/opt/pyenv-3.7.5/bin/jupyter notebook

In [ ]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

pandas' describe() function

In [ ]:
```
x = np.random.randn(1000)
```

In [ ]:
```
dataset = pd.DataFrame(x)
```

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

In [ ]:
```
dataset.describe()
```

$$\text{count} \leftrightarrow N$$
$$\text{mean} \leftrightarrow \bar{X},\ \text{std} \leftrightarrow s$$
$$25\% \leftrightarrow Q_1,\ 50\% \leftrightarrow M \equiv Q_2,\ 75\% \leftrightarrow Q_3$$

Μέτρα Ασυμμετρίας

Μέτρο ασυμμετρίας Pearson

$$\tilde{Sk}_p = \frac{3(\bar{X} - M)}{s}$$

In [ ]:
```
pearson = 3*(dataset.mean()-dataset.median())/dataset.std()
```

In [ ]:
```
pearson
```

Μέτρο ασυμμετρίας Bowley

$$Sk_b = \frac{(Q_3 - M) - (M - Q_1)}{Q_3 - Q_1}$$

In [ ]:
```
Q_1 = dataset.quantile(0.25)
M = dataset.median()
Q_3 = dataset.quantile(0.75)
bowley = ((Q_3 - M) - (M - Q_1))/(Q_3-Q_1)
```

In [ ]:
```
bowley
```

Καμπύλη Lorenz - Συντελεστής Gini

In [ ]:
```
wage = np.array([500, 1000, 1000, 1000, 2000, 2000])

wage = np.insert(wage, 0, 0.0)
```

In [ ]:
```
df = pd.DataFrame(wage, columns=['x'])
```

In [ ]:
```
df.describe()
```

$$\Phi_n = \frac{\sum_{j=1}^{n} x_j}{\sum_{j=1}^{N} x_j}$$

In [ ]:
```
df['Phi'] = df['x'].cumsum()/df['x'].sum()
```

$$RF_n = \frac{n}{N}$$

In [ ]:
```
df['RF'] = np.array(range(len(wage)))/(len(wage)-1)
```

$$\{(RF_n, \Phi_n)\}_{n=1}^{N}$$

In [ ]:
```
df.plot(x='RF',y='Phi')
plt.plot((0,1),(0,1))
```

In [ ]:
```
df
```

$$\Sigma\Phi_n = \Phi_n + \Phi_{n-1}$$

In [ ]:
```
SPhi = df['Phi'][1:].values + df['Phi'][:-1].values
```

In [ ]:
```
SPhi = np.insert(SPhi, 0, 0.0)
```

In [ ]:
```
df['SPhi'] = SPhi
```

$$\Delta RF_n = RF_n - RF_{n-1}$$

In [ ]:
```
DRF = df['RF'][1:].values - df['RF'][0:-1].values
```

In [ ]:
```
DRF = np.insert(DRF, 0, 0.0)
```

In [ ]:
```
df['DRF'] = DRF
```

In [ ]:
```
df['SPhi_DRF'] = df['SPhi'] * df['DRF']
```

In [ ]:
```
df
```

In [ ]:
```
Gini = 1 - df['SPhi_DRF'].sum()
```

In [ ]:
```
Gini
```

In [ ]:
```

```