ENSE622, Spring 2018:
Homework 4: Monte Carlo Simulation
SOLUTIONS
(2/18/18)

Additional Instructions:

1. **This Homework has a maximum point value of 120 points (20 points EC).**
2. **For full credit all work must be shown.**

1. **Problem 1 (25 pts) – Static MC Project Cost Simulation:**
Consider the Project described in HW 3.2 that was characterized by the following activity edge table in which two people were required to complete each task (each working 8 hours a day).

Activity	Nodes	Duration Est (days)				Labor Cost ($/h)		
		Min	Max	Likely		Min	Max	Likely
A	1,2	4	8	6		$ 100	$ 150	$ 120
B	1,3	2	8	4		$ 100	$ 150	$ 120
C	2,4	1	7	3		$ 100	$ 150	$ 120
D	3,4	6	12	9		$ 100	$ 150	$ 120
F	3,5	5	15	10		$ 100	$ 150	$ 120
I	3,6	7	18	12		$ 90	$ 110	$ 100
J	4,7	5	12	9		$ 90	$ 110	$ 100
K	5,7	1	3	2		$ 90	$ 110	$ 100
L	6,8	2	6	3		$ 90	$ 110	$ 100
M	7,9	10	20	15		$ 90	$ 110	$ 100
N	8,9	6	11	9		$ 90	$ 110	$ 100

(a) (4) Provide a table showing each tasks expected task cost and associated standard deviation and expected total project and associated standard deviation (same as in HW 3.2).

Activity	Nodes	Task Cost		
		Mean	SD	VAR
A	1,2	$ 11,680.0	$ 95.3	$ 9,090.1
B	1,3	$ 8,435.6	$ 127.0	$ 16,123.5
C	2,4	$ 6,488.9	$ 124.9	$ 15,591.0
D	3,4	$ 17,520.0	$ 143.0	$ 20,444.4
F	3,5	$ 19,466.7	$ 219.3	$ 48,091.0
I	3,6	$ 19,466.7	$ 187.8	$ 35,268.1
J	4,7	$ 14,133.3	$ 120.4	$ 14,485.9
K	5,7	$ 3,200.0	$ 34.0	$ 1,157.8
L	6,8	$ 5,333.3	$ 67.6	$ 4,572.3
M	7,9	$ 24,000.0	$ 174.0	$ 30,288.9
N	8,9	$ 14,133.3	$ 88.4	$ 7,817.0
Total Project Cost =		$ 143,857.8	$ 450.5	$ 202,930.1

(b) (4) Develop and submit a MATLAB program that:

a. Accepts the min, max, and likely hours per task and the min, max, and likely cost per labor hour for each task, the number of histogram bins (Nb), and the number of samples (Ns).
b. Assumes a **triangular distribution** for hours per task and cost per labor hour.
c. Provides the following output:
    i. The MC mean total cost of the project, the associated standard deviation, the associated standard error in the mean, the cumulative running mean (CRM), and the CPU time taken.
    ii. A histogram graph showing the probability distribution associated the project's total cost.
    iii. A graph of the CRM.

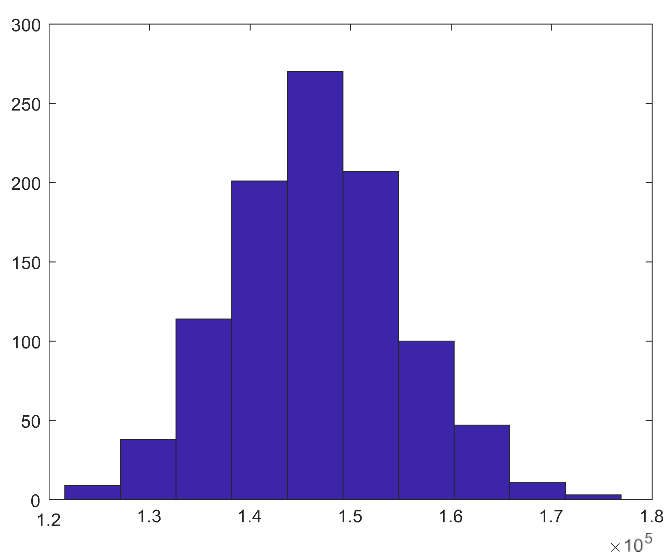
See HW4_1_MC_Cost_Dist_Tri.

