# CSV読み込み用

import pandas as pd

# 計算用

import numpy as np

from scipy.spatial import distance

# グラフ用

import matplotlib.pyplot as plt

%matplotlib inline

import seaborn as sns

sns.set\_style("whitegrid")

# データの読み込み

data = pd.read\_csv("MT\_data.csv")

data.head()

# データをMとS絞る

data = data[['M','S']]

plt.scatter(data['M'].values,data['S'].values)

# 平均

mean= np.mean(data, axis=0)

print(mean)

#標準偏差

#siguma=np.std(data, axis=0)

#print(siguma)

# 平均引いた値

data\_m = data - mean

print(data\_m)

#mean=mean- mean

print(mean)

#平均引いた値÷標準偏差

#data\_n = data\_m/siguma

#print(data\_n)

# 共分散

cov = np.cov(data.T)

print(cov)

print(np.linalg.pinv(cov))

print(mean)

# pythonの関数の計算

result\_1 = np.sqrt(np.sum(np.dot(data\_m,np.linalg.pinv(cov))\*data\_m, axis=1))

print(result\_1)

plt.plot(result\_1)

# distance.mahalanobisを使ってみた場合

data = data.to\_numpy()

data\_m =data\_m.to\_numpy()

result\_2 = []

# データ: data[i], 平均値: mean, 共分散行列の逆行列: np.linalg.pinv(cov) から距離を計算

for i in range(len(data\_m)):

 result\_2.append(distance.mahalanobis(data[i], mean, np.linalg.pinv(cov)))

# 結果

print(result\_2)

plt.plot(result\_2)