STOCHASTIC PROCESSES

HW02

Hanie Hatami(99100614)
Stochastic processes course

Contents

1 Abstract

In this exercise, we want to first draw the graphs of $XP(X)$, $X^2P(X)$, $X^3P(X)$,... in terms of X for an arbitrary number of data and bins and check their behavior for different data. The reason for this is to see that for any number of data it is reasonable to calculate up to the which highest order. Then, in the second part of the exercise, we go to the moments, and draw the graph of different moments in terms of the number of data and check them.

2 Results

First, we need to draw the graphs of $XP(X)$, $X^2P(X)$, $X^3P(X)$,... in terms of X for the number of different data with normal distribution.

2.1 Question 1

Here we only give a limited number of orders for each number (as much as it was necessary to show the behavior of the orders.) To see more orders, just enter your desired order value when prompted in the code to display the graphs. display for you.
According to our expectation, the $X^0P(X)$ graph should be the same as the data histogram because it is basically the $P(X)$ that is being drawn.

2.1.1 Normal distribution n=10^2:

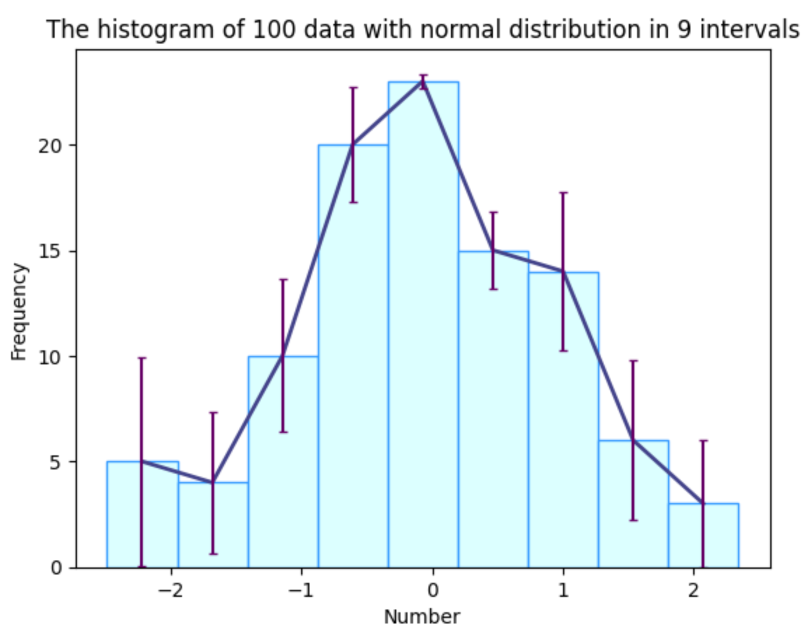

Figure 1: The histogram of 10^2 data with normal distribution in 9 intervals

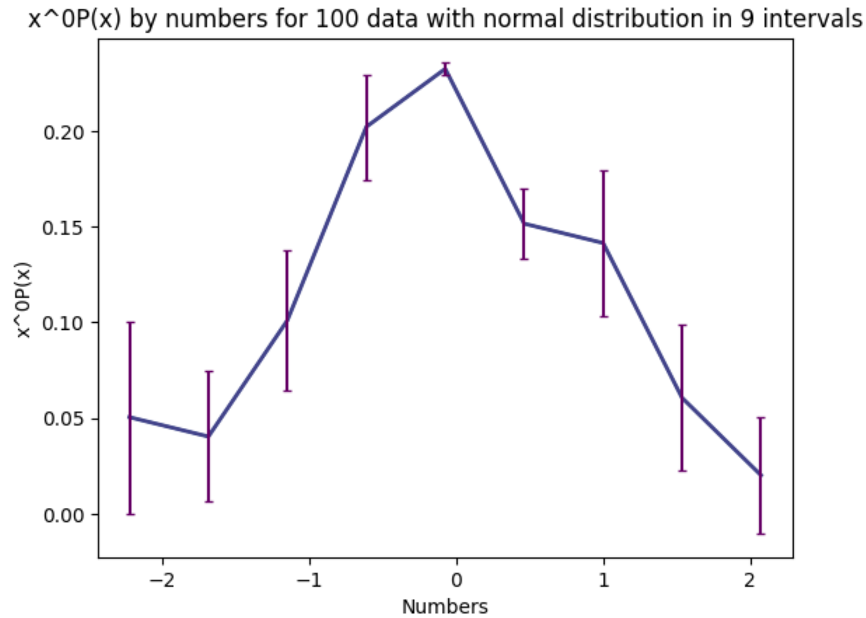

Figure 2: $X^0P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

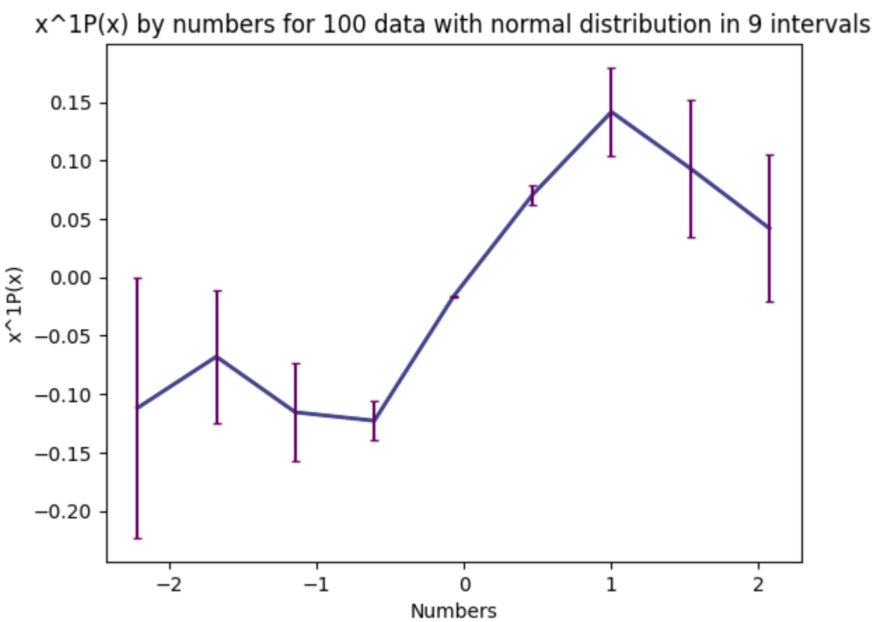

Figure 3: $X^1P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

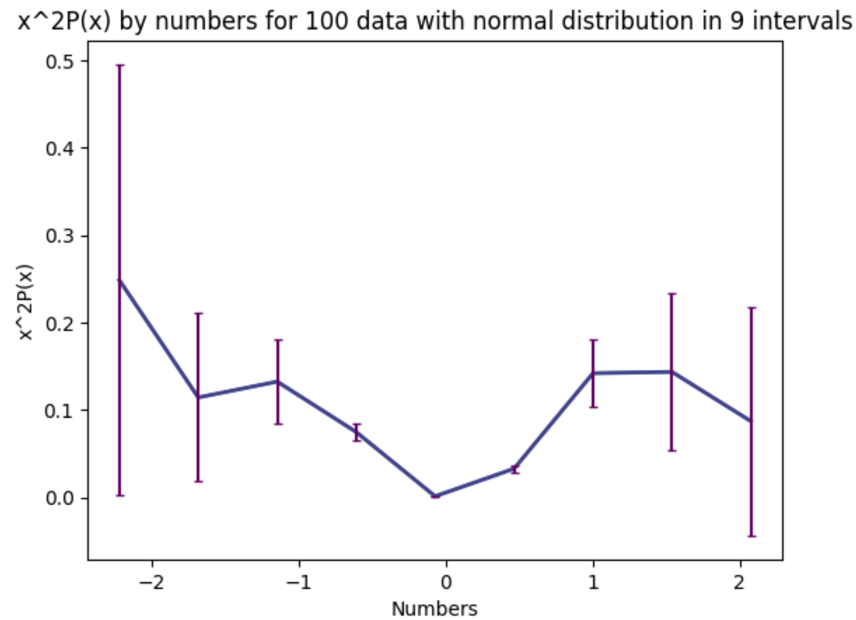

Figure 4: $X^2P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

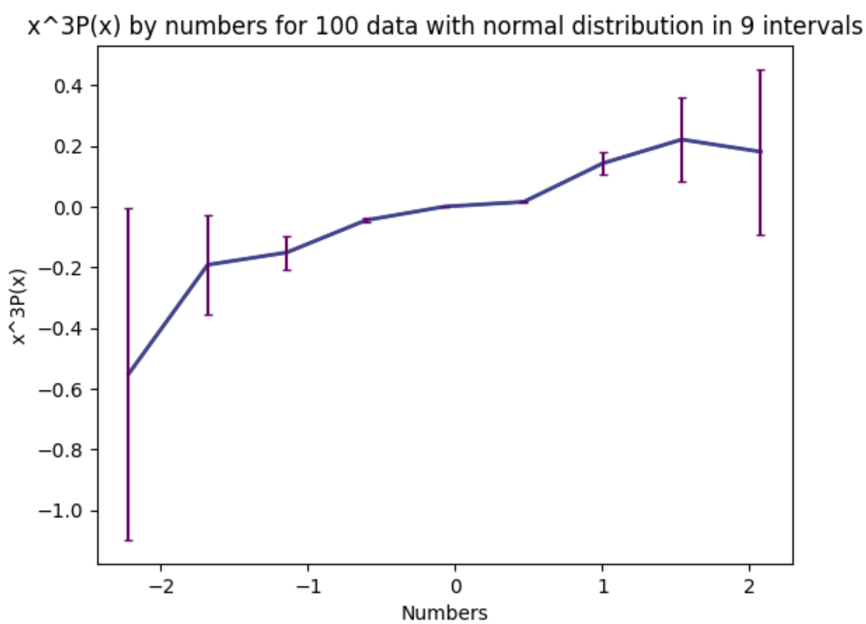

Figure 5: $X^3P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

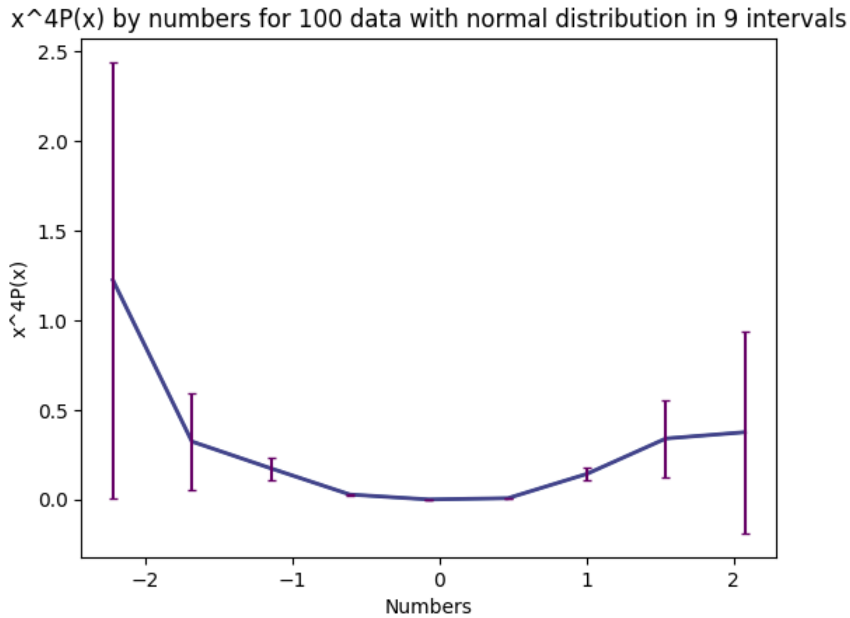

Figure 6: $X^4P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

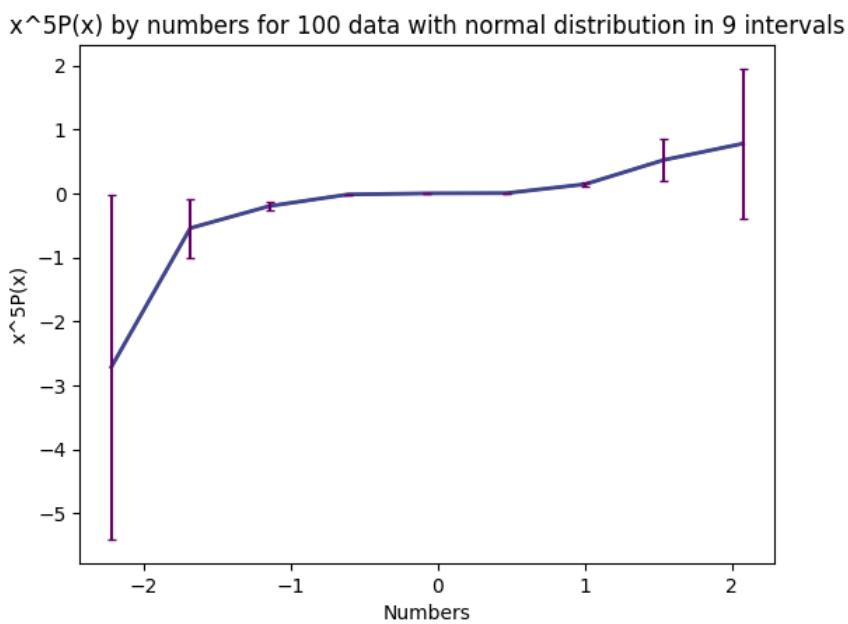

Figure 7: $X^5P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

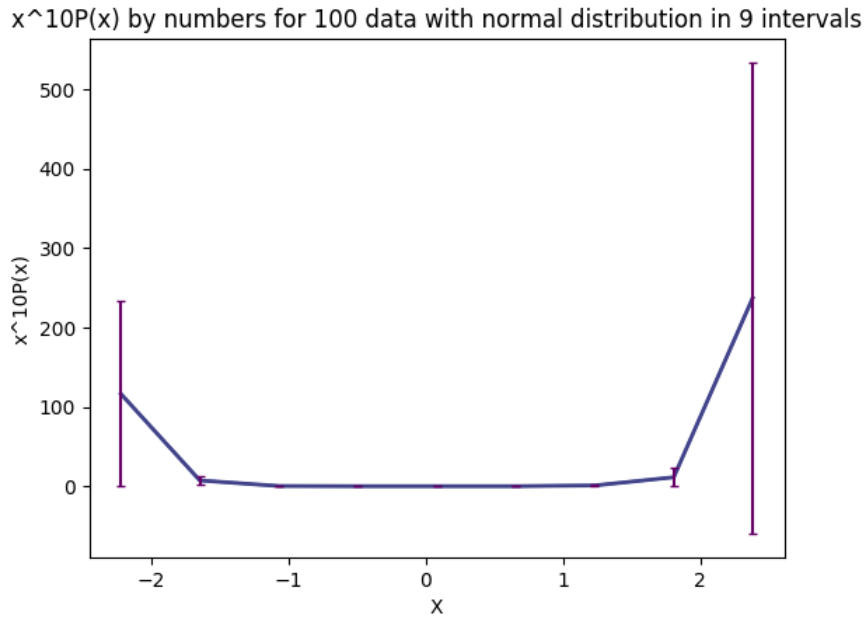

Figure 8: $X^(10)P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

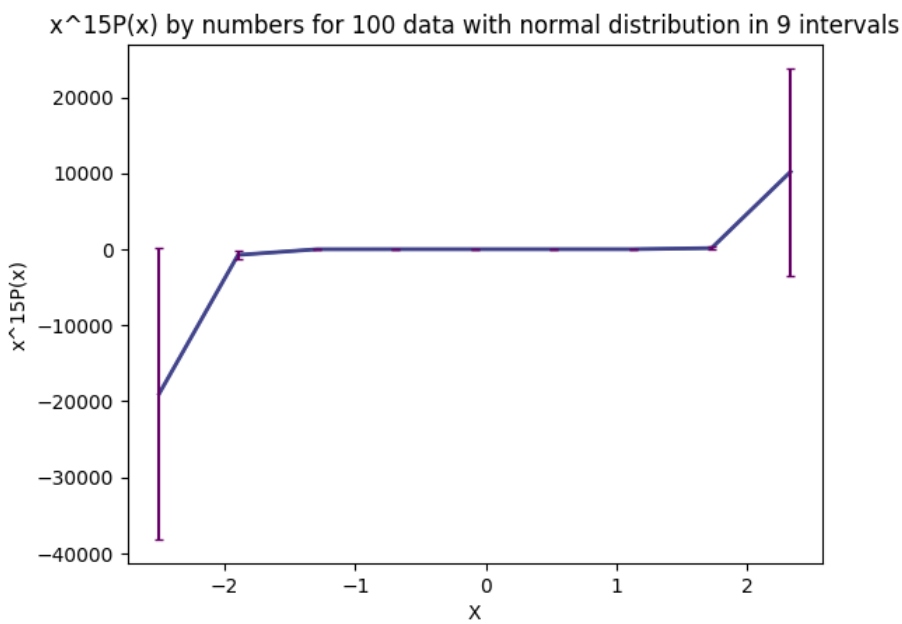

Figure 9: $X^(15)P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

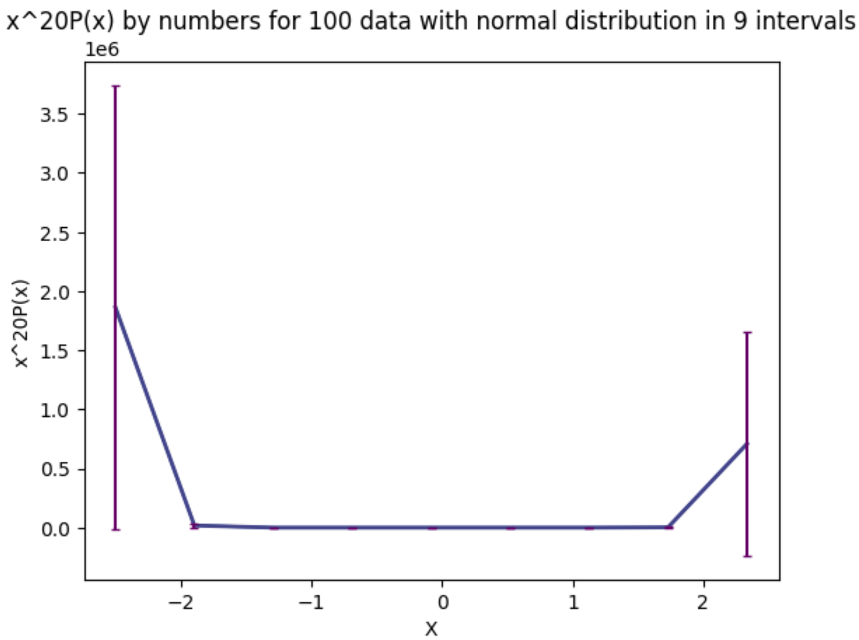

Figure 10: $X^(20)P(X)$ according to X for 10^2 data with normal distribution in 9 intervals

As we can see in the graphs, for 10^2 data, we can hardly calculate $X^2P(X)$, and from then on, all graphs show exponential behavior in the corners and the errors become large.