(c) (4) Run the program for Nb=10 and Ns = 20, 200, and 2000 and provide the resulting output below (text output should go in a table).

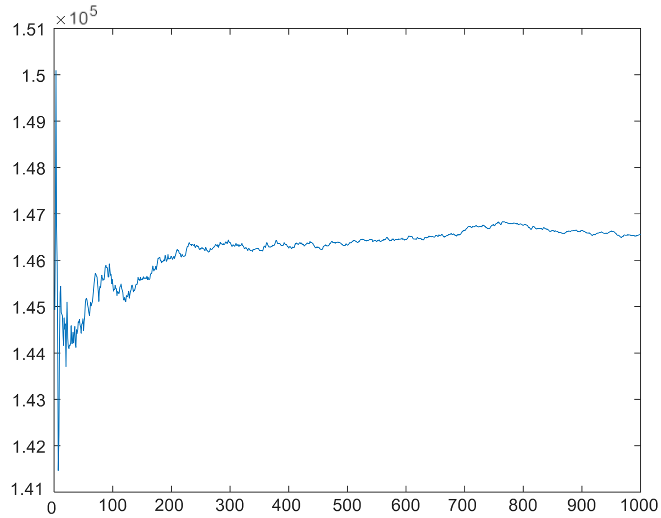
a. Text Output:

Ns = 1000
ProjCostmean = 1.4656e+05
ProjCostcumrm = 1.4656e+05
SD = 8.6017e+03
SE = 272.0099
>>

b. Graphical Output:



c. **Note that original problem was ambiguous as to whether to use normal or triangular within the simulation. As such their SD and SE may vary somewhat from this.**



(d) (3) What do you conclude from the results table?

   a. **Project Cost** appears to reach a reasonably stable mean of $146,560 +/- $270 (SE) after about 1000 runs and appears to have a standard deviation of ~ $8,600.

(e) (3) Are these results significantly different from the analytical result for project cost that you obtained in HW 3.2? Explain your answer.

   a. While the value of the mean is consistent with that found in HW 3.2 (it is within a couple SEs), the SD is larger by an order of magnitude. This is due to the fact that the beta distribution is much narrower than the triangular distribution.

(f) (1) What does the graph of the CRM tell you?

   a. It appears one needs ~400+ samples to get a stable estimate for the mean.

(g) (2) What kind of distribution is suggested by the histogram? Explain. Is this surprising? Explain.

   a. Cost distribution looks normal.
      i. It has tails (so not triangular).
      ii. It appears to be symmetric.
   b. Since the individual task durations are triangular, one might expect the result to be triangular. But it is not. So it is a bit surprising.

(h) (3) What would you report to your manager as the expected project cost? Be sure to give all the information required.

   a. Project Cost is likely to be $146,600 +/- $8,600 (SD).
   b. The estimate of the mean is good to +/- $270 (SE)

(i) (2) What is the 80% confidence project cost?

