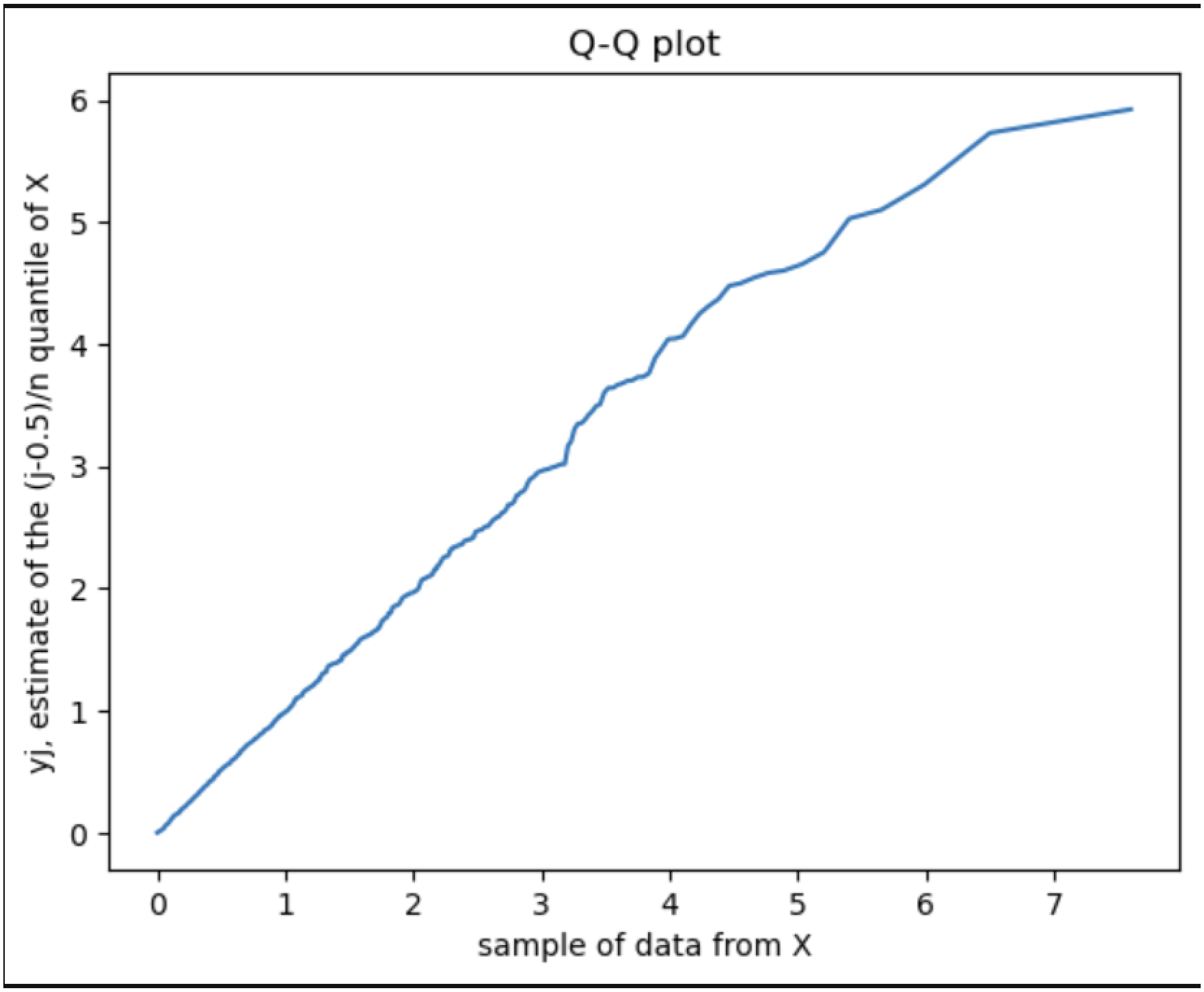
1. The result can be re-generated by running CSC446_a4_q1.ipynb
   a) We can observe approximately a straight line with a slope of 1. As a result, F or the random exponential distribution is a member of an appropriate family of distributions with appropriate parameter values



   b) Sample mean: 0.9924776997821552
   Sample variance: 0.9316510768175442
   The sample mean roughly matches the sample variance when mu = 1 because sample mean is equal to 1/mu (scale factor) in theory and sample variance is equal to 1/(mu^2) in theory. It matches the exponential distribution characteristics.

   c) Value of x0: 127.50802040130888
   Critical value for 100 degrees of freedom and 0.05 level of significance: 124.3
   H0, the null hypothesis, is rejected at the 0.05 level of significance

d) $X_{0.01,\ 100}$ = 135.8
$X_{0.10,\ 100}$ = 118.5

The null hypothesis remained rejected at 0.10 level of significance. However, the null is not rejected when the level of significance is changed to 0.01.