2.1.2 Normal distribution n=10^3:

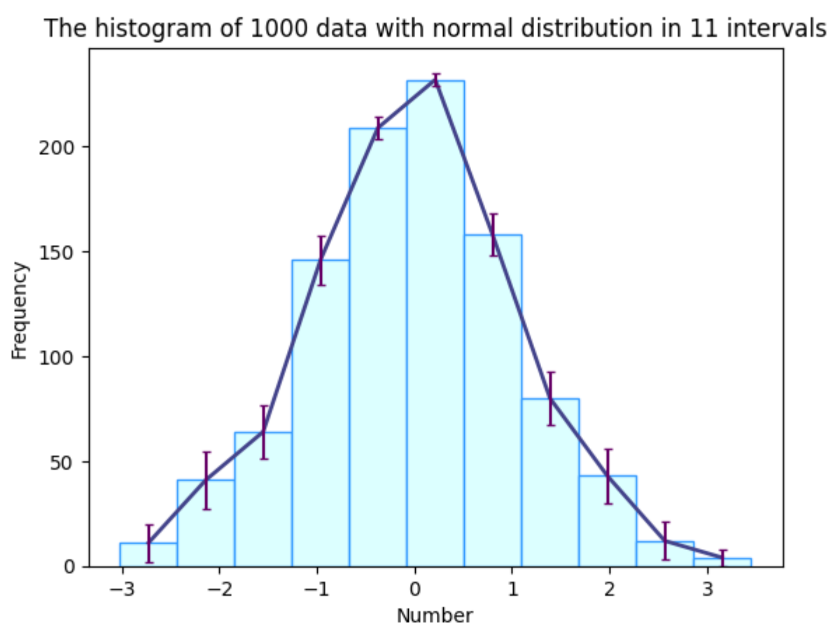

Figure 11: The histogram of 10^3 data with normal distribution in 11 intervals

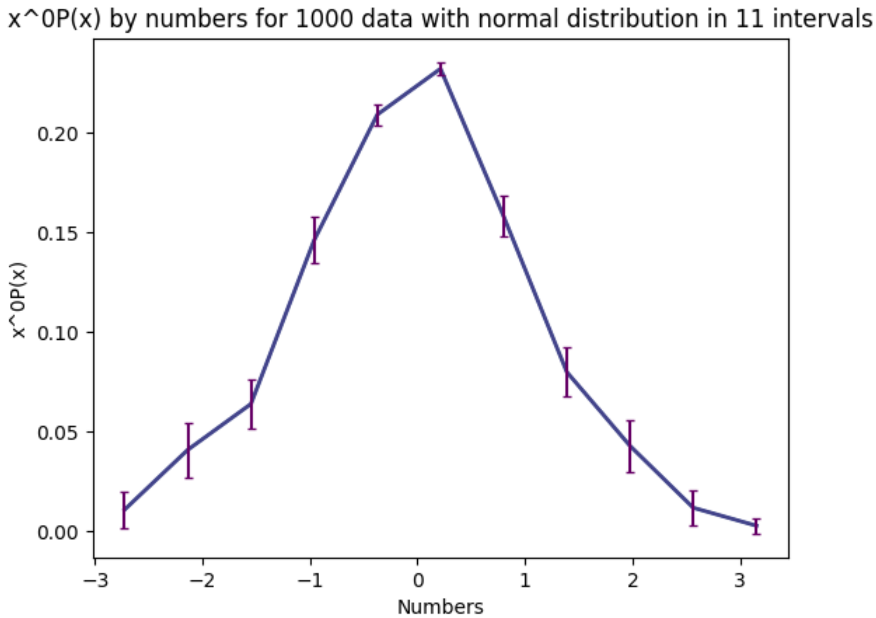

Figure 12: $X^0P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

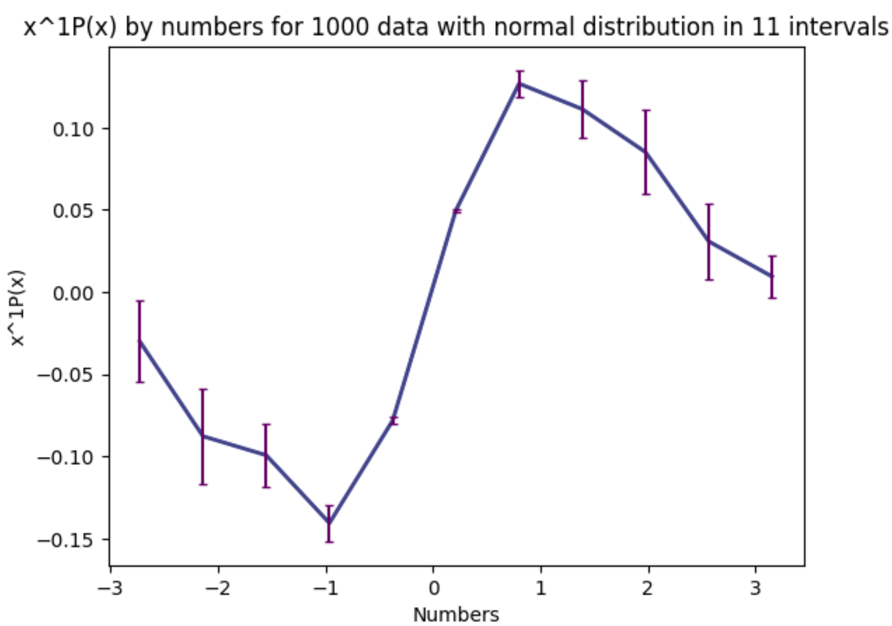

Figure 13: $X^1P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

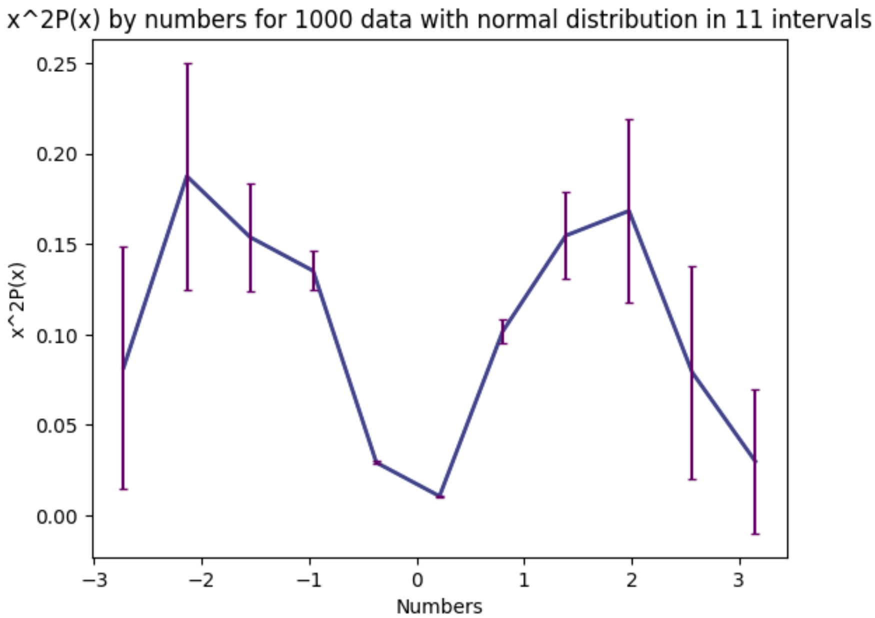

Figure 14: $X^2P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

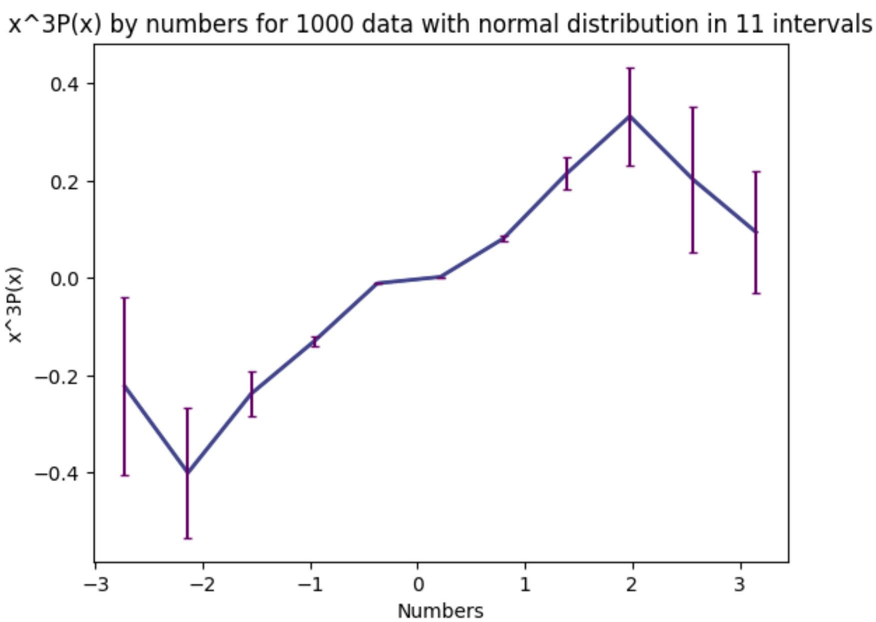

Figure 15: $X^3P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

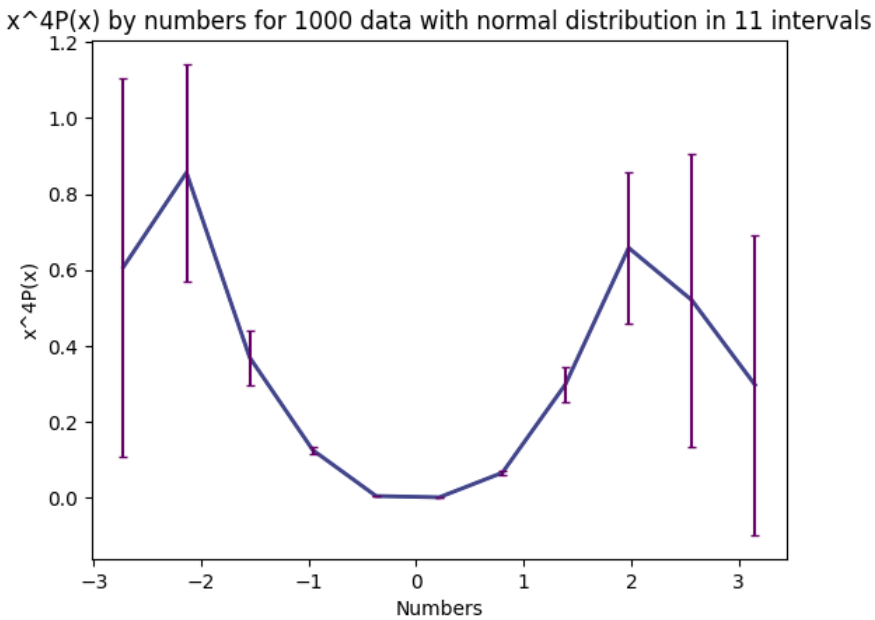

Figure 16: $X^4P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

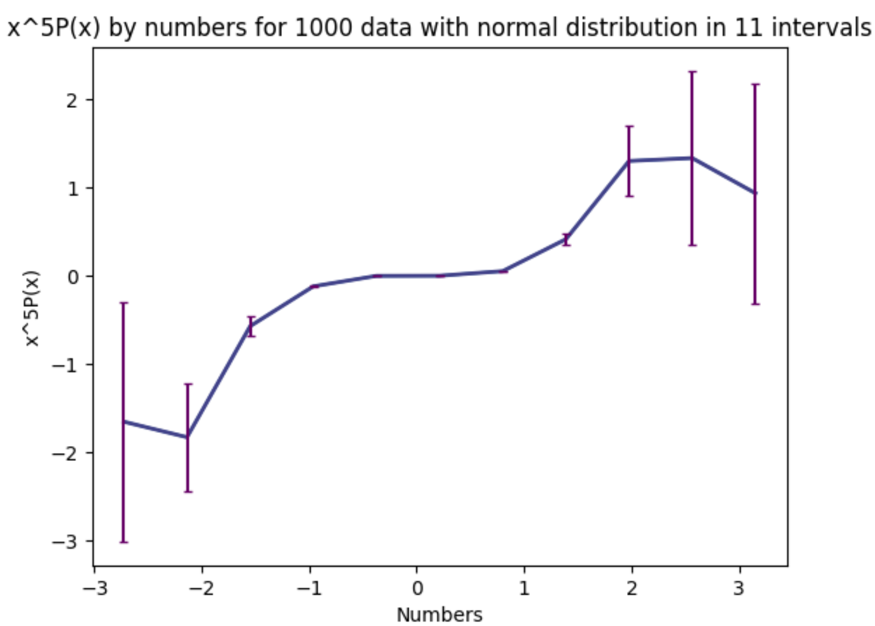

Figure 17: $X^5P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

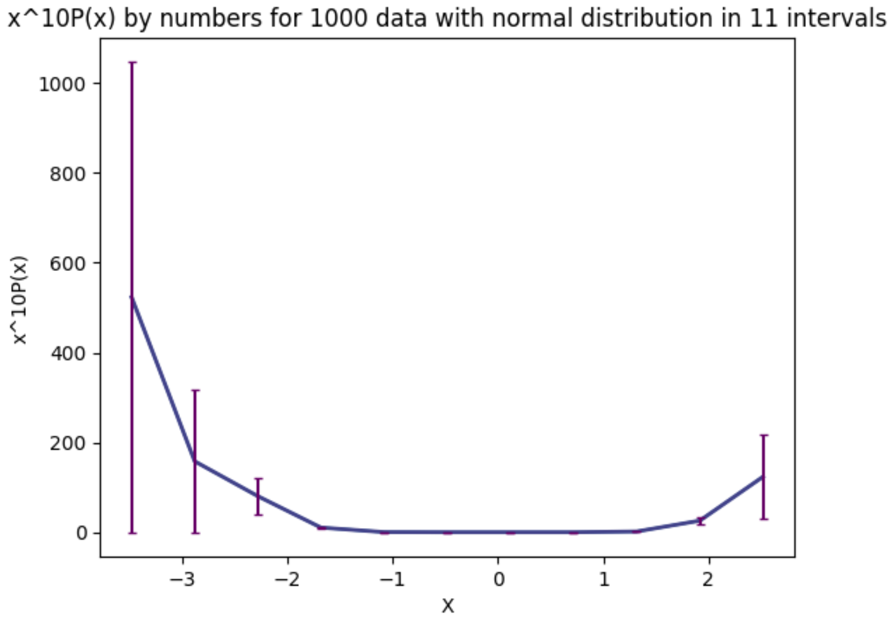

Figure 18: $X^(10)P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

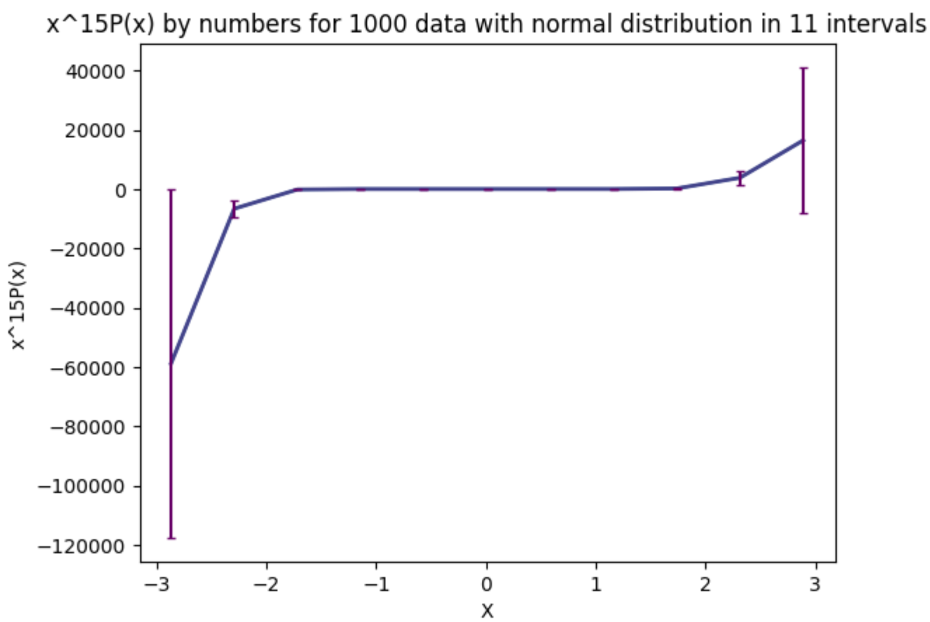

Figure 19: $X^(15)P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

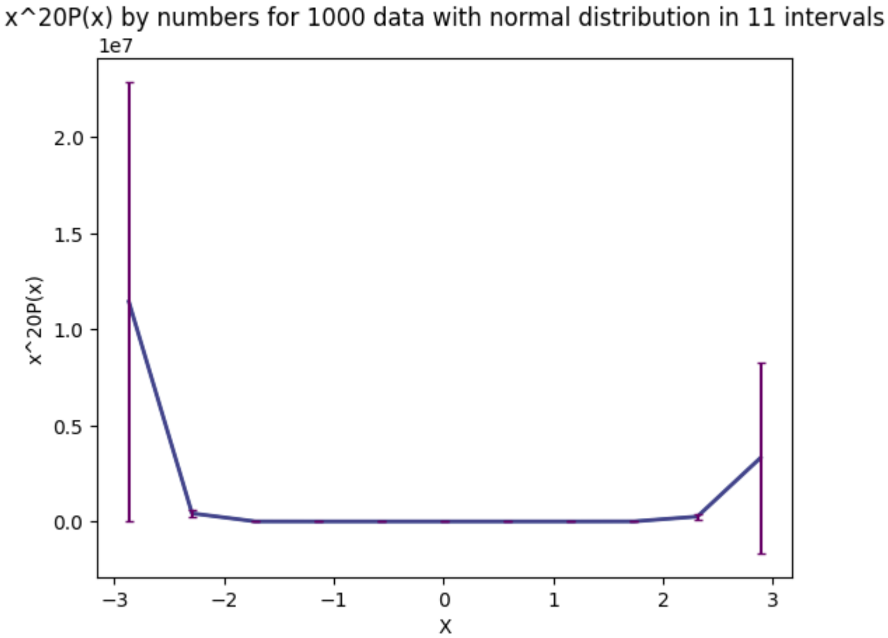

Figure 20: $X^(20)P(X)$ according to X for 10^3 data with normal distribution in 11 intervals

As we can see in the graphs, for 10^3 data, we can hardly calculate $X^3P(X)$, and from then on, all graphs show exponential behavior in the corners and the errors become large.

