```
Last login: Tue Apr 21 10:49:53 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-54-236-240-251.compute-1.amazonaws.com
Warning: Permanently added the ED25519 host key for IP address '54.236.240.251' to the list of known hosts.
Welcome to Ubuntu 18.04.3 LTS (GNU/Linux 4.15.0-1060-aws x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

  System information as of Tue Apr 21 14:51:46 UTC 2020

  System load:  1.15               Processes:           884
  Usage of /:   28.5% of 30.96GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.7.244
  Swap usage:   0%

 * Kubernetes 1.18 GA is now available! See https://microk8s.io for docs or
   install it with:

     sudo snap install microk8s --channel=1.18 --classic

 * Multipass 1.1 adds proxy support for developers behind enterprise
   firewalls. Rapid prototyping for cloud operations just got easier.

     https://multipass.run/

 * Canonical Livepatch is available for installation.
   - Reduce system reboots and improve kernel security. Activate at:
     https://ubuntu.com/livepatch

91 packages can be updated.
40 updates are security updates.


Last login: Fri Apr  3 19:45:17 2020 from 107.13.161.147
ubuntu@ip-172-31-7-244:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
10:53, 04/21; num of cores:96
final_T_large

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_D_range, t_func] = [96, None, None, 25, 0.5, 1.5, [80], None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[100, 101, 105, 110], [15], [6, 7, 8], [[0.0003, 0.0001, 5e-05], [0.0003, 0.0001,
5e-05]]]

---------------------------------------
[pattern_seed, day, sd_R, u_D] = [2, 6, 15, 80]

max(u_O) =  157.3 mean(u_O) =  93.7
O_threshold = 100
means of Order:

89.6 98.6 46.6 141.0 55.2

79.0 112.6 68.9 73.6 77.3

113.8 157.3 101.0 72.1 113.5

85.1 99.5 129.4 81.3 100.2

78.0 96.1 106.4 75.3 91.5

target policy:

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 1

0 0 1 0 0

number of reward locations:  9
O_threshold = 101
number of reward locations:  8
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

---------------------------------------
Value of Behaviour policy:50.643
O_threshold = 100
```

```
MC for this TARGET:[57.983, 0.207]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.73, -1.92, -2.23]][[-2.98, -57.98, -7.34]]
std:[[0.9, 0.88, 0.61]][[0.47, 0.0, 0.32]]
sd_MSE:[array([0.34, 0.36, 0.28])][array([0.29, 0.  , 0.47])]
MSE:[[3.8, 4.46, 5.34]][[9.1, 3361.68, 53.98]]
RMSE(-DR):[[0.0, 0.66, 1.54]][[5.3, 3357.88, 50.18]]
***
==============
O_threshold = 101
MC for this TARGET:[55.16, 0.209]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.06, -0.21, -0.73]][[-2.08, -55.16, -4.52]]
std:[[0.94, 0.93, 0.57]][[0.45, 0.0, 0.32]]
sd_MSE:[array([0.11, 0.12, 0.09])][array([0.19, 0.  , 0.29])]
MSE:[[0.89, 0.91, 0.86]][[4.53, 3042.63, 20.53]]
RMSE(-DR):[[0.0, 0.02, -0.03]][[3.64, 3041.74, 19.64]]
**
==============
O_threshold = 105
MC for this TARGET:[56.952, 0.212]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.51, -3.67, -3.92]][[-5.42, -56.95, -6.31]]
std:[[1.01, 0.98, 0.63]][[0.44, 0.0, 0.32]]
sd_MSE:[array([0.78, 0.78, 0.51])][array([0.49, 0.  , 0.41])]
MSE:[[13.34, 14.43, 15.76]][[29.57, 3243.3, 39.92]]
RMSE(-DR):[[0.0, 1.09, 2.42]][[16.23, 3229.96, 26.58]]
***
==============
O_threshold = 110
MC for this TARGET:[55.704, 0.207]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.26, -3.37, -3.79]][[-6.28, -55.7, -5.06]]
std:[[0.95, 0.93, 0.67]][[0.45, 0.0, 0.32]]
sd_MSE:[array([0.67, 0.68, 0.52])][array([0.59, 0.  , 0.33])]
MSE:[[11.53, 12.22, 14.81]][[39.64, 3102.49, 25.71]]
RMSE(-DR):[[0.0, 0.69, 3.28]][[28.11, 3090.96, 14.18]]