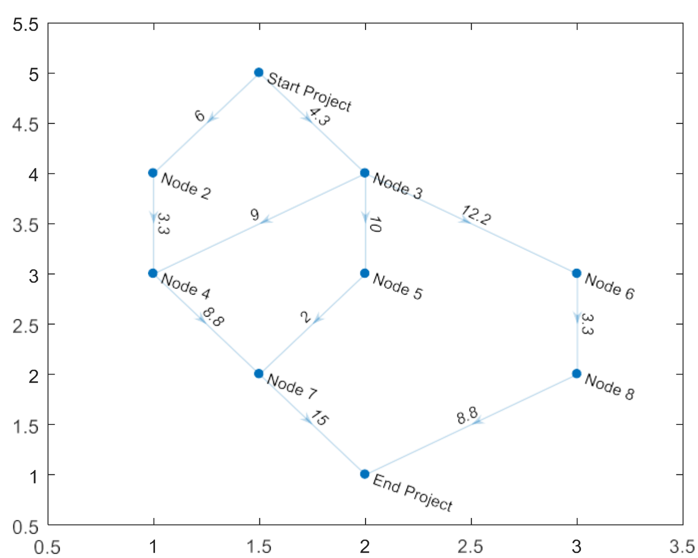
   a. 80% confidence project cost ~ $154,000 (using Excel NORM.INV() function).

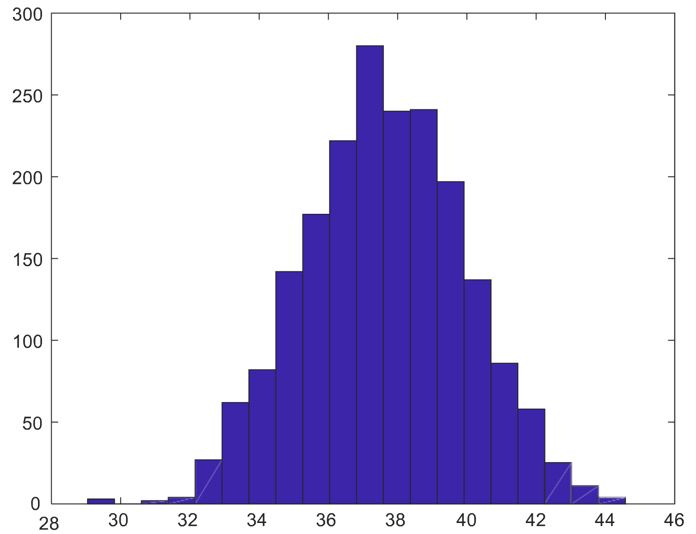
2. **Problem 2 (20 pts) – Static MC Critical Path Simulation:**
   (a) (2) Using the information from Problem 1, provide a table showing the project activities and the expected mean duration and standard deviation for each activity.
   (b) (4) Develop and submit a MATLAB program that:
      a. Accepts the expected hours per mean duration and standard deviation for each task.
      b. Assumes a normal distribution of activity durations.
      c. Note that you will have to find the longest path for each sample (since it may change).
      d. Provides the following output:
         i. The MC mean total duration of the project, the associated standard deviation, the associated standard error in the mean, the CRM, and the CPU time taken.
         ii. A histogram graph showing the distribution of project durations.
         iii. A graph of the CRM.

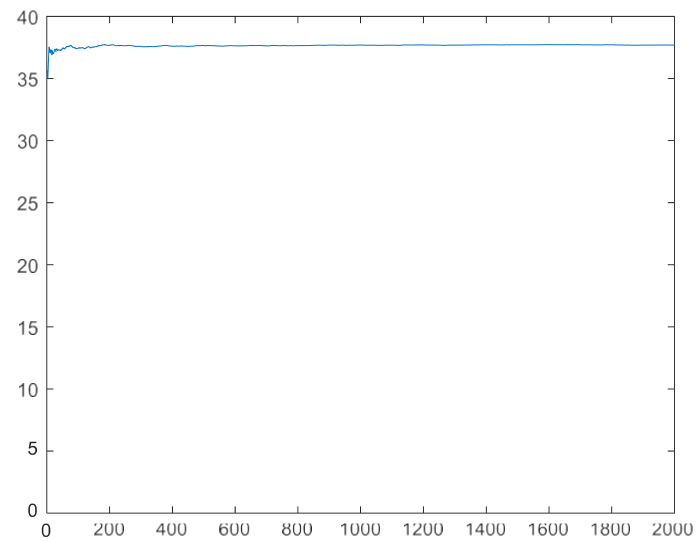
   See HW_4_2_MC_CP
   (c) (4) Run the program for Nb=10 and Ns = 20, 200, and 2000 and provide the resulting output below (text output should go in a table).
      a. Text Output for Ns = 2000:

```
Crit_Path_Len =  37.4500
SD =    2.3022
SE =    0.0728
NSamePath =  1978
>>
```

      b. Graphical Output:







(d) (2) What do you conclude from the results table?

a. The more samples taken,
   i. The more accurate the mean becomes (i.e., the smaller the SE).
   ii. The longer the simulation takes to run.
b. The SD does not change significantly with increasing Ns.
c. One generally needs Ns ~1000 times to get meaningful statistics on alternative paths.

(e) (2) Are these results significantly different from the analytical result for critical path that you obtained in HW 3.2? Explain your answer.

a. The analytical results for HW 3.2 were:
   i. CP Length = 37.1 d +/- 2.5 d.
   ii. CP = 1,1, 3,3,3,4,6,7,9
b. The MC results are:

        i. Crit_Path_Len = 37.45 +/- 2.3 d
        ii. SE in mean is 0.07 d
        iii. NSamePath = 1978 of the MC paths were the same as for the analytical case.
    c. Analysis:
        i. The difference is means is statistically significant (given the SE for the MC mean)
        ii. The reason for the difference is that the MC simulation gives rise to alternative CPs.

(f) (1) What does the graph of the CRM tell you?
    a. The CRM tells you how stable the MC estimate of the mean is.

(g) (2) What kind of distribution is suggested by the histogram? Explain. Is this surprising? Explain.
    a. Normal
    b. Since the individual task durations are triangular, one might expect the result to be triangular. But it is not. So it is a bit surprising.

(h) (2) What would you report to your manager as the estimated duration of the project? Be sure to give all the information required.
    a. There is a 50% chance that the Project can be completed in 37.5 days.
    b. The uncertainty associated with this estimate is ~2.3 days.

(j) (1) What is the 80% confidence project duration?
    a. 80% confidence project cost ~ 39.4 days (using Excel NORM.INV() function).

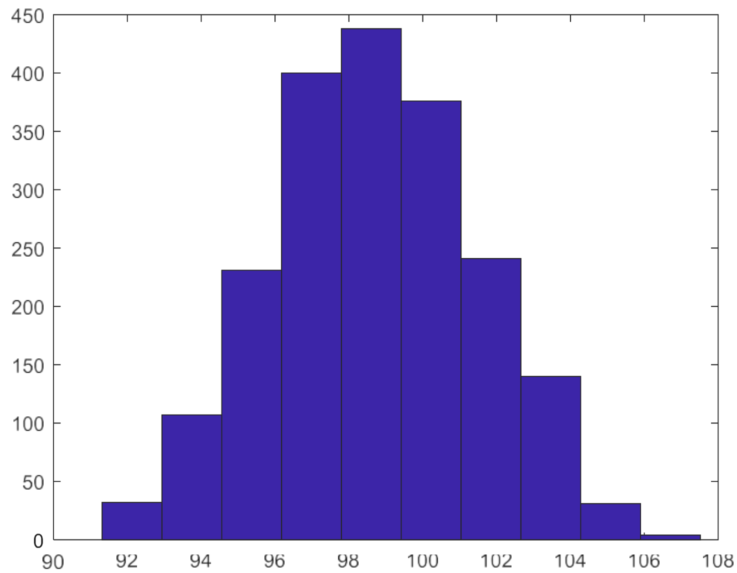
3. **Problem 3 (20 pts) – Static MC Shortest Path Sensitivity Analysis:**
Consider the communications network described in HW 3.1. Assume that one is uncertain about each of the costs associate with each link by +/- 10%. Assume a uniform distribution for the cost of each link.