2.1.3 Normal distribution n=10^4:

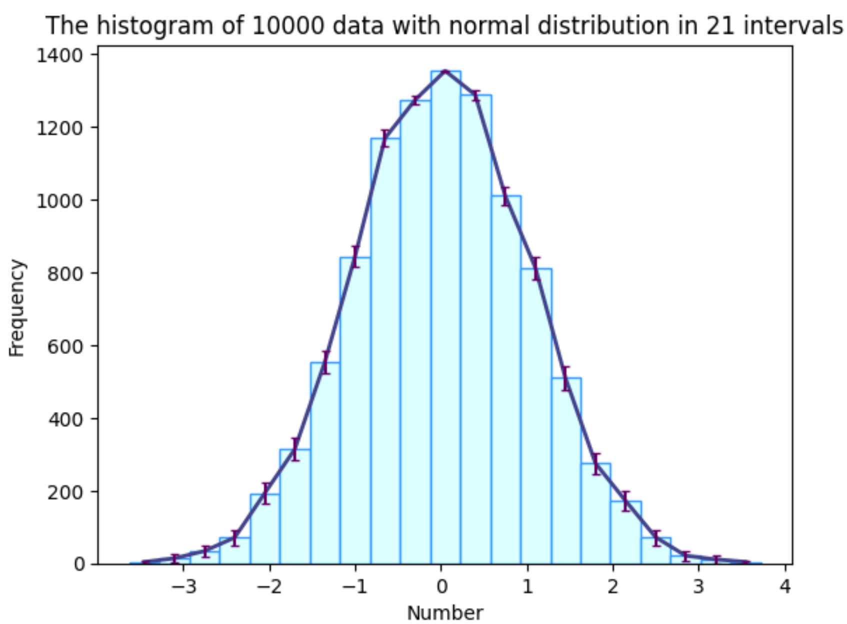

Figure 21: The histogram of 10^4 data with normal distribution in 21 intervals

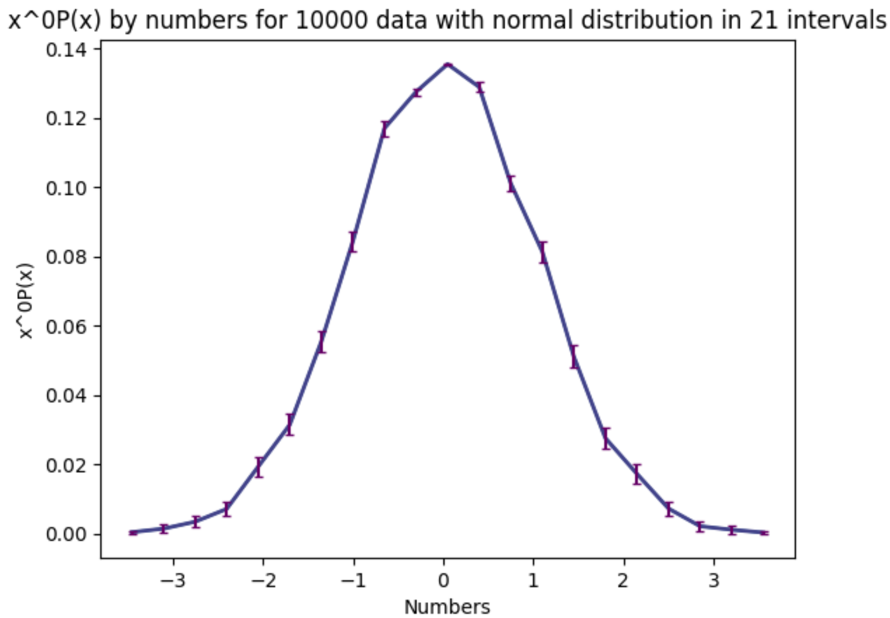

Figure 22: $X^0P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

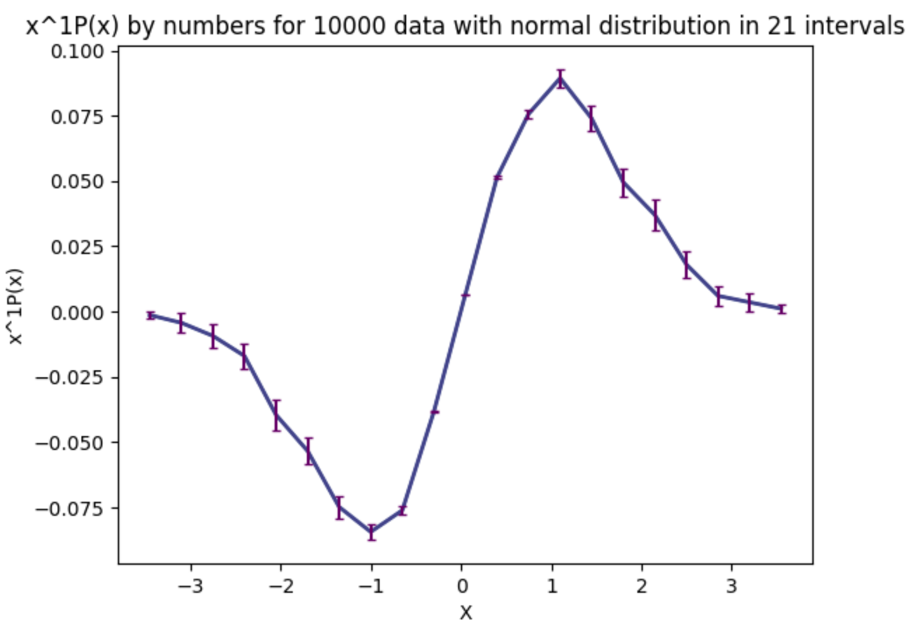

Figure 23: $X^1P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

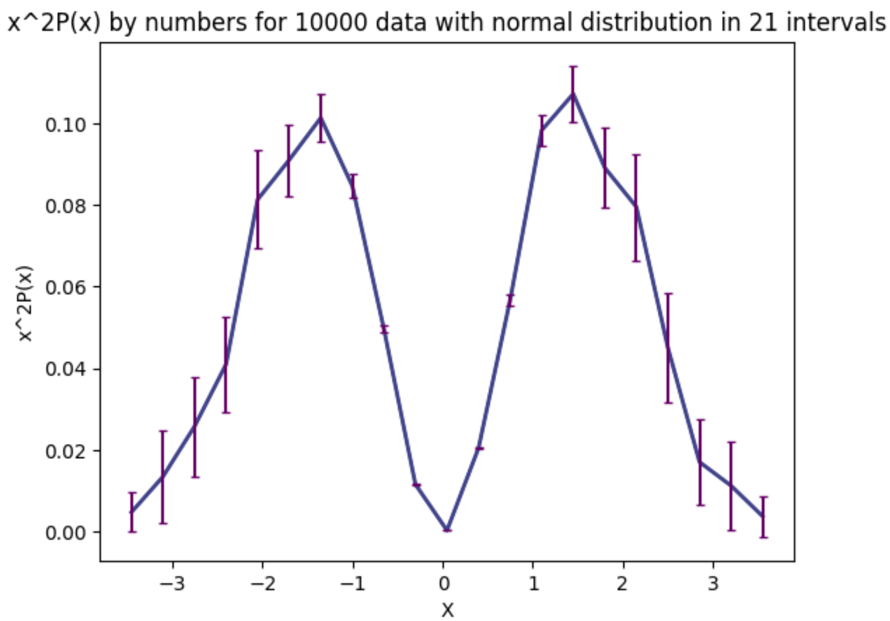

Figure 24: $X^2P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

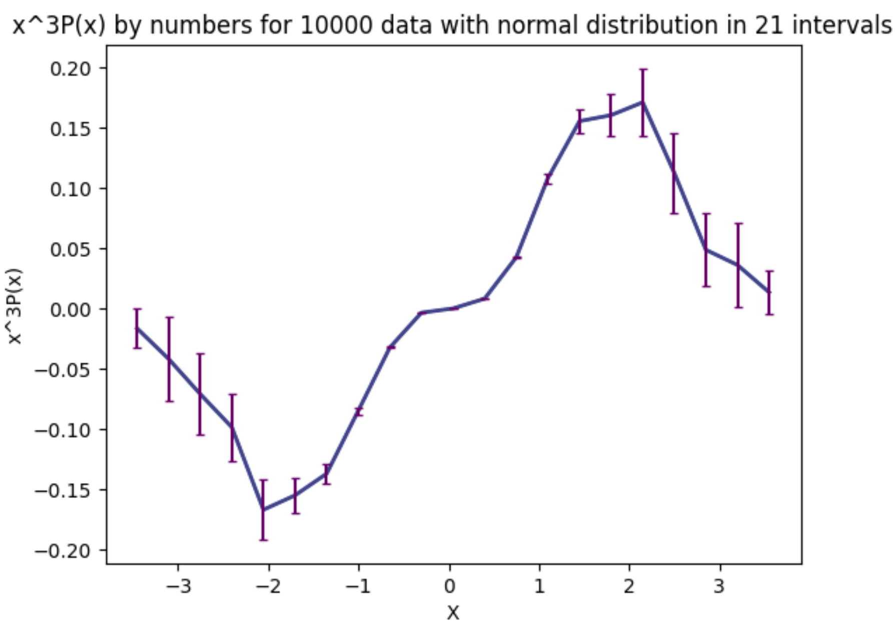

Figure 25: $X^3P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

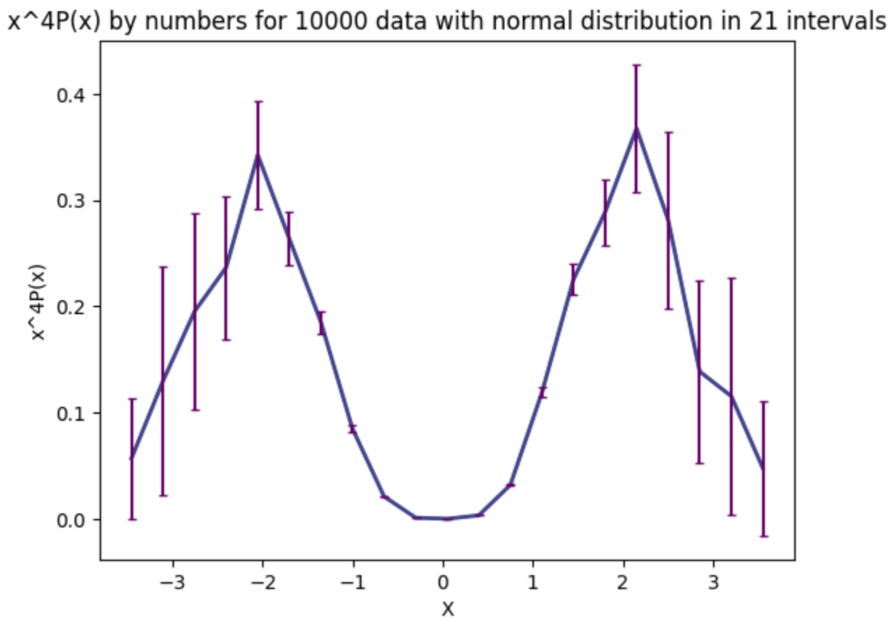

Figure 26: $X^4P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

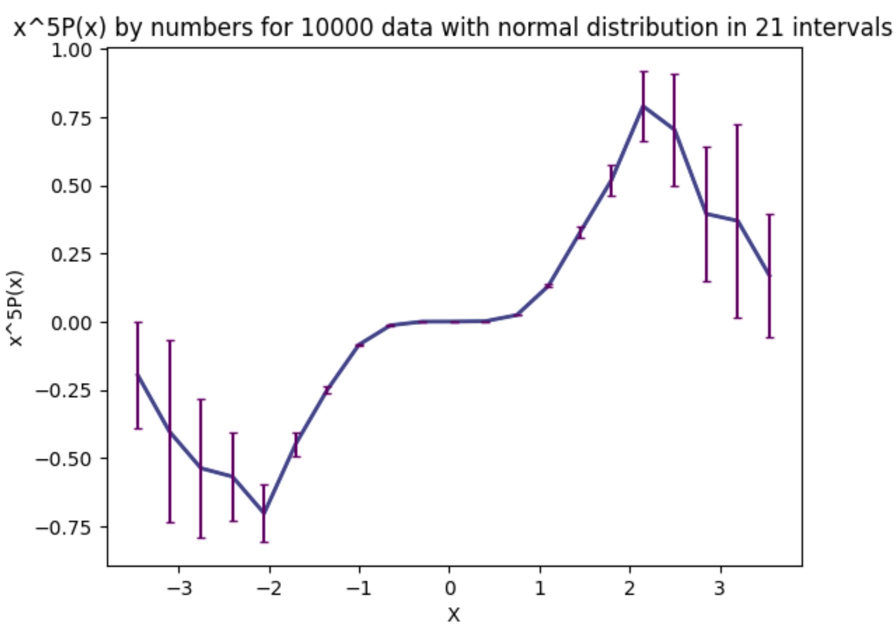

Figure 27: $X^5P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

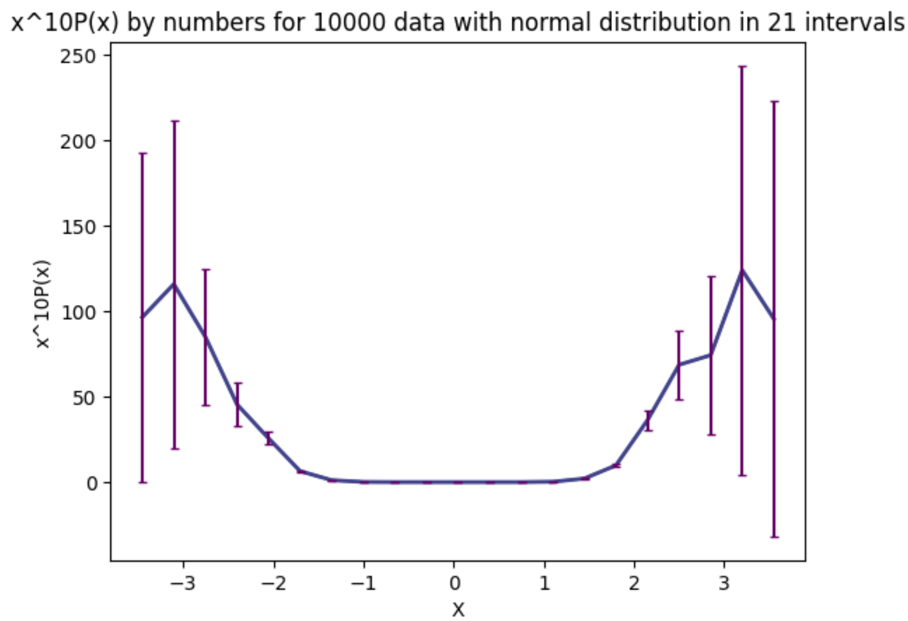

Figure 28: $X^(10)P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

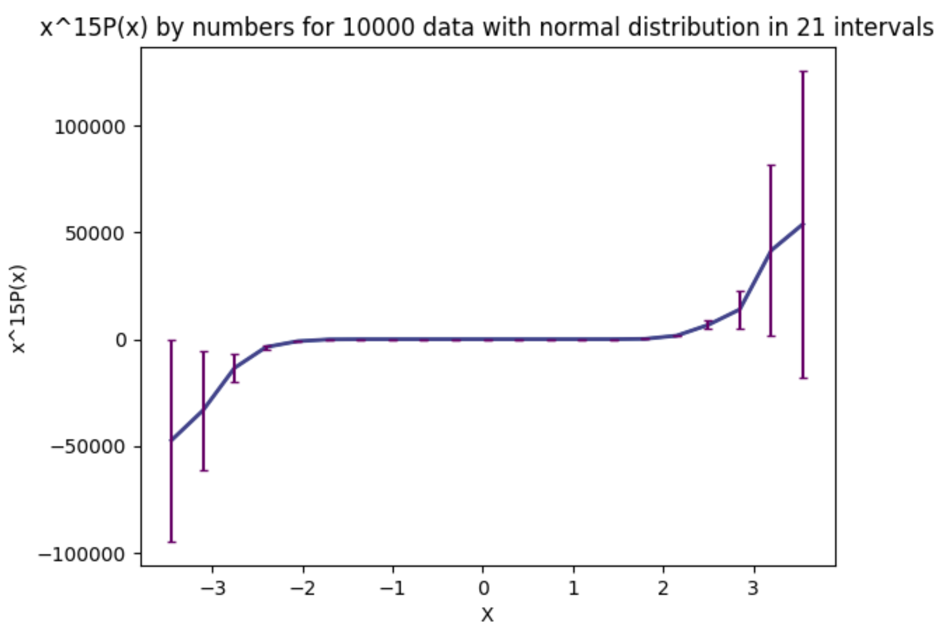

Figure 29: $X^(15)P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

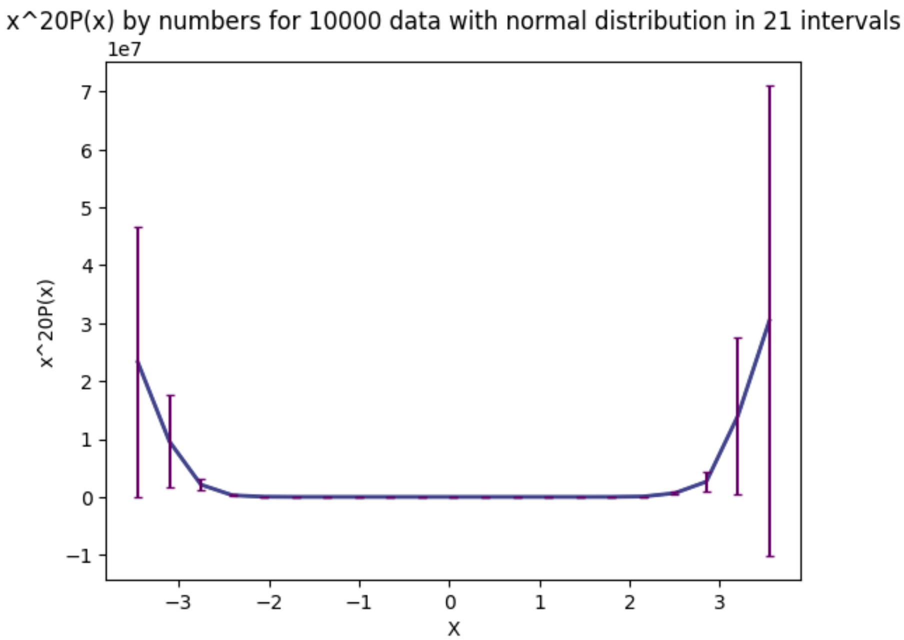

Figure 30: $X^(20)P(X)$ according to X for 10^4 data with normal distribution in 21 intervals

As we can see in the graphs, for 10^4 data, we can hardly calculate $X^4P(X)$ or $X^5P(X)$, and from then on, all graphs show exponential behavior in the corners and the errors become large.

