hw3 Problem2

October 23, 2018

1 Assignment 3

1.0.1 MACS 30000, Dr. Evans

1.0.2 Zeyu Xu

Due Wednesday, Oct. 24 at 11:30 AM

```
In [1]: # Import packages
        import numpy as np
        import matplotlib.pyplot as plt
```

1.0.3 1. Simulation in Sociology, Moretti (2002)

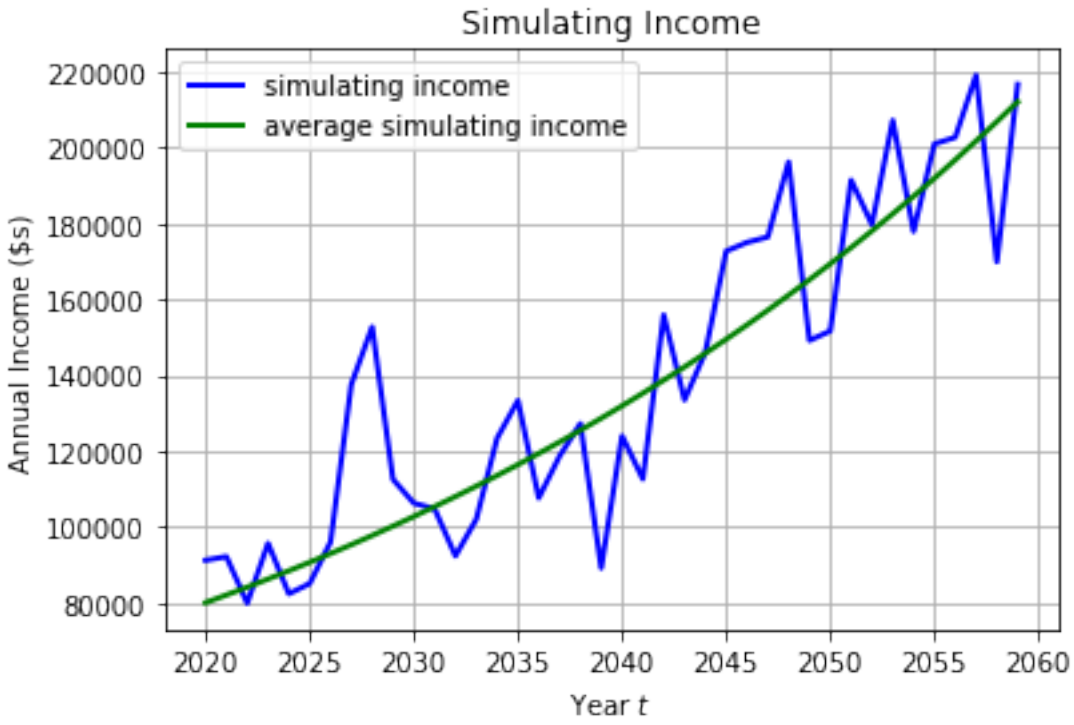
See the attached pdf.

1.0.4 2. Simulating your income

(a) Here is the code to define regression functions, simulate errors, predict incomes and plot lifetime income paths.