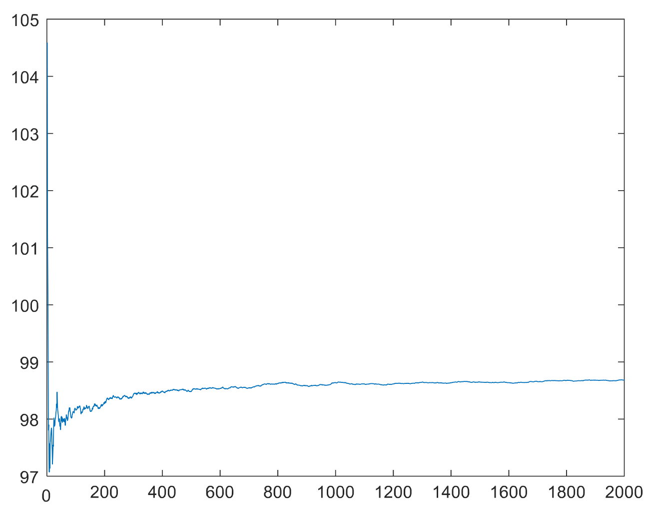
Link	Cost ($/Tb)		Link	Cost ($/Tb)		Link	Cost ($/Tb)
So-N1	**60**		**N2-N4**	**50**		**N5-N4**	**10**
So-N2	**30**		**N2-N5**	**70**		**N3 -Si**	**50**
N1-N2	**10**		**N3-N4**	**10**		**N4-Si**	**30**
N2-N1	**10**		**N4-N3**	**10**		**N5-Si**	**10**
N1-N3	**30**		**N4-N5**	**10**			

(a) (5) Develop and submit a MATLAB program that:
   a. Accepts the upper and lower bounds for the costs of each link, the number of bins, and the number of samples.
   b. Note that the shortest path for each sample (since it may change from sample to sample).
   c. Provides the following output:
      i. The MC mean cost, the associated standard deviation, the associated standard error in the mean, the CRM, and the CPU time taken.
      ii. The number of times the shortest path differed from the analytical shortest path (from HW 3.1)
      iii. A histogram graph showing the distribution of project durations.
      iv. A graph of the CRM.

   See HW_4_3_MC_Short_Path.

(b) (5) Run the program Nb=10 and for Ns = 20, 200, and 2000 and provide the resulting output below (text output should go in a table).
   a. Text Output:
      Ns = 2000
      Nb = 10
      MeanPathLen = 98.6802
      SD = 2.7692
      SE = 0.0619
      >>
   b. Graphical Output:





(c) (2) What do you conclude from the results table?

a. The more samples taken,
   i. The more accurate the mean becomes (i.e., the smaller the SE).
   ii. The longer the simulation takes to run.
b. The SD does not change significantly with increasing Ns.
c. One generally needs Ns ~1000 times for the CRM to settle down and to get meaningful statistics.

(d) (2) Are these results significantly different from the analytical result you obtained in HW 3.1? Explain your answer.

a. Yes, the results are significantly different.
b. The MC mean shortest path is $98.68 /Tb +/- $0.06 Tb (SE).
c. The analytical result is $100 /Tb
d. The difference between the two exceeds 3* SE.

e. This is because the MC result allows alternative paths.

(i) (1) What does the graph of the CRM tell you?

a. The CRM tells you how stable the MC estimate of the mean is.

(e) (2) What kind of distribution is suggested by the histogram? Explain. Is this surprising? Explain.

a. Normal
b. Since the individual links have uniform distributions, one might expect the result to be a uniform distribution. But it is not. So it is a bit surprising.

(f) (3) What would you report to your manager as the estimated route and cost for transmitting data? Be sure to give all the information required.

a. The mean shortest path is expected to cost $98.7/Tb +/- $2.8 Tb (SD).
b. It is unclear which path will be shortest.
c. A separate study is needed to determine the uncertainty associated with the using the analytical solution's shortest path.