2.1.4 Normal distribution n=10^5:

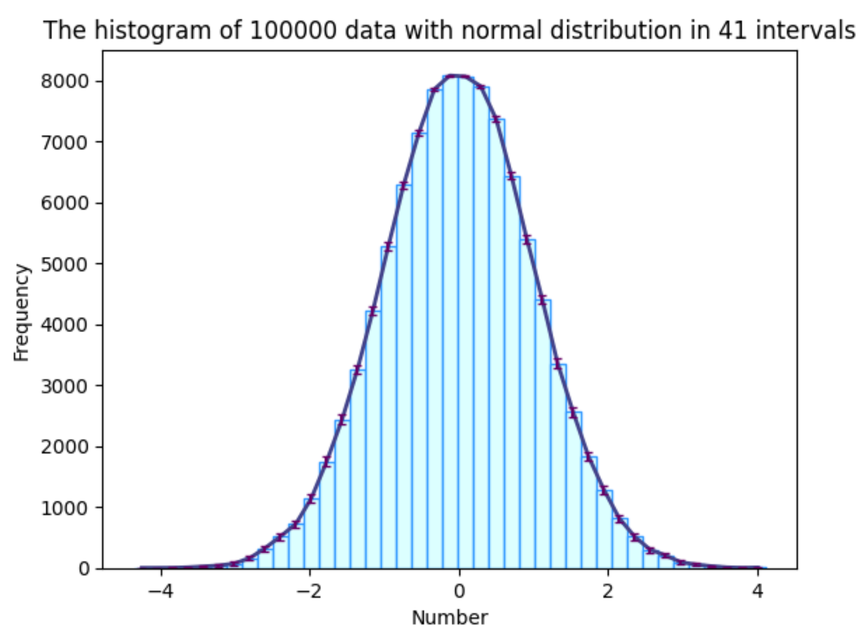

Figure 31: The histogram of 10^5 data with normal distribution in 41 intervals

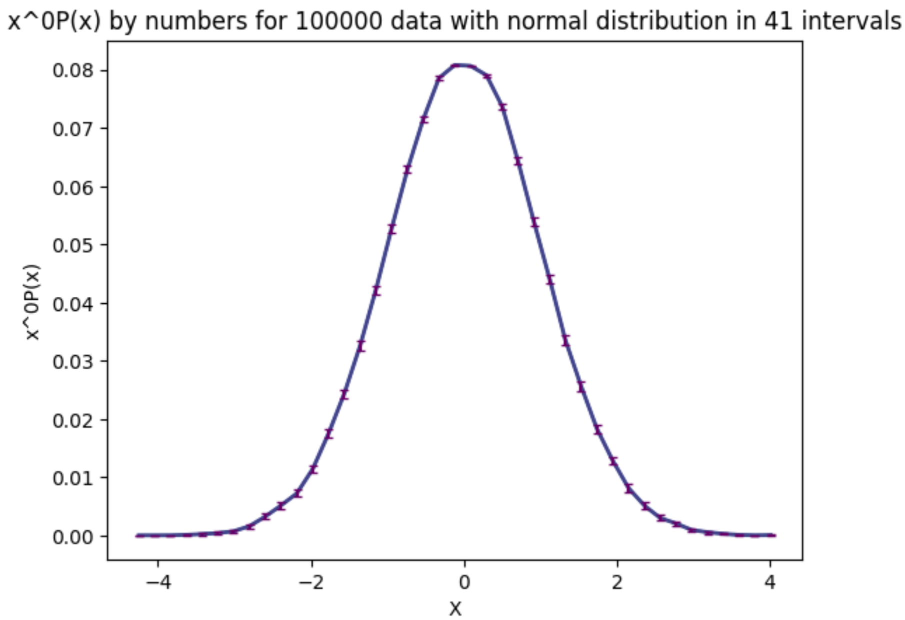

Figure 32: $X^0P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

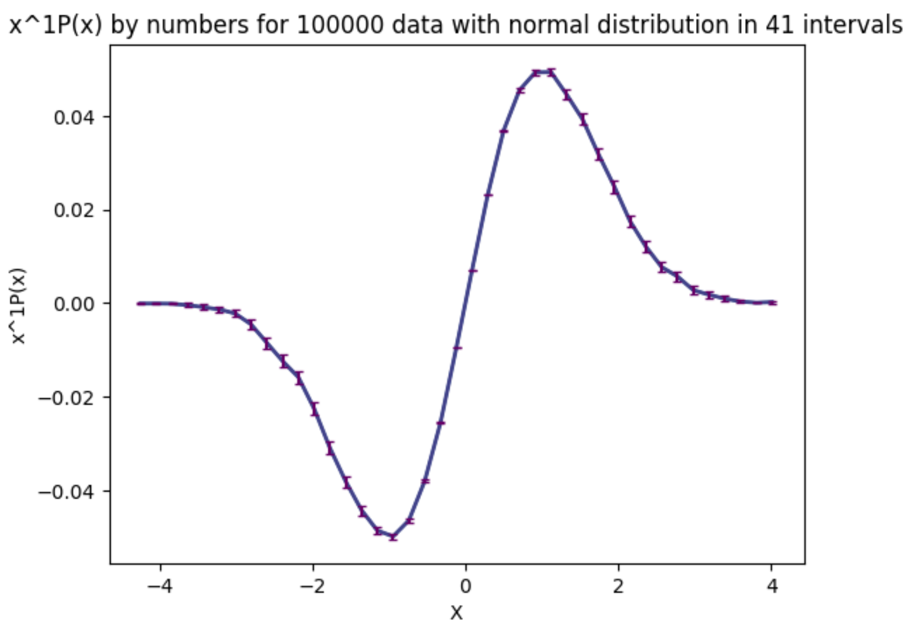

Figure 33: $X^1P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

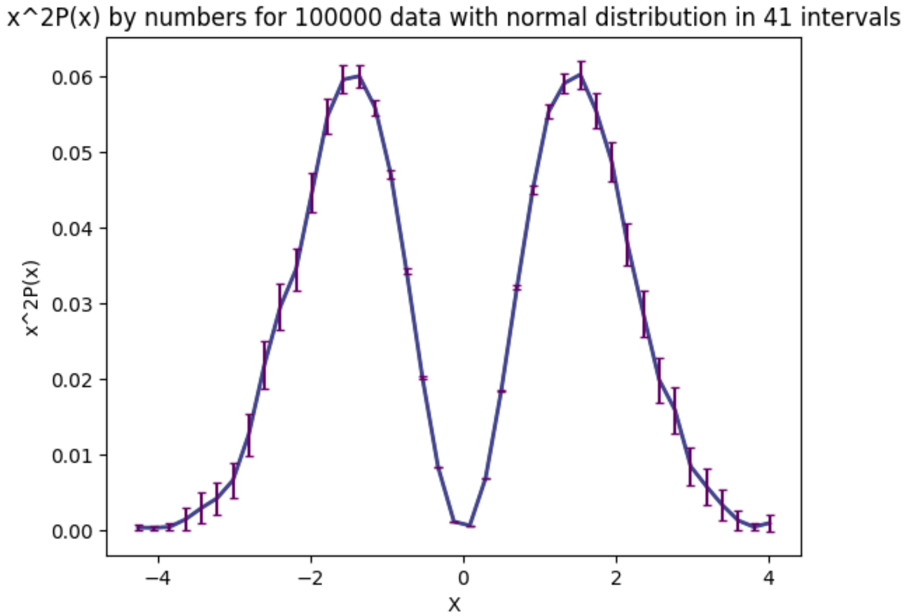

Figure 34: $X^2P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

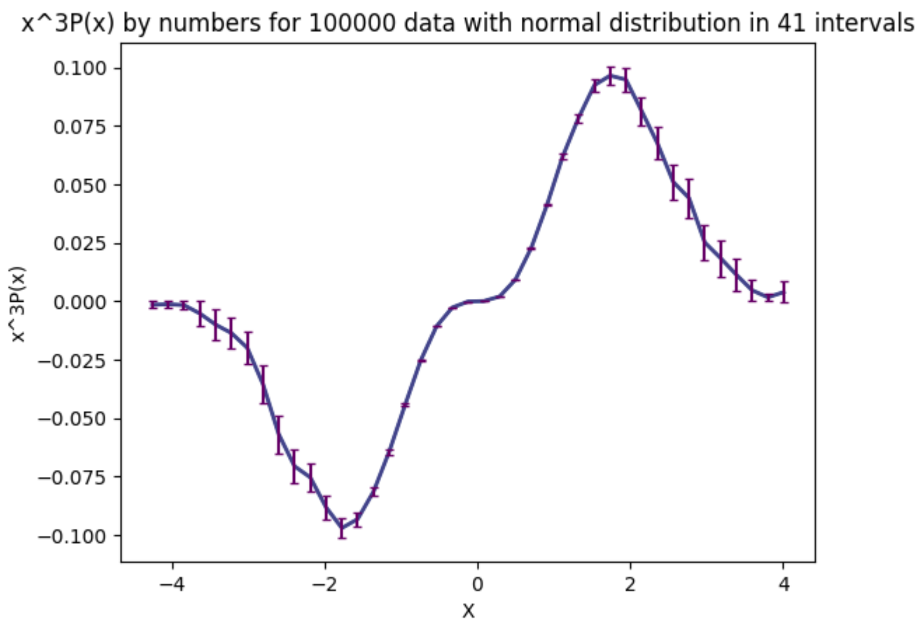

Figure 35: $X^3P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

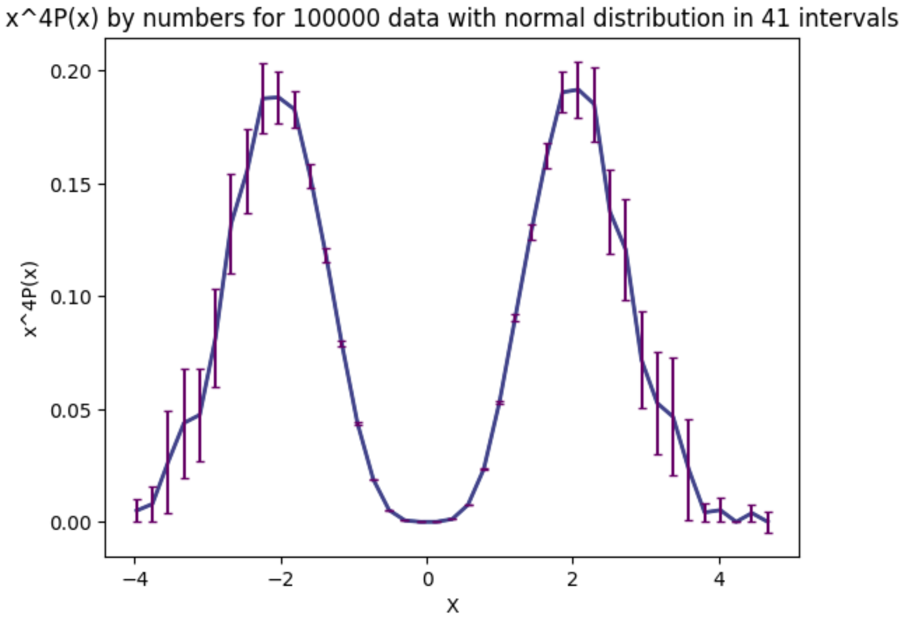

Figure 36: $X^4P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

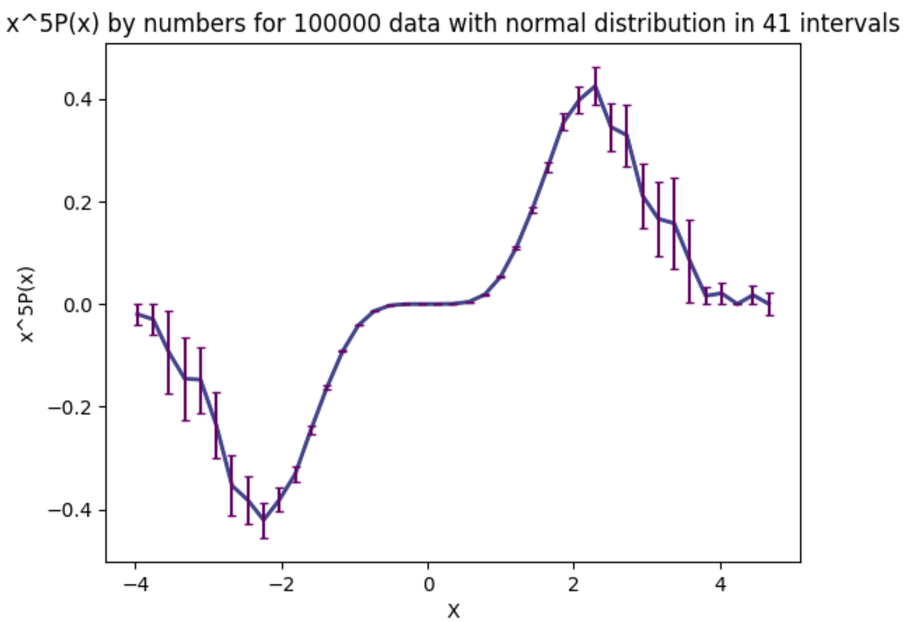