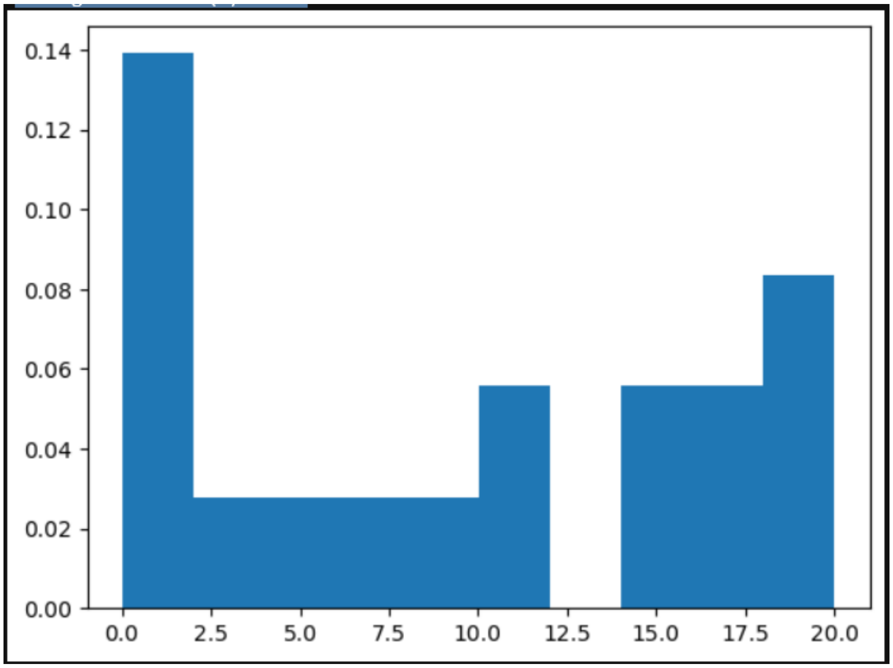
2. The result can be re-generated by running CSC446_a4_q2.ipynb
    a) Mean of the sample: 1.11
    Expected class: [32.9558961075189, 36.581044679345986, 20.302479797037023, 7.511917524903699, 2.0845571131607765, 0.4627716791216925, 0.08561276063751312]
    Value of x0: 15.527046432155455
    Critical value when class interval is 4 and the level of significance is 0.01: 13.28
    The null hypothesis is reject because X0 > X0.05, 4

    b) Mean of the sample(given): 1.0
    Expected class: [36.787944117144235, 36.787944117144235, 18.393972058572118, 6.1313240195240395, 1.5328310048810099, 0.30656620097620196, 0.051094366829366994]
    Value of x0: 25.11418844904105
    Critical value when class interval is 4 and the level of significance is 0.01: 13.28
    The null hypothesis is rejected because X0 > X0.05, 4

    c) Part b with a given mean produces a greater X0 value. Although the null hypothesis is rejected for both cases, part a) verifies if the sample is generated from the poisson distribution while no raw data is available (we only know the data of class intervals, and use the suggested estimator to determine x0). Part (b) is more of verifying if the sample is generated from the poisson distribution with a specific parameter.

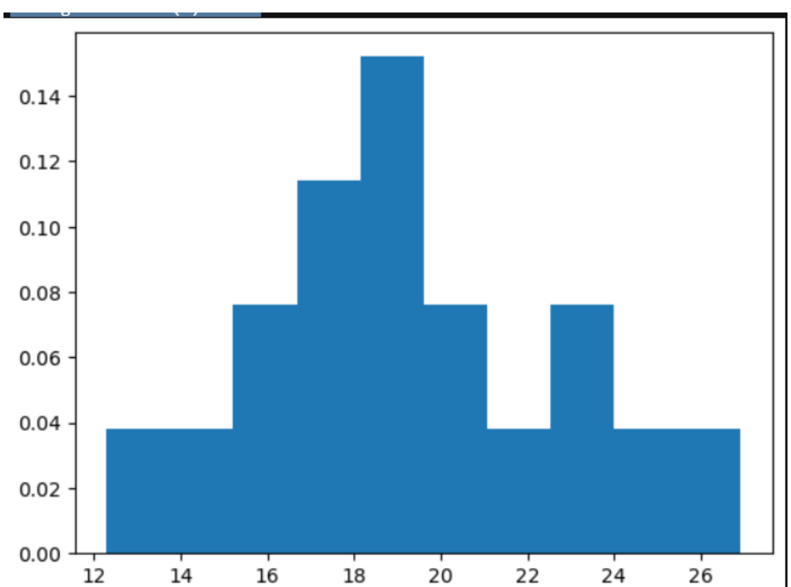
3. The result can be re-generated by running CSC446_a4_q3.ipynb
EAR(1) model
Sample mean: 20.27777777777778
Sample variance: 16.683006535947726
Sum of Xt*Xt+1 where t=[1, 18]: 7112
lag-1 autocovariance: 7.164669571532293
autocorrelation: 0.42945913592338486
mu: 0.049315068493150684