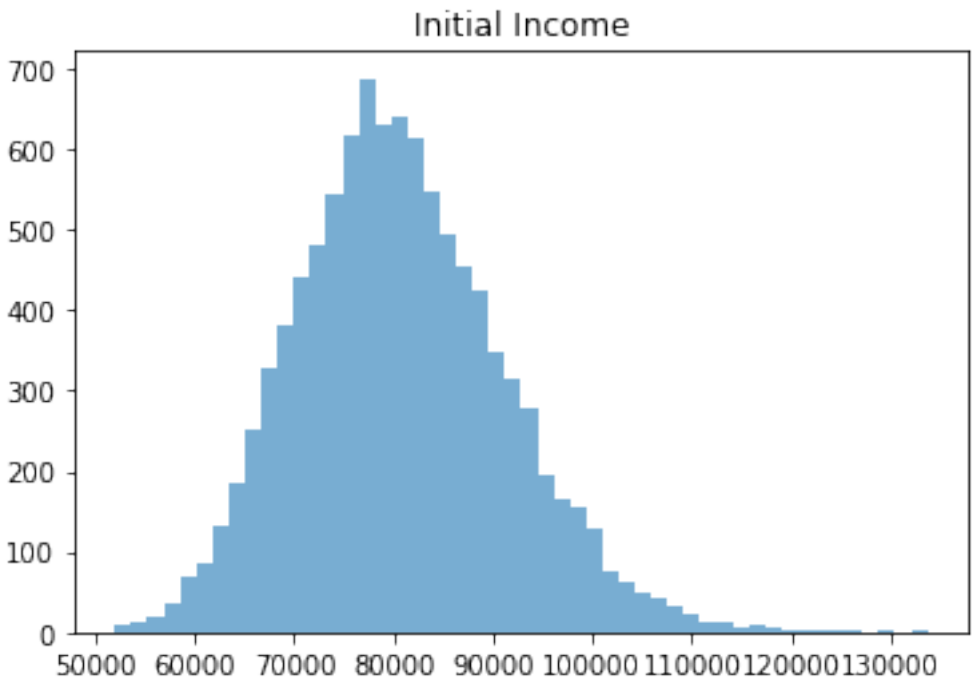
```
In [2]: sigma=0.13
        inc0 = 80000
        g = 0.025
        def predict(t, income, error, rho=0.4, initincome = inc0):
            lnincome = (1 - rho) * (np.log(initincome) + g * (t-2020)) + rho * np.log(income) +
            return np.exp(lnincome)
        Error = np.zeros((10000,40))
        for i in range(10000):
            Error[i,:] = np.exp(sigma * np.random.randn(40))
        income = np.ones((10000,40))
        for i in range(10000):
            income[i,0] = np.exp(np.log(inc0)+np.log(Error[i,0]))
            for j in range(1,40):
                income[i,j] = predict(2020+j, income[i,j-1], Error[i,j])
        time = np.array(list(range(2020,2060)))
        Income0 = np.array([np.exp(np.log(inc0) + g * (t-2020)) for t in time])
        fig, ax = plt.subplots()
        ax.plot(time, income[0,:], 'b-', label = 'simulating income', linewidth=2)
```

```
ax.plot(time, Income0, 'g-', label = 'average simulating income', linewidth=2)
ax.set_title('Simulating Income')
ax.set_xlabel(r'Year $t$')
ax.set_ylabel(r'Annual Income (\$s)')
ax.legend()
ax.grid()
plt.show()
```



(b) Here is the code to plot a histogram with 50 bins of year t = 2020 initial income for each of the 10,000 simulations and find out the required percentage.

```
In [3]: fig2, ax2 = plt.subplots()
        ax2.hist(income[:,0], alpha = 0.6, bins = 50)
        ax2.set(title = "Initial Income")
        plt.show()
        Pct1 = len([x for x in income[:,0] if x > 100000])/10000
        Pct2 = len([x for x in income[:,0] if x < 70000])/10000
        print("percent of the class earn more than $100,000 in the first year out of the progra
        print("percent of the class earn less than $70,000 in the first year out of the progra
```



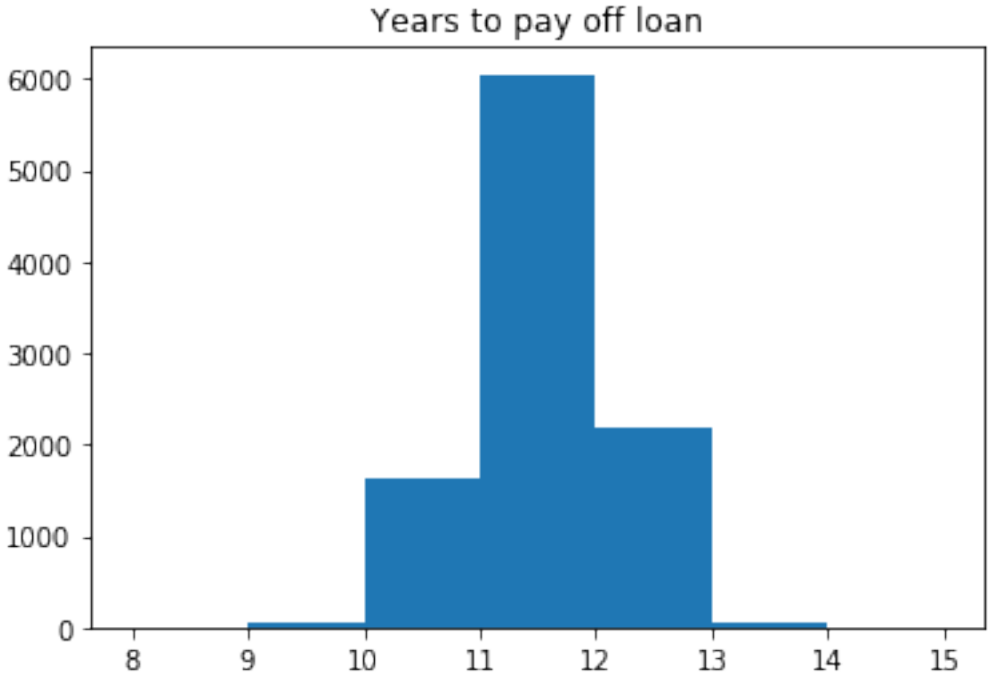

```
percent of the class earn more than $100,000 in the first year out of the program:  0.0422
percent of the class earn less than $70,000 in the first year out of the program:  0.1514
```

As shown above: 4.22% of my class will earn more than $100,000 in the first year out of the program. 15.14% of my class will earn less than $70,000 in the first year out of the program. Judge from the histogram, the distribution normally distributed.