Figure 37: $X^5P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

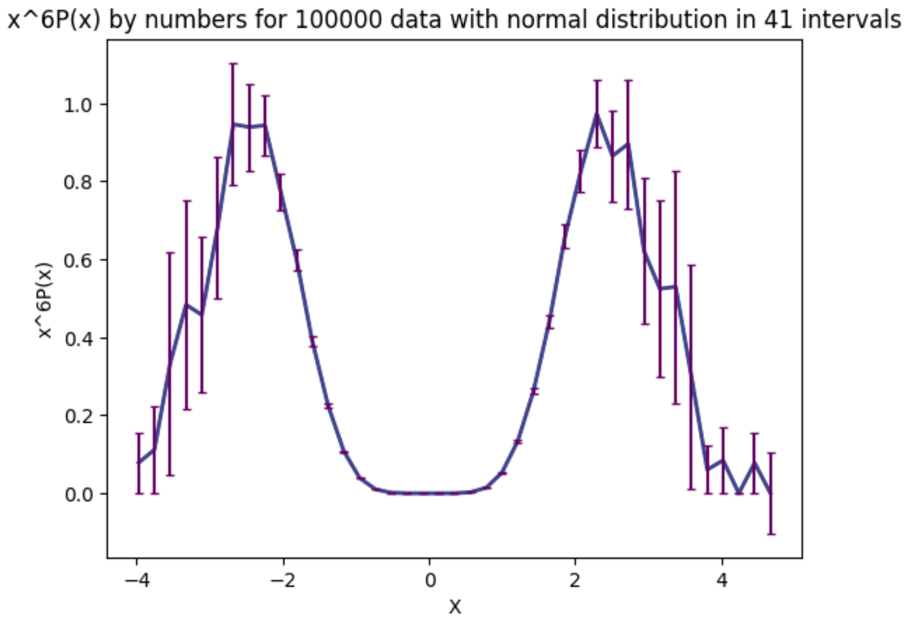

Figure 38: $X^6P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

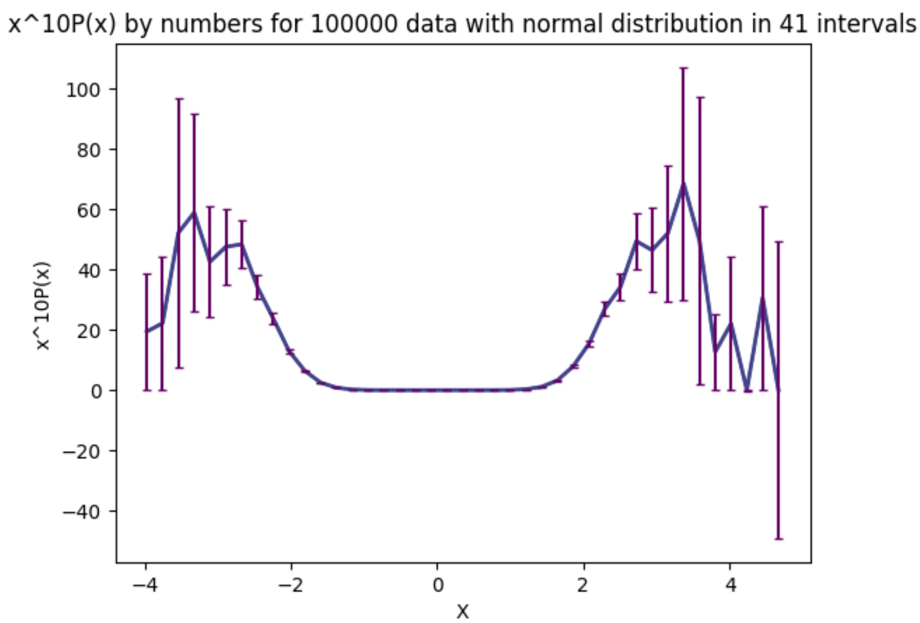

Figure 39: $X^(10)P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

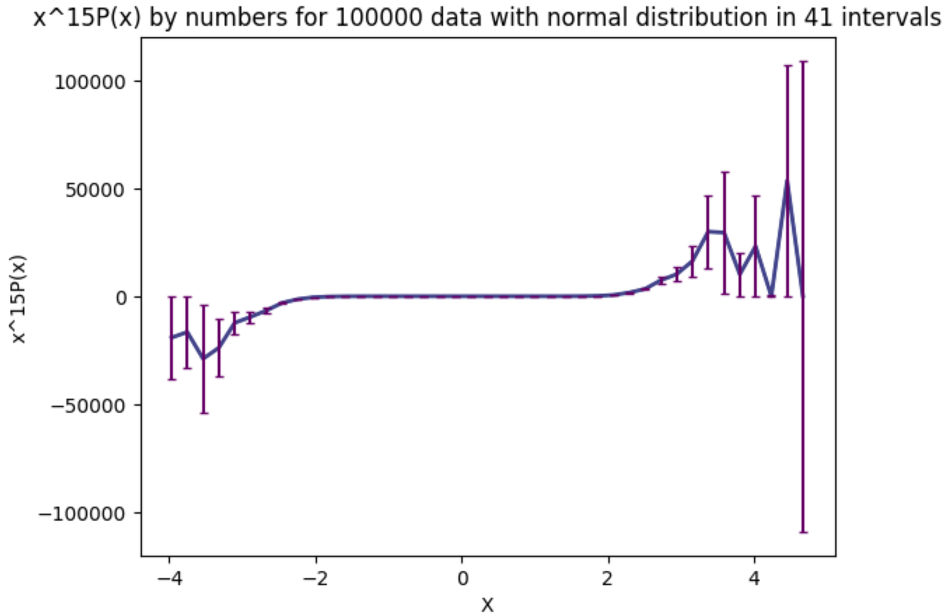

Figure 40: $X^(15)P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

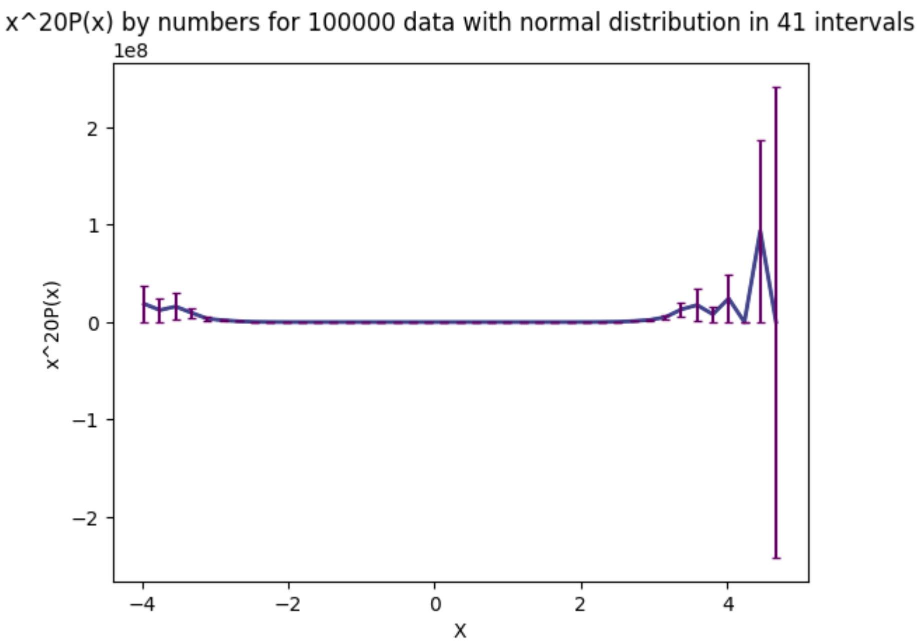

Figure 41: $X^(20)P(X)$ according to X for 10^5 data with normal distribution in 41 intervals

As we can see in the graphs, for 10^5 data, we can hardly calculate $X^5P(X)$, and from then on, all graphs show exponential behavior in the corners and the errors become large.