***
==============
[[3.8000e+00 4.4600e+00 5.3400e+00 9.1000e+00 3.3617e+03 5.3980e+01]
 [8.9000e-01 9.1000e-01 8.6000e-01 4.5300e+00 3.0426e+03 2.0530e+01]
 [1.3340e+01 1.4430e+01 1.5760e+01 2.9570e+01 3.2433e+03 3.9920e+01]
 [1.1530e+01 1.2220e+01 1.4810e+01 3.9640e+01 3.1025e+03 2.5710e+01]]


time spent until now: 71.5 mins

12:04, 04/21

---------------------------------------
[pattern_seed, day, sd_R, u_D] = [2, 7, 15, 80]

max(u_O) =  157.3 mean(u_O) =  93.7
O_threshold = 100
means of Order:

89.6 98.6 46.6 141.0 55.2

79.0 112.6 68.9 73.6 77.3

113.8 157.3 101.0 72.1 113.5

85.1 99.5 129.4 81.3 100.2

78.0 96.1 106.4 75.3 91.5

target policy:

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 1

0 0 1 0 0

number of reward locations:  9
O_threshold = 101
number of reward locations:  8
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!
```

```
--------------------------------------
Value of Behaviour policy:50.643
O_threshold = 100
MC for this TARGET:[58.003, 0.174]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.53, -1.74, -2.14]][[-2.92, -58.0, -7.36]]
std:[[0.91, 0.88, 0.59]][[0.51, 0.0, 0.28]]
sd_MSE:[array([0.32, 0.34, 0.26])][array([0.31, 0.  , 0.41])]
MSE:[[3.17, 3.8, 4.93]][[8.79, 3364.0, 54.25]]
RMSE(-DR):[[0.0, 0.63, 1.76]][[5.62, 3360.83, 51.08]]
***
==============
O_threshold = 101
MC for this TARGET:[55.18, 0.172]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.16, -0.03, -0.66]][[-2.06, -55.18, -4.54]]
std:[[0.91, 0.89, 0.55]][[0.5, 0.0, 0.28]]
sd_MSE:[array([0.11, 0.11, 0.09])][array([0.22, 0.  , 0.25])]
MSE:[[0.85, 0.79, 0.74]][[4.49, 3044.83, 20.69]]
RMSE(-DR):[[0.0, -0.06, -0.11]][[3.64, 3043.98, 19.84]]
**
==============
O_threshold = 105
MC for this TARGET:[56.973, 0.171]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.41, -3.57, -3.96]][[-5.42, -56.97, -6.33]]
std:[[1.01, 0.99, 0.63]][[0.49, 0.0, 0.28]]
sd_MSE:[array([0.75, 0.76, 0.53])][array([0.54, 0.  , 0.36])]
MSE:[[12.65, 13.72, 16.08]][[29.62, 3245.58, 40.15]]
RMSE(-DR):[[0.0, 1.07, 3.43]][[16.97, 3232.93, 27.5]]
***
==============
O_threshold = 110
MC for this TARGET:[55.724, 0.169]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.26, -3.37, -3.85]][[-6.28, -55.72, -5.08]]
std:[[1.03, 1.0, 0.68]][[0.47, 0.0, 0.28]]
sd_MSE:[array([0.67, 0.69, 0.53])][array([0.6 , 0.  , 0.28])]
MSE:[[11.69, 12.36, 15.28]][[39.66, 3104.72, 25.88]]
RMSE(-DR):[[0.0, 0.67, 3.59]][[27.97, 3093.03, 14.19]]
***
==============
[[3.8000e+00 4.4600e+00 5.3400e+00 9.1000e+00 3.3617e+03 5.3980e+01]
 [8.9000e-01 9.1000e-01 8.6000e-01 4.5300e+00 3.0426e+03 2.0530e+01]
 [1.3340e+01 1.4430e+01 1.5760e+01 2.9570e+01 3.2433e+03 3.9920e+01]
 [1.1530e+01 1.2220e+01 1.4810e+01 3.9640e+01 3.1025e+03 2.5710e+01]]


[[3.1700e+00 3.8000e+00 4.9300e+00 8.7900e+00 3.3640e+03 5.4250e+01]
 [8.5000e-01 7.9000e-01 7.4000e-01 4.4900e+00 3.0448e+03 2.0690e+01]
 [1.2650e+01 1.3720e+01 1.6080e+01 2.9620e+01 3.2456e+03 4.0150e+01]
 [1.1690e+01 1.2360e+01 1.5280e+01 3.9660e+01 3.1047e+03 2.5880e+01]]


time spent until now: 149.1 mins

13:22, 04/21

--------------------------------------
[pattern_seed, day, sd_R, u_D] = [2, 8, 15, 80]

max(u_O) =  157.3 mean(u_O) =  93.7
O_threshold = 100
means of Order:

89.6 98.6 46.6 141.0 55.2

79.0 112.6 68.9 73.6 77.3

113.8 157.3 101.0 72.1 113.5

85.1 99.5 129.4 81.3 100.2

78.0 96.1 106.4 75.3 91.5

target policy:

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 1

0 0 1 0 0

number of reward locations:  9
```

```
O_threshold = 101
number of reward locations:  8
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
target 1 in 4 DONE!
