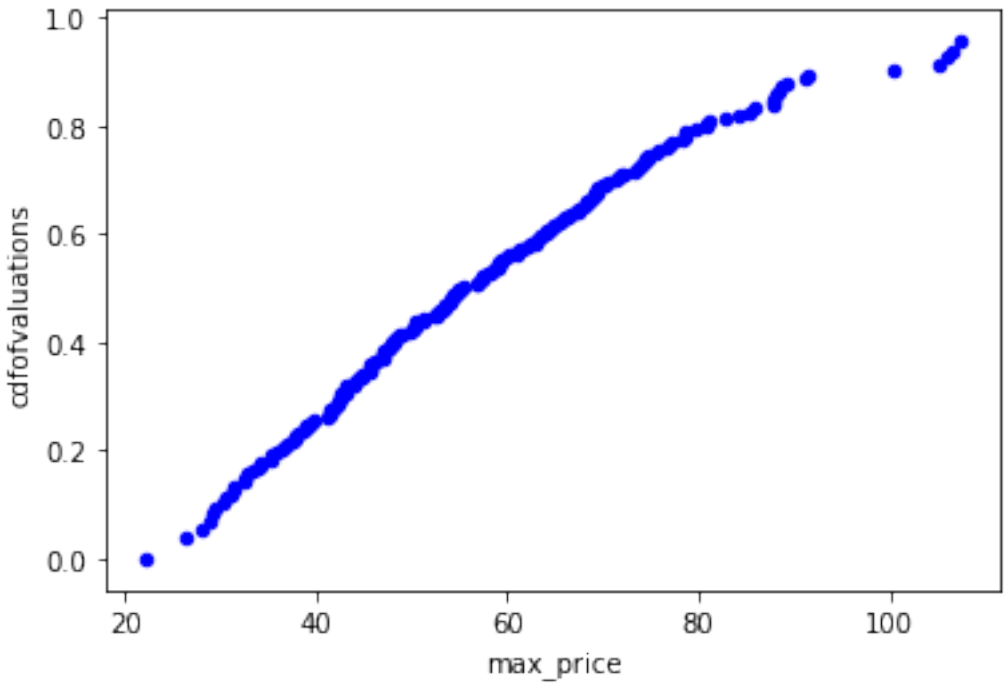
```
c:\users\satti\appdata\local\programs\python\python37\lib\site-
packages\ipykernel_launcher.py:6: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

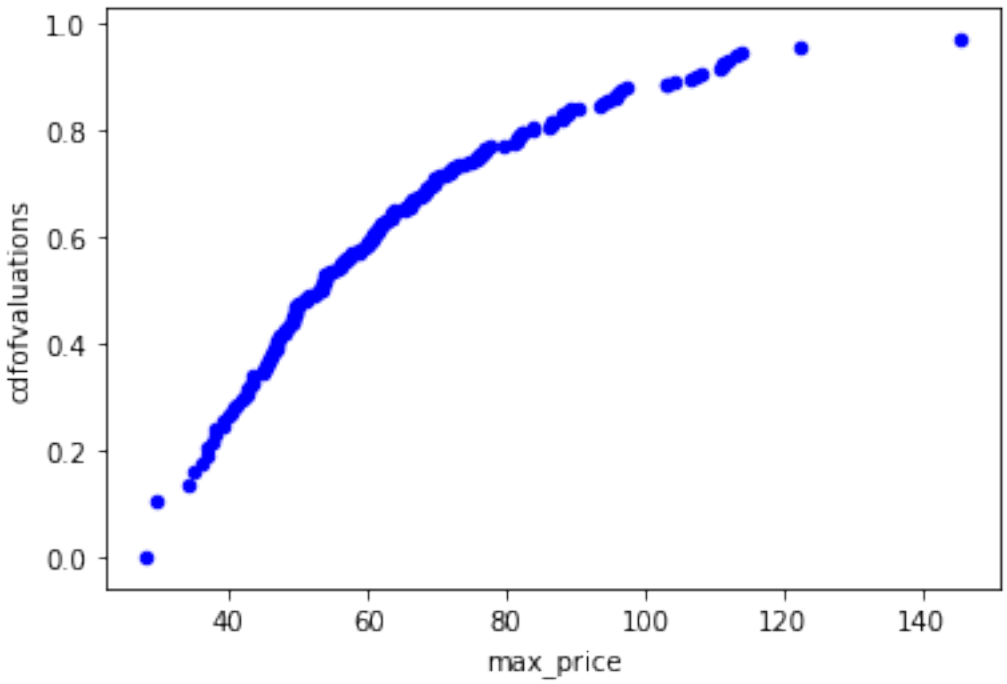
[288]:
```
sorted_dataset_4["cdfofvaluations"] = range(200)
calc2output = []
for i in range(len(sorted_dataset_4['prob'])):
    prob = sorted_dataset_4.iloc[i]['prob']
    #print(cdfcalc2(prob))
    calc2output.append(cdfcalc2(prob))
    #print(cdfcalc(prob))
sorted_dataset_4['cdfofvaluations'] = calc2output
```

[289]:
```
sorted_dataset_5["cdfofvaluations"] = range(200)
calc3output = []
for i in range(len(sorted_dataset_5['prob'])):
    prob = sorted_dataset_5.iloc[i]['prob']