Histogram for EAR(1) model



AR(1) model
Sample mean: 20.27777777777778
Sample variance: 16.683006535947726
Sum of Xt*Xt+1 where t=[1, 18]: 7112
lag-1 autocovariance: 7.164669571532293
variance for AR(1) model: 13.6060737325809
autocorrelation: 0.42945913592338486
mu: 20.27777777777778

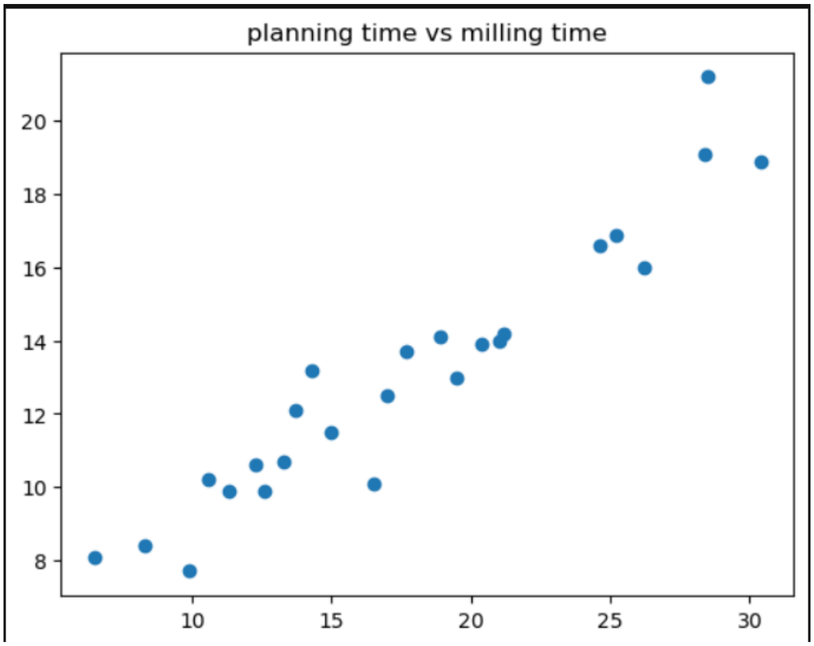
Histogram for AR(1) model



Comparing 2 histograms, AR(1) model fits better considering there are no visible gaps between intervals and the distribution is relatively identical to the normal distribution.

4. The result can be re-generated by running CSC446_a4_q4.ipynb

a)



From the scatter plot, we can observe that planning time has roughly a positive correlation. They are dependent of each other.

b)

Mean of milling time: 17.732

Variance of milling time: 45.03893333333341

Mean of planning time: 13.059999999999997

Variance of planning time: 12.684166666666746

Sum of $X_{plan,j}$ * $X_{mill,\ j}$ where j=[1,25]: 6337.92

Covariance between 2 samples: 22.850916666666762

Correlation between 2 samples: 0.9560456027353942

c)

X1 = u1 + o1 * Z1

X2 = u2 + o2 * (p * Z1 + Z2 * ((1-p**2) * 0.5))

Mean of milling time, u1: 17.732

Standard deviation of milling time, o1: 22.519466666666705

Mean of planning time, u2: 13.059999999999997

Standard deviation of planning time, u2: 6.342083333333373

Independent random number, R1: 0.13436424411240122

Independent random number, R2: 0.8474337369372327

Standard normal random variable, z1: 1.1534672601458011

Standard normal random variable, z2: -1.6426712518326385
Value of X1: 43.70746751594474
Value of X2: 19.60599075338846

5. The result can be re-generated by running CSC446_a4_q4.ipynb and the single server java files

a) 5000 samples are collected from running the single server java files. Performance data can found in the Q5_simulations_5000 samples file

Point estimator: 5.1190698
Standard deviation: 0.6274773473639065
Confidence interval: 5.1190698 +/- 0.4484422
Prediction interval: 5.1190698 +/- 1.4873146

b) Best case: 0.5675120171097143
Wrose case: 0.3293724171097141

Since the best case 0.568 is great than 0.5, additional replications are needed to reach a decision.

c) 50000 samples are collected from running the single server java files. Performance data can found in the Q5_simulations_50000_samples file

Point estimator: 5.488656000000001
Standard deviation: 0.31883488552749895
Confidence interval: 5.488656 +/- 0.227863
Prediction interval: 5.488656 +/- 0.755737
Best case: 0.7165192424876254
Wrose case: 0.26079275751237585

We can see that the best case is getting greater, and the worst case is getting smaller. However, additional replications are still needed to reach a decision. Simulations with different sample size would need a different close enough value