2.1.5 Normal distribution n=10^6:

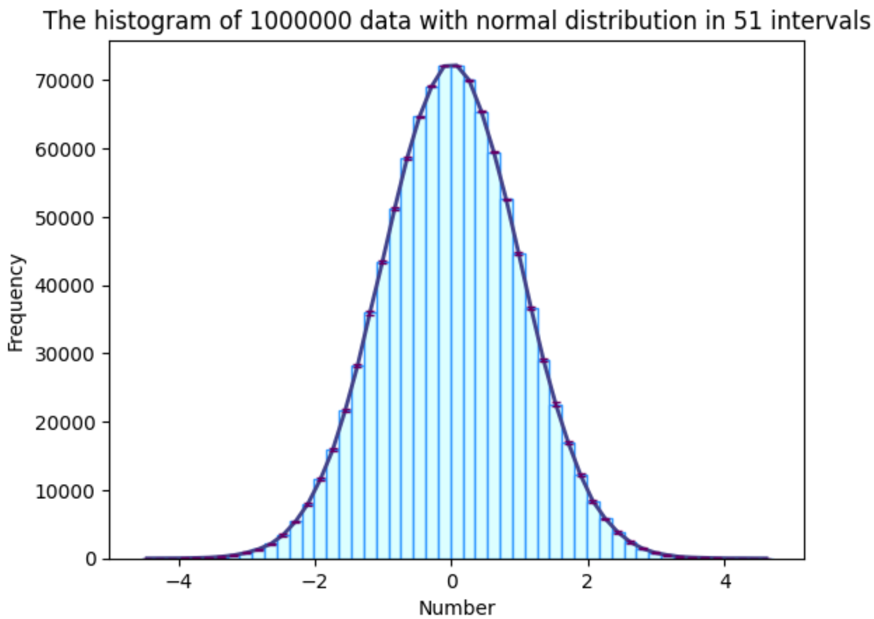

Figure 42: The histogram of 10^6 data with normal distribution in 51 intervals

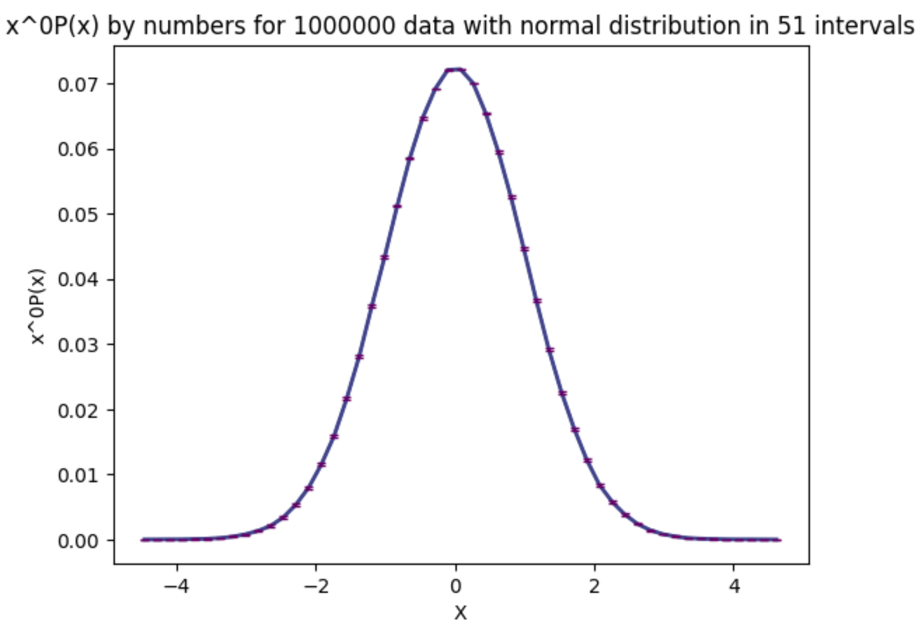

Figure 43: $X^0P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

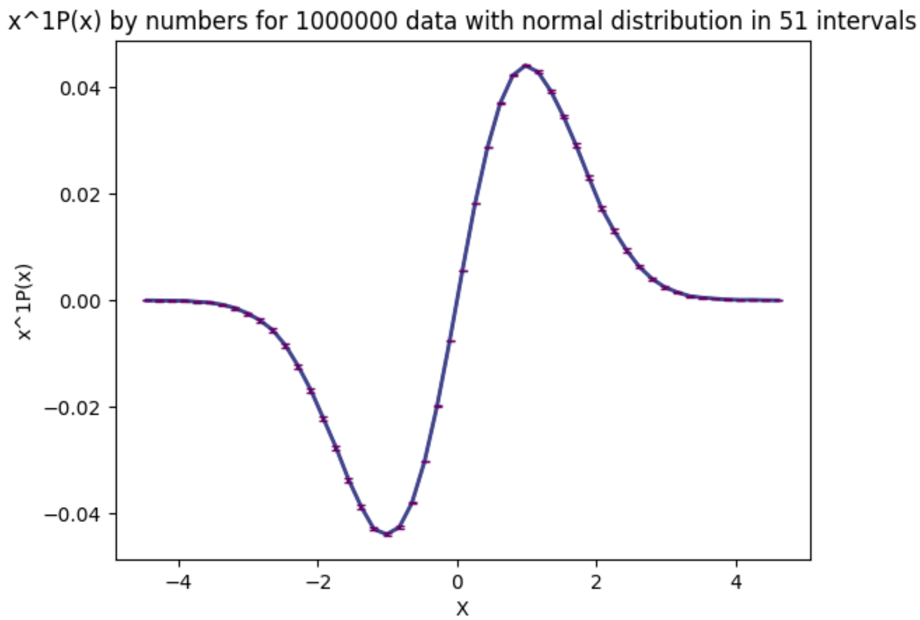

Figure 44: $X^1P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

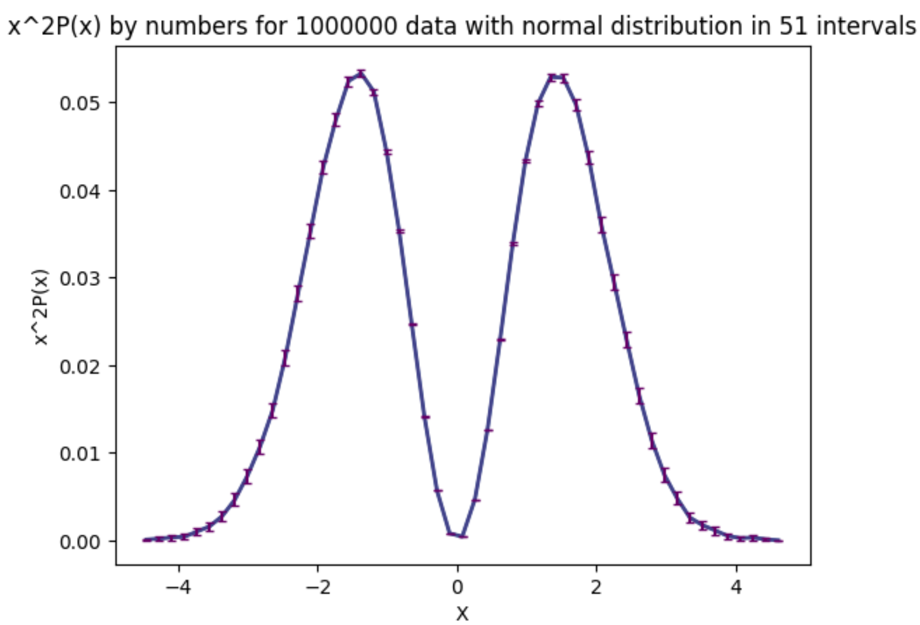

Figure 45: $X^2P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

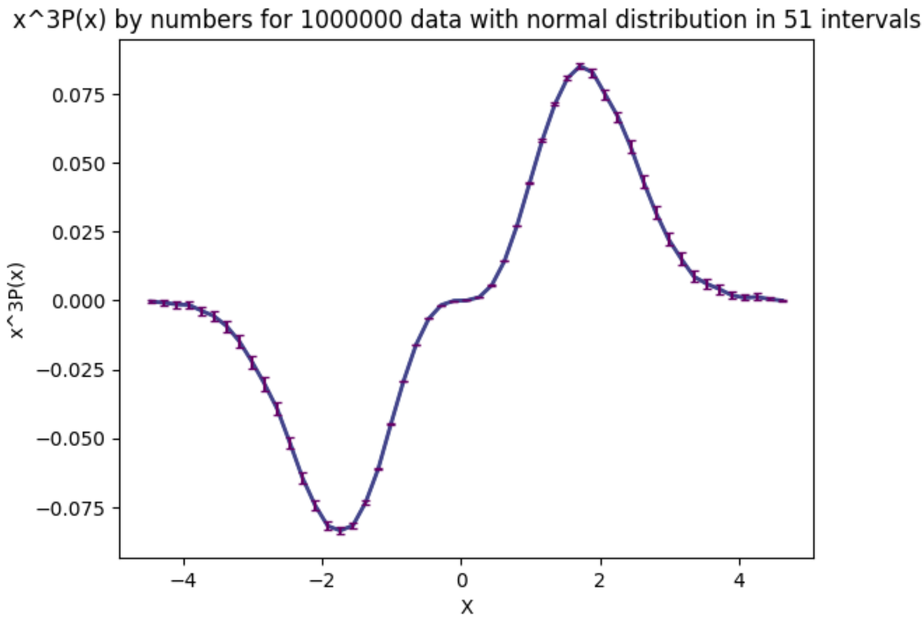

Figure 46: $X^3P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

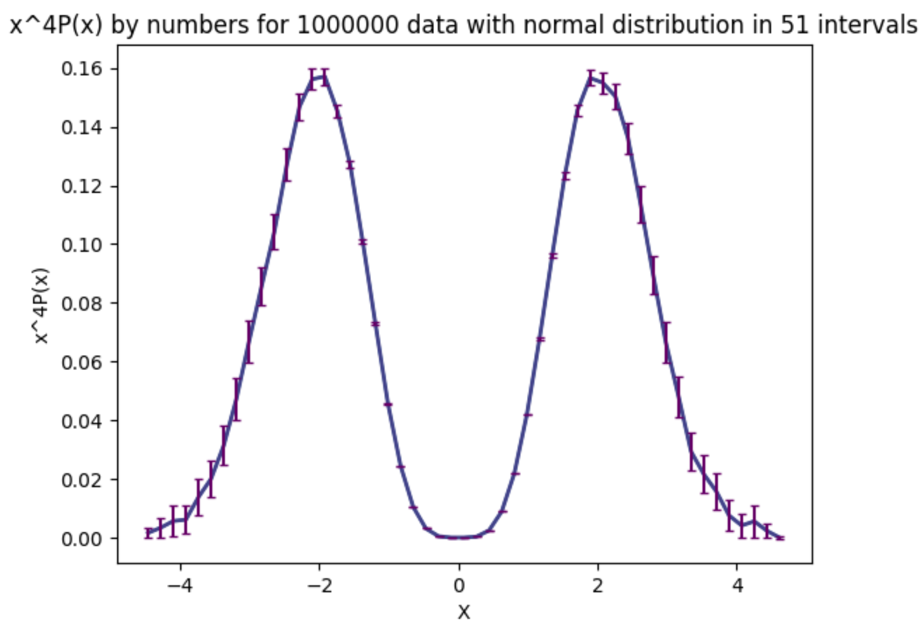

Figure 47: $X^4P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

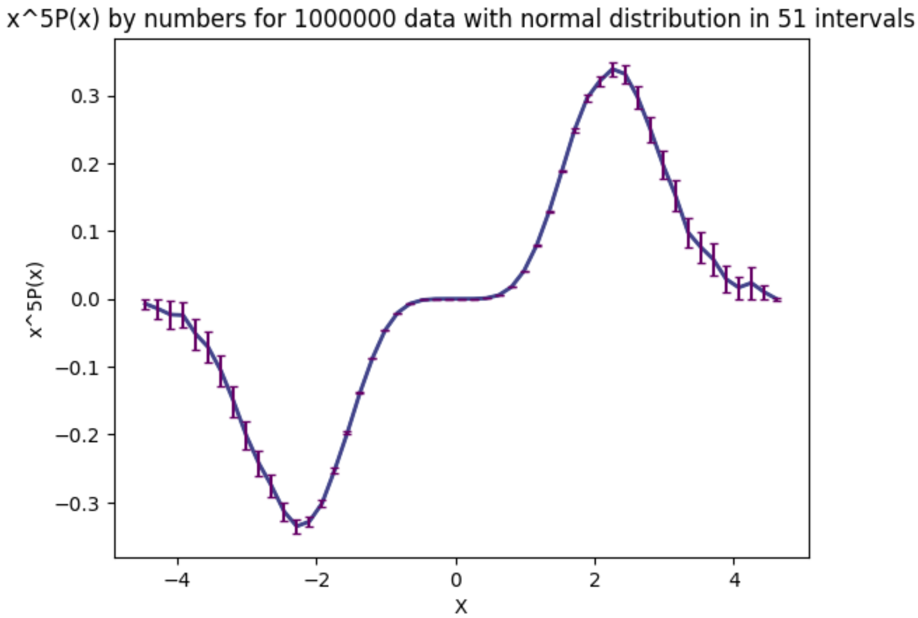

Figure 48: $X^5P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

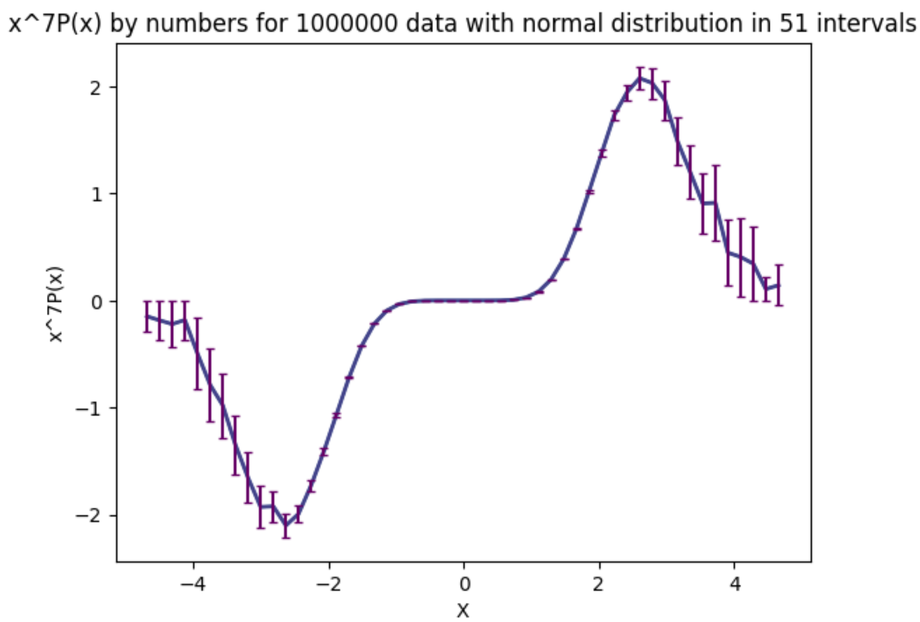

Figure 49: $X^7P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

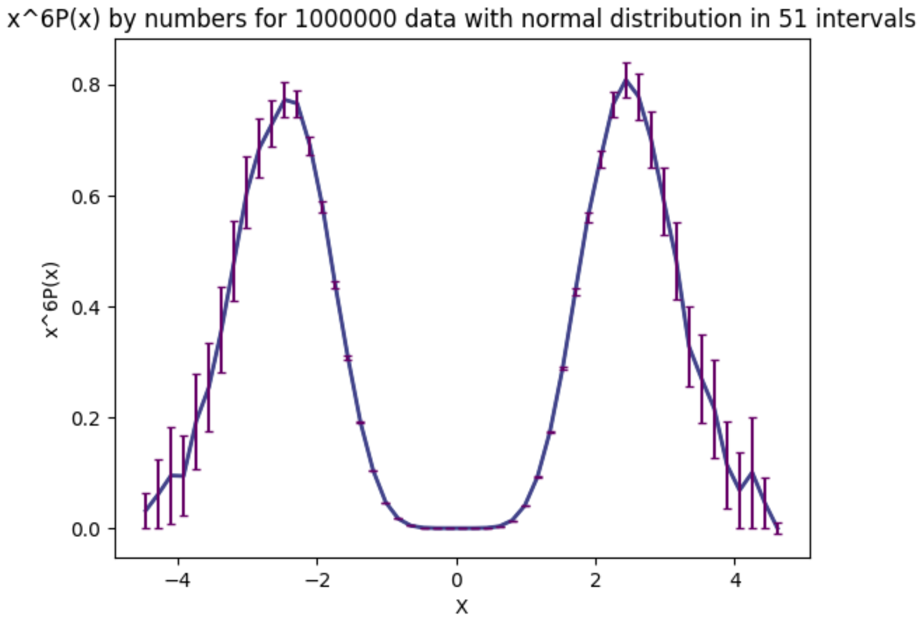

Figure 50: $X^6P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

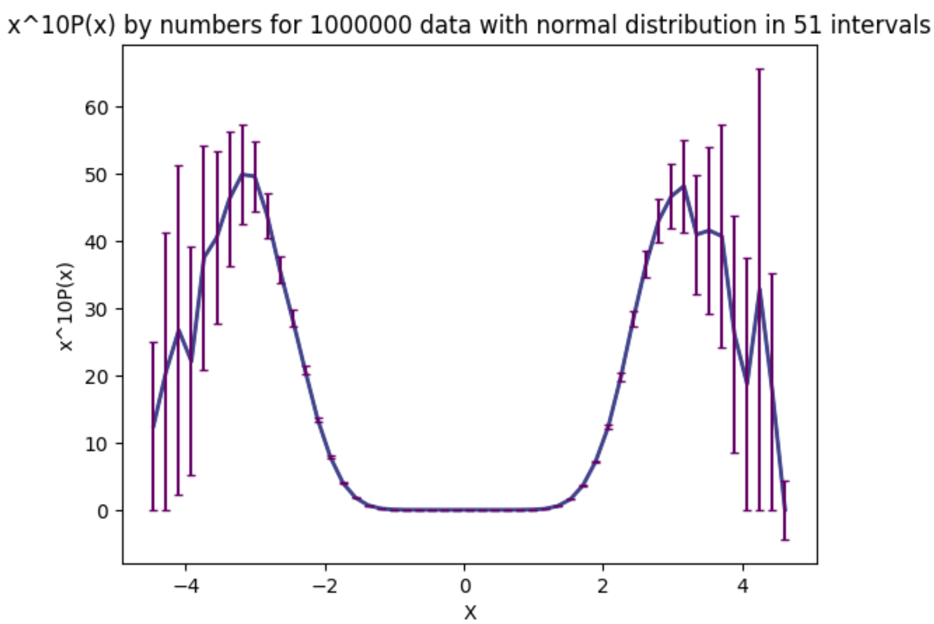

Figure 51: $X^(10)P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

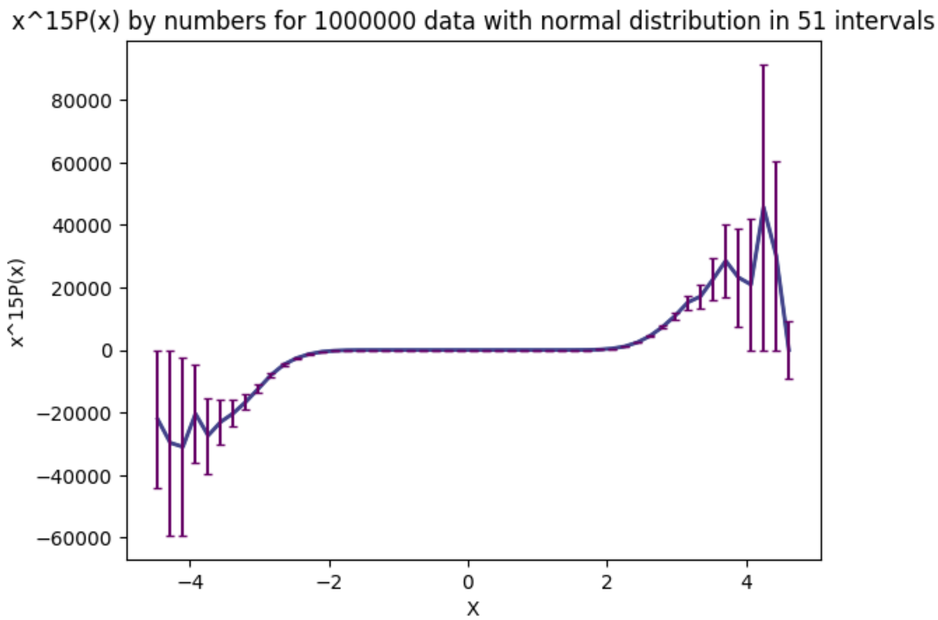

Figure 52: $X^(15)P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

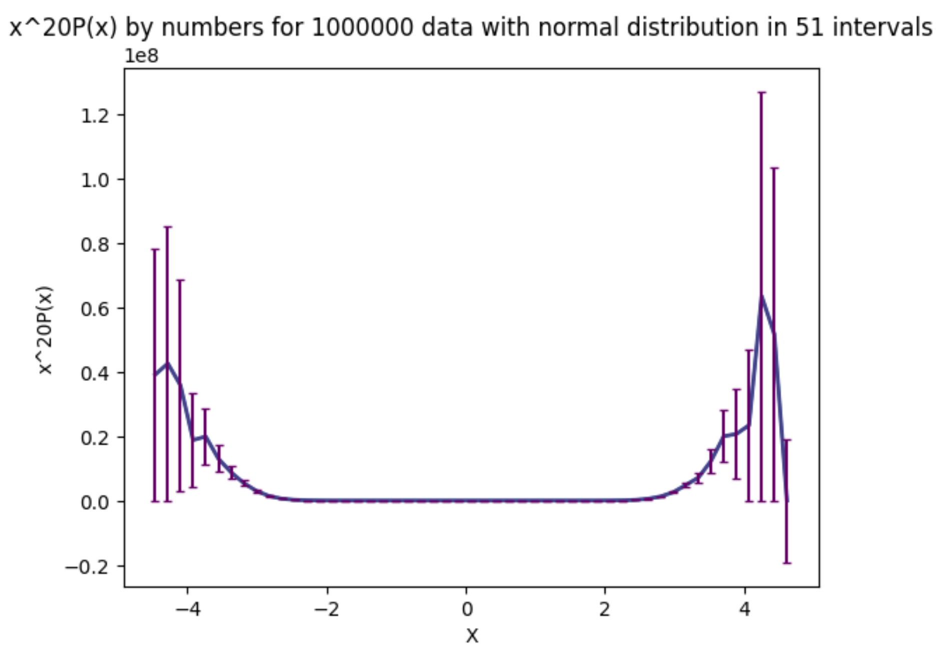

Figure 53: $X^(20)P(X)$ according to X for 10^6 data with normal distribution in 51 intervals

As we can see in the graphs, for 10^6 data, we can hardly calculate $X^6P(X)$, and from then on, all graphs show exponential behavior in the corners and the errors become large.

2.2 Question 2

In this part of the exercise, we want to draw different moments according to the number of data and check their behavior with the increase of the number of data.