(c) Here is the code to plot the histogram of how many years it takes to pay off the loan in each of my 10,000 simulations and find out the required percentage.

```
In [4]: def pay_year(income):
            loan = 95000
            i = 0
            while i <= 40 and loan > 0:
                loan -= income[i] * 0.1
                i += 1
            return i
        Pay_year = np.zeros((10000))
        for i in range(10000):
            Pay_year[i] = pay_year(income[i,:])
        fig3, ax3 = plt.subplots()
        ax3.hist(Pay_year, np.arange(min(set(Pay_year))-1,max(set(Pay_year))+3))
        ax3.set(title = "Years to pay off loan")
        plt.show()
```

```
Pct3 = len([x for x in Pay_year if x <= 10])/10000
print("percent of the simulations able to pay off the loan in 10 years: ",Pct3)
```



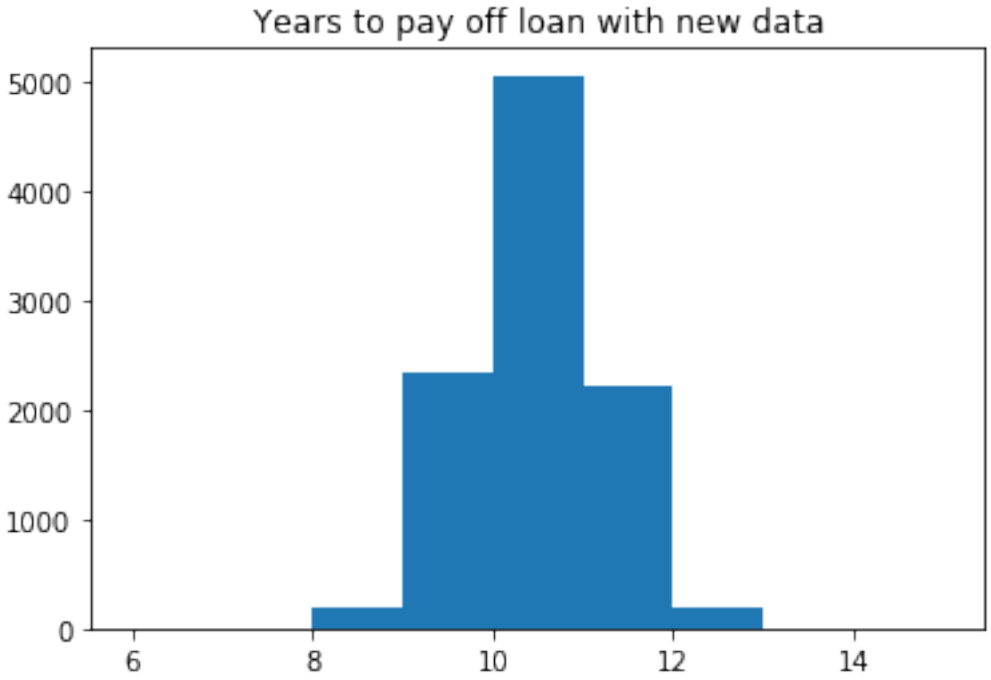

```
percent of the simulations able to pay off the loan in 10 years:  0.1689
```

16.89% of the simulations are able to pay off the loan in 10 years.

(d) Correlation matrix for the now six variables

```
In [5]: sigma_new =0.17
        inc0_new = 90000
        Error_new = np.zeros((10000,40))
        for i in range(10000):
            Error_new[i,:] = np.exp(sigma_new * np.random.randn(40))
        income_new = np.ones((10000,40))
        for i in range(10000):
            income_new[i,0] = np.exp(np.log(inc0_new)+np.log(Error_new[i,0]))
            for j in range(1,40):
                income_new[i,j] = predict(2020+j, income_new[i,j-1], Error_new[i,j], initincome
        Pay_year_new = np.zeros((10000))
        for i in range(10000):
            Pay_year_new[i] = pay_year(income_new[i,:])
        fig4, ax4 = plt.subplots()
        ax4.hist(Pay_year_new, np.arange(min(set(Pay_year_new))-1,max(set(Pay_year_new))+3))
```

```
ax4.set(title = "Years to pay off loan with new data")
plt.show()
Pct4 = len([x for x in Pay_year_new if x <= 10])/10000
print("percent of the simulations able to pay off the loan in 10 years: ",Pct4)
```




```
percent of the simulations able to pay off the loan in 10 years:  0.7593
```

75.93% of the simulations are able to pay off the loan in 10 years.