4. **Problem 4 (25 pts) – Time-Driven Dynamic MC Simulation of an AV:**
Consider an autonomous vehicle (AV) attempting to move in the "X direction" subject to random gusts of wind (blowing in the +/- Y direction). Assume the AV has a mass of 2 kg and is traveling at a speed of 60 km/h. Assume that the impulse (change in momentum) imparted to the AV by the wind gusts over a period of 5 seconds is characterized by a normal distribution with a mean of 0 kg km/h and a standard deviation of 10 kg km/hr.

(a) (2) Draw a schematic diagram of the system showing the AV, its motion, and the effect of the wind.

(b) (2) Write down the equations of motion (in X and Y directions) for dt = 5 sec.
   a. Random impulse (change in momentum):
      i. dPy(i) = m dVy(i) = m ay(i) dt
      ii. ay(i) = dVy(i)/dt
   b. X direction:
      i. Vx(n) = Vxo
      ii. X(n) = X(n-1) + Vx(n) dt
   c. Y direction:
      i. Vy(n) = Vy(n-1) + a(n) dt = Vy(n-1) + dVy(n-1)
      ii. Y(n) = Y(n-1) + Vy(n) dt + ½ ay(n) dt^2
      = Y(n-1) + Vy(n) dt + ½ dVy(n) dt

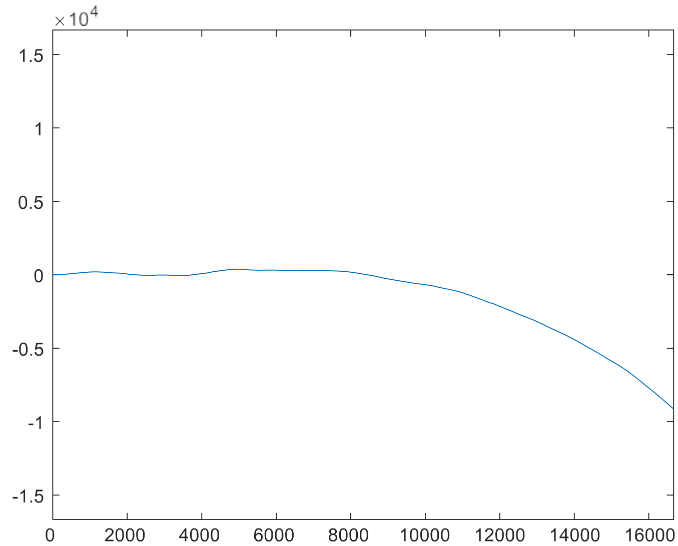
(c) (3) Develop and submit a MATLAB program that:
   a. Accepts the speed and mass of the AV, the time step size, the mean and standard deviation of the impulse delivered to the AV, the duration of the simulation (Ts), and number of bins. All units should be converted to MKS.
   b. Provides the following output:
      i. The final location of the AV (at t = Ts).
      ii. A 2-D plot of the trajectory of AV (at dt time intervals).

See HW_4_4_c_MC_AV_Wind_1_Run and HW_4_4_e_MC_AV_Wind_Ns_Run.

(d) (3) Run the program for Ts = 100 s and 1,000 s and provide the resulting output below (text output should go in a table).
   a. Text Output:

```
Nts =  200
EndTime    =    1005
YEnd  =  -9.1517e+03
Xend = 1.6667e+04
>>
```

   b. Graphical Output:



(e) (4) Develop and submit a MATLAB program that:

a. Accepts the speed and mass of the AV, the time step size, the mean and standard deviation of the impulse delivered to the AV, the duration of the simulation (Ts), the number of samples (Ns), and number of bins (Nb). All units should be converted to MKS.

b. Provides the following output:

i. The mean ending Y coordinate (in m), its associated standard deviation, and standard error in the mean.

ii. A histogram graph showing the distribution of ending Y locations.