2.2.1 Moment 0:

As we expect, the $0th$ moment for all numbers of data is equal to zero and without error because every number to the power of 0 is equal to 1.

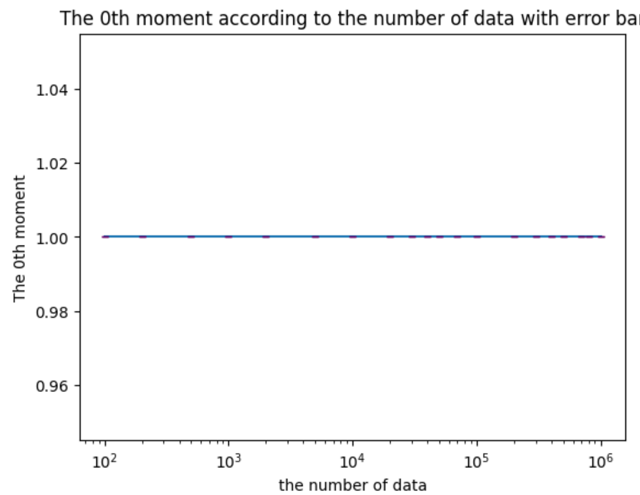

Figure 54: The $0th$ moment according to the number of data with error bar

2.2.2 Moment 1:

For the $1th$ moment, because we manually set the average of all data series to 0, the graph moves to zero, but the more the number of data, the more accurate this average is, and that is why the error rate is smaller.all of the

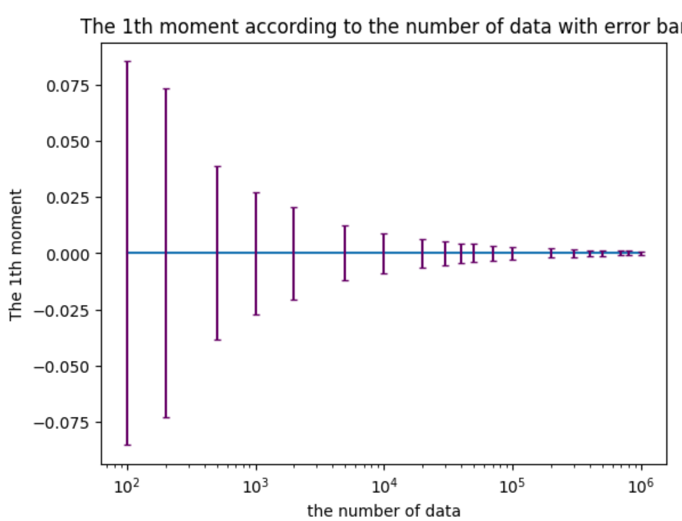

Figure 55: The $1th$ moment according to the number of data with error bar

2.2.3 Moment 2:

As we have obtained from the previous section, the $2th$ moment is calculated as the limit moment for the number of 1062 data, and the error bar increases greatly, and this continues in the next moments as well. For any amount of data, we can see how the error bar increases at that specific moment.

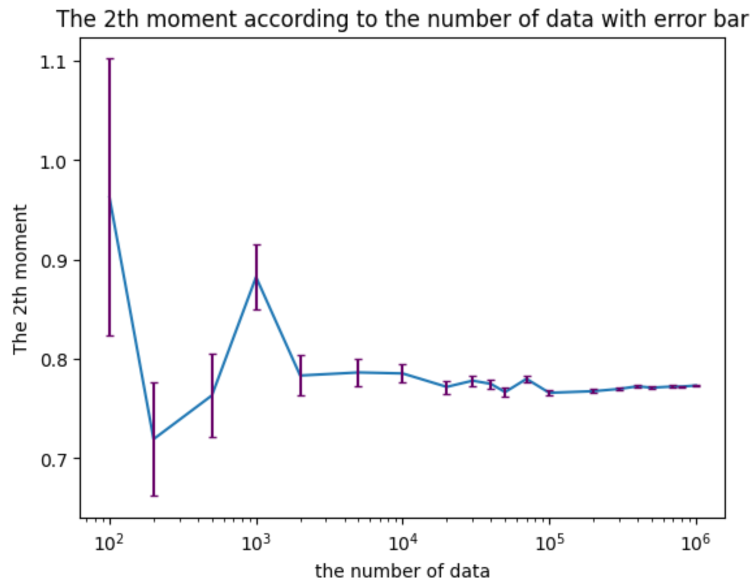

Figure 56: The $2th$ moment according to the number of data with error bar

2.2.4 Moment 3:

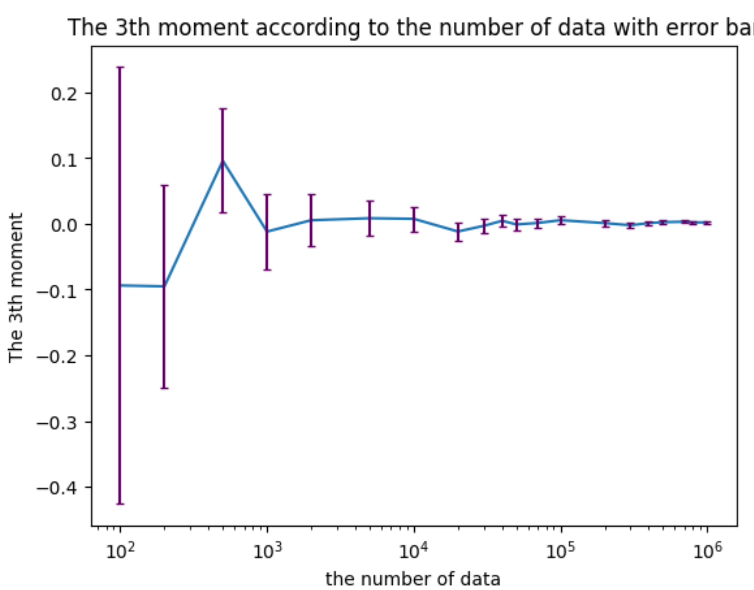

Figure 57: The $3th$ moment according to the number of data with error bar

2.2.5 Moment 4:

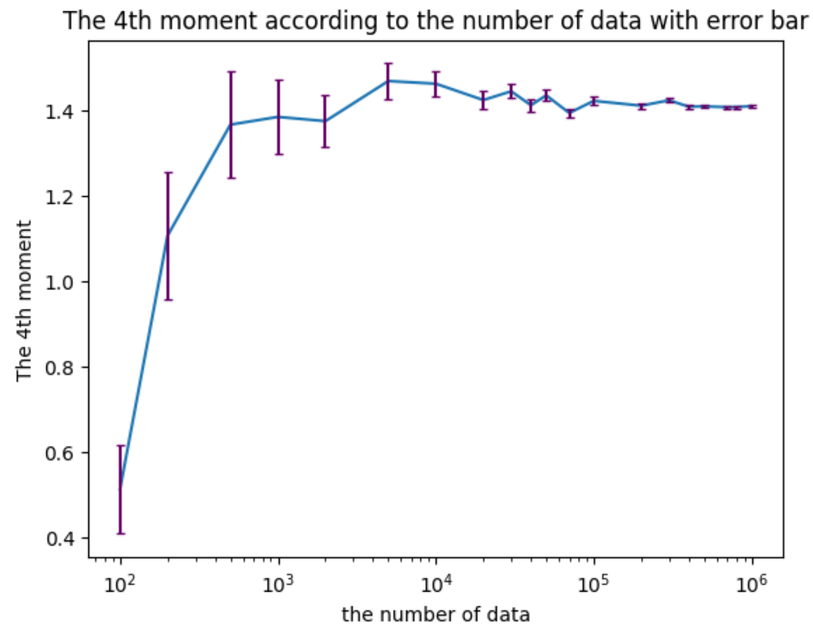

Figure 58: The $4th$ moment according to the number of data with error bar

2.2.6 Moment 5:

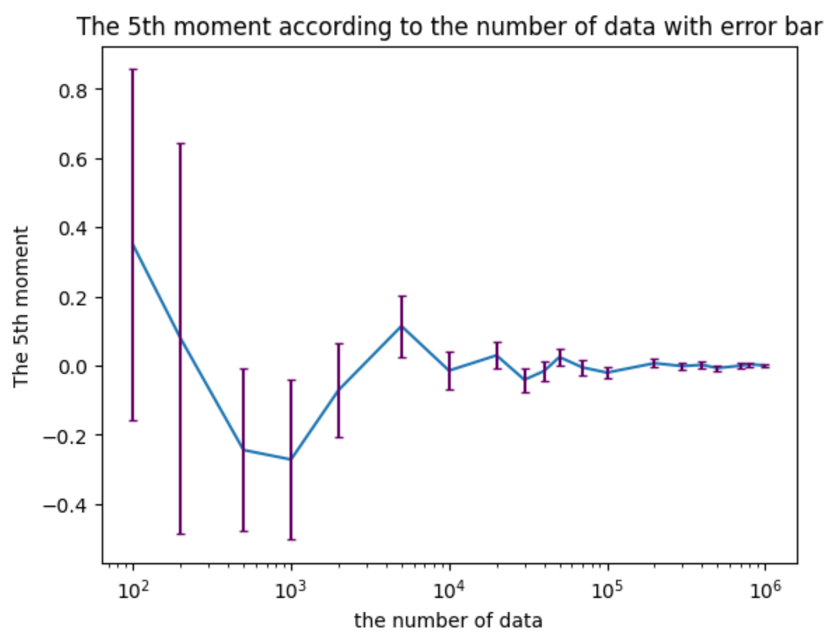

Figure 59: The $5th$ moment according to the number of data with error bar

2.2.7 Moment 6:

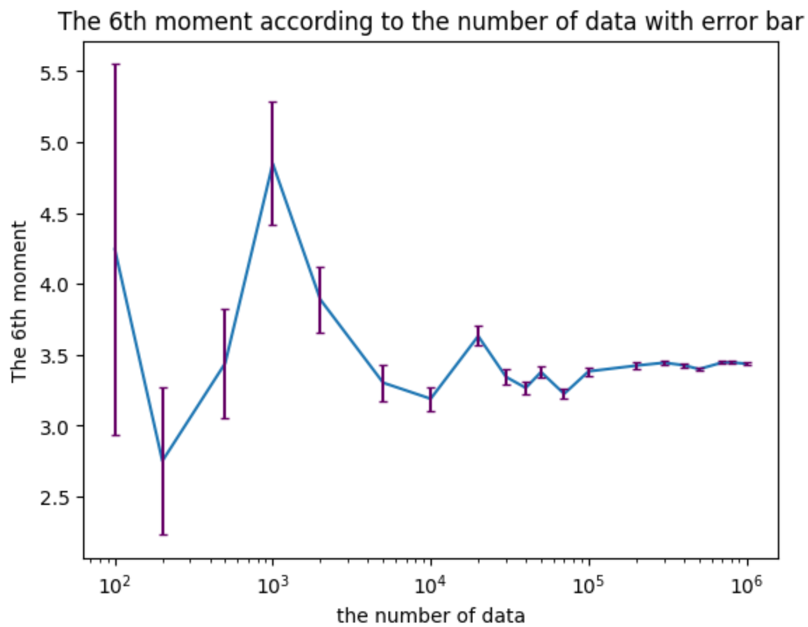

Figure 60: The $6th$ moment according to the number of data with error bar

2.2.8 Moment 7:

As we expect, from this moment on, the upward graph is no longer certain, even considering the error bars, this is because we set the maximum number of data to 10^6.