    #print(cdfcalc5(prob))
    calc3output.append(cdfcalc3(prob))
    #print(cdfcalc(prob))
sorted_dataset_5['cdfofvaluations'] = calc3output
```

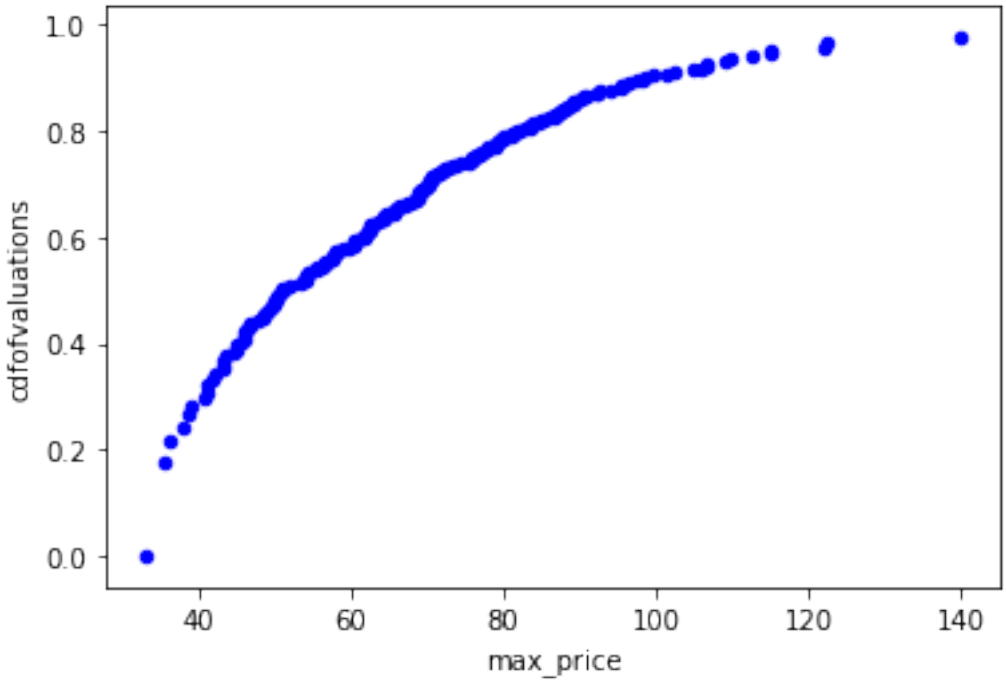
[290]:
```
import matplotlib.pyplot as plt
sorted_dataset_3.
 ↪plot(kind='scatter',x='max_price',y='cdfofvaluations',color='blue')
plt.show()
```



[291]:
```
import matplotlib.pyplot as plt
sorted_dataset_4.
 ↪plot(kind='scatter',x='max_price',y='cdfofvaluations',color='blue')
plt.show()
```



[292]:
```
import matplotlib.pyplot as plt
sorted_dataset_5.
 ↪plot(kind='scatter',x='max_price',y='cdfofvaluations',color='blue')
plt.show()
```


1.0
0.8
0.6
0.4
0.2
0.0
cdfofvaluations
40
60
80
100
120
140
max_price