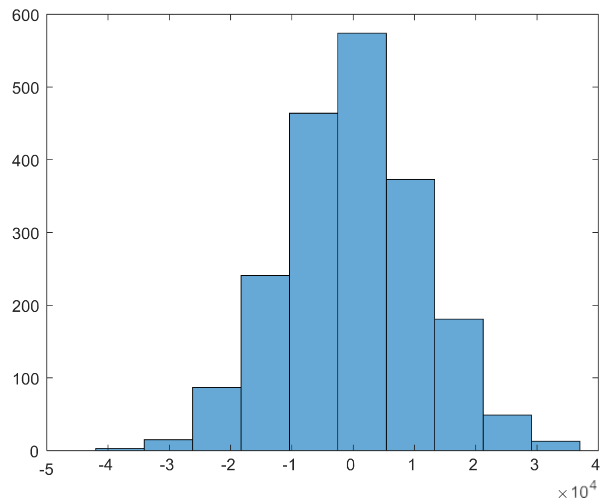
See MC_AV_Wind DistTime.

(f) (4) Run the program Ts = 1000 s and Nb=10 and provide the resulting output below (text output should go in a table).

a. Text Output for Results:

Tend = 1000
dt = 5
Nts = 200
Ns = 2000
Nb = 10
Yfmean = 88.1627
SDYf = 1.1286e+04
SEYfm = 252.3525
>>

b. Graphical Output:



(g) (2) What do you conclude from the results table?

a. The mean value for Yend ~ 0 m (88 m +/- 252 m).
b. It has a very large SD (~11,300 m)
c. This means that in most runs the AV will be blown off course by almost as much as the x distance it travels.

(h) (2) Are these results consistent with the result you got from the simulation you developed in (c) and ran in (d)? Explain.

a. Yes, in running the model in (c) and (d) multiple times, most of the time it was blown way off from the Y = 0 by an amount that was comparable to the X distance gone.

(i) (2) What kind of distribution is suggested by the histogram? Explain. Is this surprising? Explain.

a. It appears to be symmetric and normal
b. This is not surprising since the wind distribution is normal.

(j) (1) How might this kind of simulation be used to evaluate the effectiveness of a self-correcting guidance algorithm?

a. It would provide a model for the system and the environment that could be used to test the self-correcting guidance algorithm.

5. **Problem 5 (20 pts) – System Availability Event-Driven Dynamic MC Simulation:** Consider a system consisting of a single autonomous vehicle (AV) that is supposed to provide 24/7365 “coverage” of some region and that it is characterized by a mean time between failure (MTBF) of 100 hours and mean time to repair (MTTR) of 20 hours. For simplicity, assume that as soon as it fails, it goes into maintenance and that as soon as it is repaired it returns to operation. Assume an exponential distribution of failure times and repair times.

   (k) (2) Calculate the expected system’s steady state operational availability (Aoss) analytically.

      a. Ao = MTBF/(MTBF+MTTR) = 100/120 = 0.833

   (a) (4) Develop and submit an event-driven MATLAB stochastic simulation that:

      a. Accepts the AV’s MTBF, MTTR, number of bins, and number of failures (Nf).
      b. Runs until all failure have occurred and been fixed.
      c. Provides the following output:
         i. The MC mean system availability, the associated standard deviation, the associated standard error in the mean, the CRM, and the CPU time taken.
         ii. A histogram graph showing the distribution of Ao
         iii. A graph of the CRM.

      See HW_4_5_MC_Ao_1_Elem_DE_Sim

      (Note that this is a minor variation of the program provided in Slide 33 of Lecture 4.5 (with almost the same name).