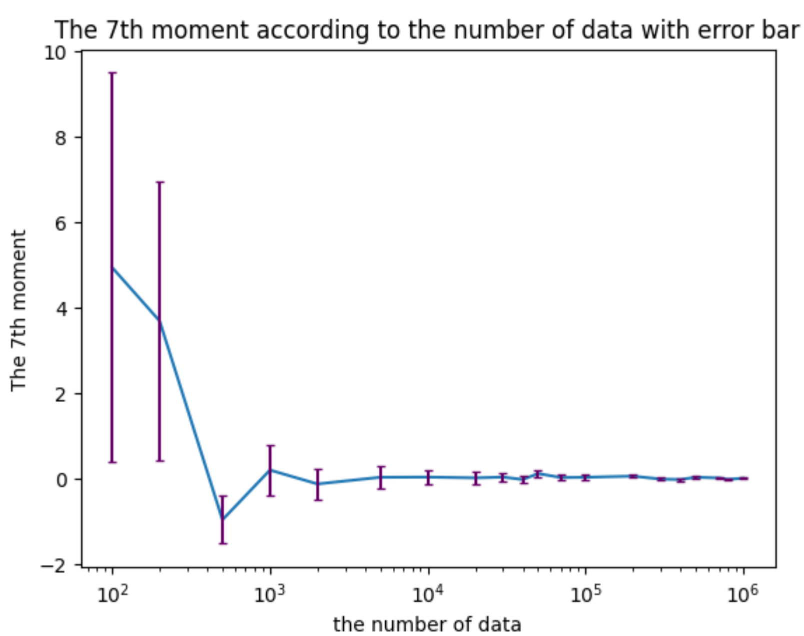

Figure 61: The $7th$ moment according to the number of data with error bar

2.2.9 Moment 10:

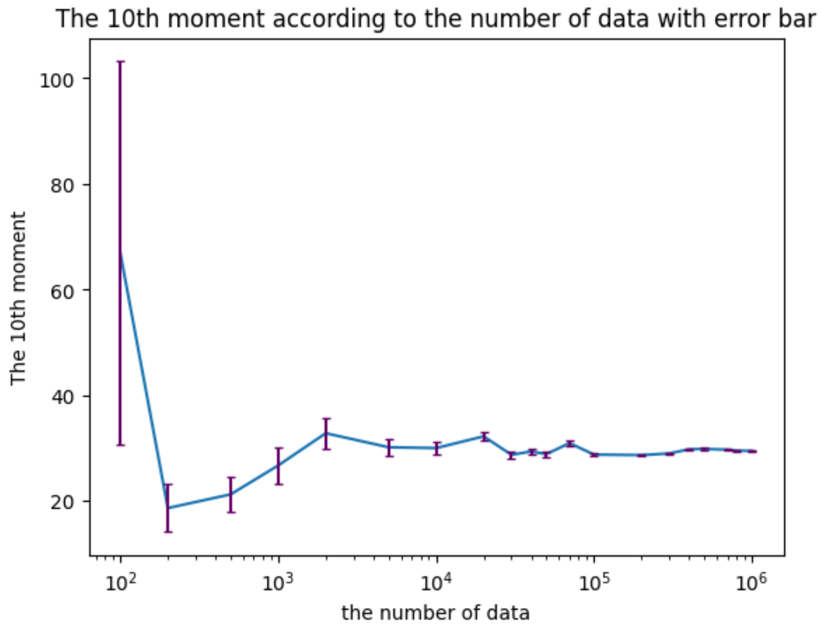

Figure 62: The $10th$ moment according to the number of data with error bar

2.2.10 Moment 12:

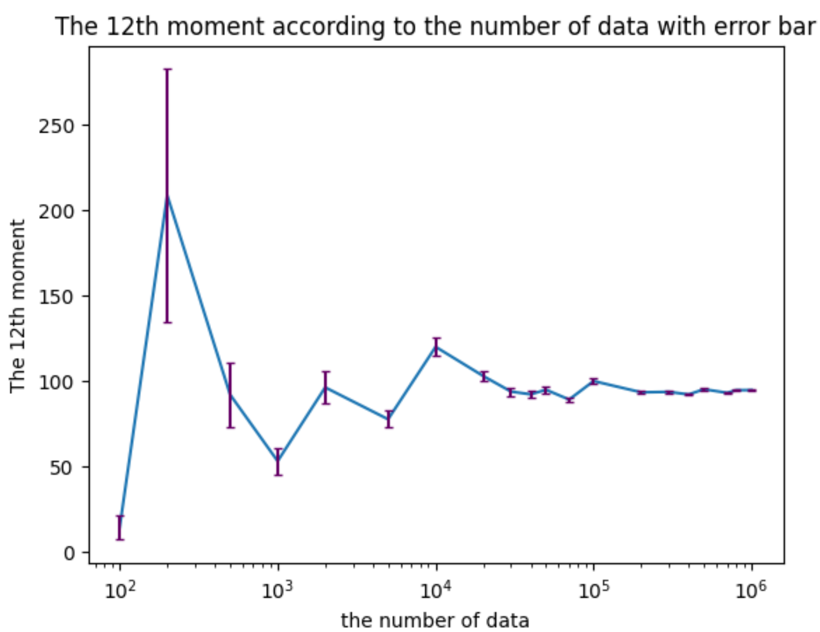

Figure 63: The $12th$ moment according to the number of data with error bar

2.2.11 Moment 15:

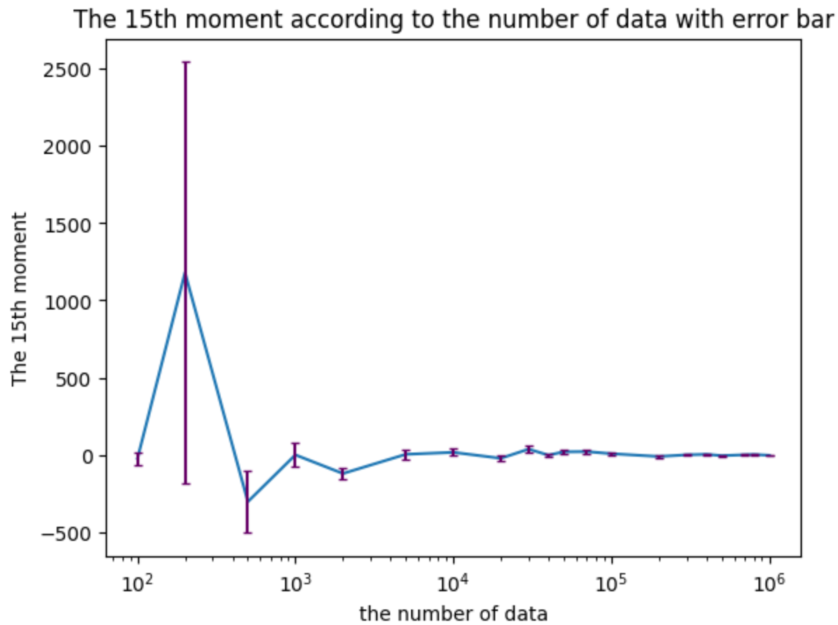

Figure 64: The $15th$ moment according to the number of data with error bar

2.2.12 Moment 20:

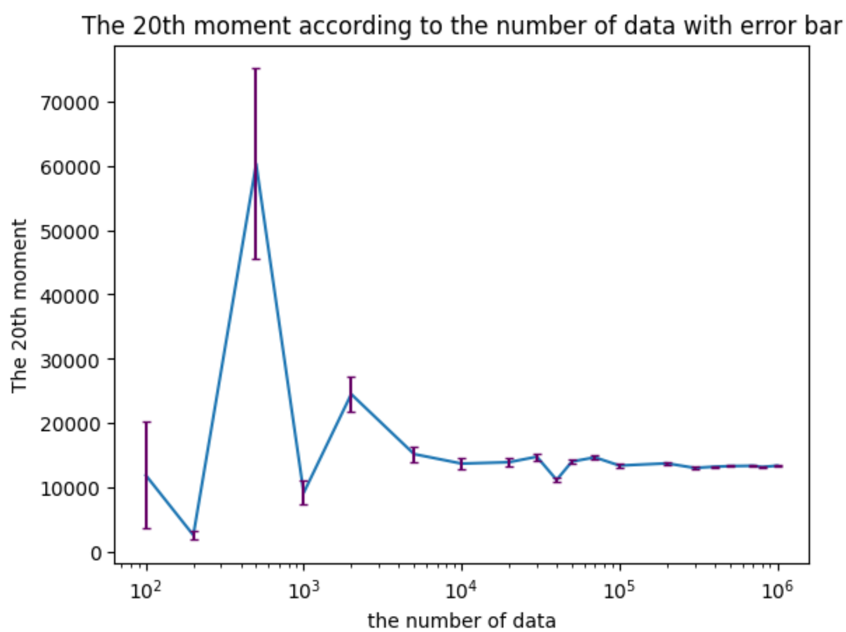

Figure 65: The $20th$ moment according to the number of data with error bar

3 Conclusion

As we have seen, we can only calculate up to $X^n P(X)$ in 10^n number of data, and more than that, in addition to a large increase in the error rate, the corners of the function behave exponentially, and the calculation above the limit of $X^n P(X)$ is illogical.

4 Code

4.1 Question 1:

```python
#QUESTION 1:
# Import needed libraries
import numpy as np
import matplotlib.pyplot as plt

# Generate n random numbers with normal distribution(for normal distribution
    run this block.)
n = int(input('Enter numbers of random data you want to gerenate: '))
random_number = np.random.normal(0, 1, 1000000)
data = np.random.choice(random_number, n , replace = False)
# number of bins we want
bins = int(input('Enter the number of bins you want:'))

# Calculate the bin centers and the histogram values and the bin wodth
hist_values, bin_edges = np.histogram(data, bins=bins)
bin_centers = (bin_edges[:-1] + bin_edges[1:]) / 2

# Calculate the error for each interval
def Error_for_each_interval(data, bins, bin_edges):
  errors = []
  for i in range(bins):
    numinbins = []
    for j in data:
      if(j >= bin_edges[i]  and j <= bin_edges[i+1]):
        numinbins.append(j)
        binmid = (bin_edges[i] + bin_edges[i+1]) /2
    if len(numinbins) != 0:
      mu = binmid * len(numinbins) / (n)
      sigma2 = ((binmid**2)*len(numinbins))- (n * (mu**2)) / (n-1)
      vari = (sigma2)**(1/2)
      filter_arr = []
      for element in data:
        if element >=bin_edges[i]  and element <= bin_edges[i+1]:
          filter_arr.append(element)
      # Remove Probability < 0
      if (len(filter_arr) - vari) <= 0:
          vari = len(filter_arr)
      errors.append(vari)
    if len(numinbins) == 0:
      errors.append(0.0)
  errors = np.array(errors)
  return errors

# Create the histogram
def plot(x,bins,bin_centers, hist_values,yerrors):

  plt.hist(x, bins=bins , color='#DCFFFF' , edgecolor='#3399FF')

  # Plot the distribution function
  plt.plot(bin_centers, hist_values, '#43438A', linewidth=2)

  # Plot the error bars
```

```python
  plt.errorbar(bin_centers, hist_values, yerr=yerrors, fmt='none', color='
    #660066', capsize=2)

  # Set the labels and title
  plt.xlabel('Number')
  plt.ylabel('Frequency')
  plt.title(f'The histogram of {n} data with normal distribution in {bins}
    intervals')

  # Display the plot
  plt.show()

plot(data, bins,bin_centers,hist_values , Error_for_each_interval(data, bins,
     bin_edges))

data = np.sort(data)

xpxlist = []
nxlist = []
dataorder = []
binmid = []
for i in range(bins):
  binmidbin = (bin_edges[i] + bin_edges[i+1]) /2
  binmid.append(binmidbin)
  numinbins = []
  for j in data:
    if(j >= bin_edges[i]  and j < bin_edges[i+1]):
      numinbins.append(j)
      dataorder.append(j)
  nx = len(numinbins)
  nxlist.append(nx)

pxlist = []
kol = sum(nxlist)
for element in nxlist:
  px = element / kol
  pxlist.append(px)

totalaverage = np.average(data)
errorsbins = Error_for_each_interval(data, bins, bin_edges)
sigmakol = (sum((errorsbins)**2)**(1/2))

yerrors = []
for i in range(len(binmid)):
  newerr = ((((binmid[i])**2)*((sigmakol/((kol)**2)))**2)+((errorsbins[i]/kol
    )**2))**(1/2)
  yerrors.append(newerr)

mnmnt = int(input('Enter numbers of moments you want to gerenate: '))
for x in range(0 ,mnmnt+1):
  xpxlist =[]
  yerrors1=[]
  for i in range(len(binmid)):
    xpx = (binmid[i]**x) * pxlist[i]
    yerrorstemp = abs((binmid[i]**x))*yerrors[i]
    xpxlist.append(xpx)
    yerrors1.append(yerrorstemp)
```

```
  binmid = np.sort(binmid)
  plt.plot(binmid , xpxlist)
  plt.plot(binmid, xpxlist, '#43438A', linewidth=2)
  plt.errorbar(binmid, xpxlist, yerr=yerrors1, fmt='none', color='#660066', capsize=2)
  plt.xlabel('X')
  plt.ylabel(f'x^{x}P(x)')
  plt.title(f'x^{x}P(x) by numbers for {n} data with normal distribution in {bins} intervals')
  plt.show()
```

4.2 Question 2:

```
#QUESTION 2:
# Import needed libraries
import numpy as np
import matplotlib.pyplot as plt
from scipy import stats

n_list = [100,200,500,1000,2000,5000,10000,20000,30000,40000,50000,70000,100000
                ,200000,300000,400000,500000,700000,800000,1000000]

random_number = np.random.normal(0, 1, 1000000)
random_number = np.sort(random_number)
moment_n = int(input('Enter numbers of moments you want to gerenate: '))
for n_m in range(moment_n+1):
  y_axis = []
  errors = []
  for n in n_list:
    data = np.random.choice(random_number, n , replace = False)
    vari = np.var(data)
    clean_data =[]
    for element in data:
      if element < 2 * vari and element > -2 * vari:
        clean_data.append(element)
    clean_data = np.sort(clean_data)
    y = stats.moment(clean_data, moment = n_m)
    y_axis.append(y)
    error = np.sqrt((np.sum(((clean_data**(n_m)) - y)**2)) / (len(clean_data)*(len(clean_data)-1)))
    errors.append(error)
  plt.plot(n_list,y_axis)
  plt.errorbar(n_list,y_axis,yerr=errors, fmt='none', color='#660066', capsize=2)
  plt.xscale("log")
  # Set the labels and title
  plt.xlabel('the number of data')
  plt.ylabel(f'The {n_m}th moment')
  plt.title(f'The {n_m}th moment according to the number of data with error bar')
  plt.show()
```