   (b) (4) Run the program for Nf = 20, 200, and 2000 and provide the resulting output below (text output should go in a table).

      a. Text Output:

         Inputs:

            MTBFa = 100
            MDTa = 20
            Nf = 2000
            Nr = 2000

         Intermediate Calculations:

            MTBFmc = 96.3009
            SDtbcf = 97.7065
            SEtbcf = 2.1848
            MDTmc = 20.4437
            SDdt = 19.4236
            SEdr = 0.4343
            T = 2.3349e+05
            Aossan = 0.8333

Results:

Aossmc = 0.8249

MeanAomc = 0.8249

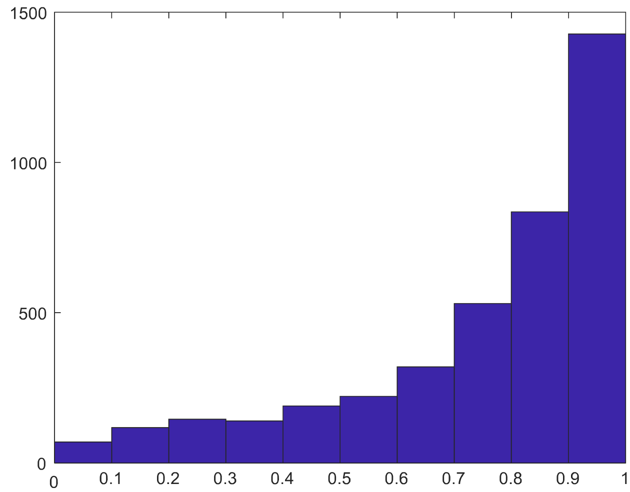
SDao = 0.2487

SEao = 0.0056

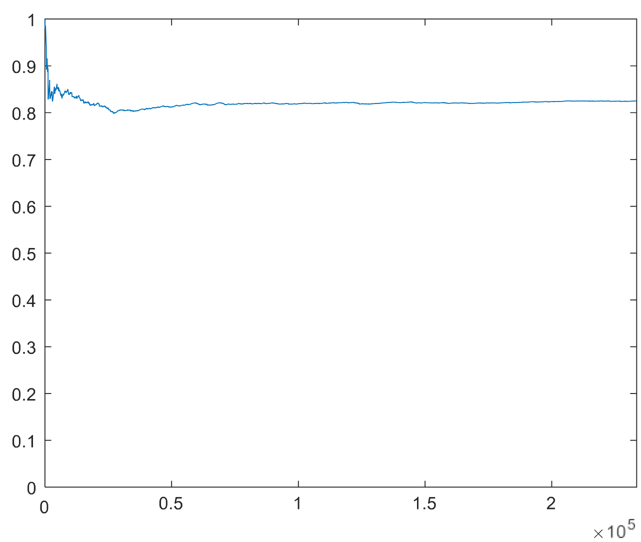
Elapsed time is 1.870890 seconds.

>>

b. Graphical Output



Note that histogram double counts each availability.



Note that x-axis is in time, not failures. Since the average up/down cycle is 120 hr, total time is ~ 2000 failure * 120 hr/failure.

(g) (2) What do you conclude from the results table?
   a. The more samples taken,
      i. The more accurate the mean becomes (i.e., the smaller the SE).
      ii. The longer the simulation takes to run.
   b. The SD does not change significantly with increasing Ns.
   c. One generally needs Nf ~500-1000 for the CRM to settle down and to get meaningful statistics.

(c) (2) Are these results significantly different from the analytical result you obtained in part (a)? Explain your answer.
   a. Aossan = 0.8333
   b. Aossmc = 0.8249 +/- 0.006 (SE)
   c. Since the MC result is within 3 SEs of the analytical result, there is no significant difference.

(j) (2) What does the graph of the CRM tell you?
   a. The CRM tells you how stable the MC estimate of the mean is.

(d) (2) What kind of distribution is suggested by the histogram? Explain. Is this surprising? Explain.
   a. It appears to be “exponential” over the range Ao = 0-1.
   b. This is surprising, given the distributions in all the other problems appeared to be normal.

(h) (2) What would you report to your manager as the estimated operational availability of the AV? Be sure to give all the information required.
   a. The mean steady state system availability is expected to be 0.825 with a standard deviation of ~ 0.25.
   b. The uncertainty associated with value of the long term mean is 0.006 (SE) (or 0.02 for 3*SE).