 target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

---------------------------------------
Value of Behaviour policy:50.657
O_threshold = 100
MC for this TARGET:[57.987, 0.182]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.6, -1.81, -2.18]][[-2.97, -57.99, -7.33]]
std:[[0.81, 0.78, 0.51]][[0.46, 0.0, 0.24]]
sd_MSE:[array([0.27, 0.29, 0.23])][array([0.29, 0.  , 0.36])]
MSE:[[3.22, 3.88, 5.01]][[9.03, 3362.84, 53.79]]
RMSE(-DR):[[0.0, 0.66, 1.79]][[5.81, 3359.62, 50.57]]
***
==============
O_threshold = 101
MC for this TARGET:[55.163, 0.181]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.04, -0.15, -0.71]][[-2.07, -55.16, -4.51]]
std:[[0.79, 0.77, 0.5]][[0.46, 0.0, 0.24]]
sd_MSE:[array([0.09, 0.08, 0.08])][array([0.2 , 0.  , 0.22])]
MSE:[[0.63, 0.62, 0.75]][[4.5, 3042.63, 20.4]]
RMSE(-DR):[[0.0, -0.01, 0.12]][[3.87, 3042.0, 19.77]]
***
==============
O_threshold = 105
MC for this TARGET:[56.96, 0.177]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.39, -3.56, -3.91]][[-5.4, -56.96, -6.3]]
std:[[0.87, 0.86, 0.49]][[0.45, 0.0, 0.24]]
sd_MSE:[array([0.62, 0.63, 0.39])][array([0.5 , 0.  , 0.31])]
MSE:[[12.25, 13.41, 15.53]][[29.36, 3244.44, 39.75]]
RMSE(-DR):[[0.0, 1.16, 3.28]][[17.11, 3232.19, 27.5]]
***
==============
O_threshold = 110
MC for this TARGET:[55.708, 0.178]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.15, -3.26, -3.72]][[-6.28, -55.71, -5.05]]
std:[[0.86, 0.86, 0.52]][[0.46, 0.0, 0.24]]
sd_MSE:[array([0.55, 0.56, 0.4 ])][array([0.59, 0.  , 0.25])]
MSE:[[10.66, 11.37, 14.11]][[39.65, 3103.6, 25.56]]
RMSE(-DR):[[0.0, 0.71, 3.45]][[28.99, 3092.94, 14.9]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[3.8000e+00 4.4600e+00 5.3400e+00 9.1000e+00 3.3617e+03 5.3980e+01]
 [8.9000e-01 9.1000e-01 8.6000e-01 4.5300e+00 3.0426e+03 2.0530e+01]
 [1.3340e+01 1.4430e+01 1.5760e+01 2.9570e+01 3.2433e+03 3.9920e+01]
 [1.1530e+01 1.2220e+01 1.4810e+01 3.9640e+01 3.1025e+03 2.5710e+01]]


[[3.1700e+00 3.8000e+00 4.9300e+00 8.7900e+00 3.3640e+03 5.4250e+01]
 [8.5000e-01 7.9000e-01 7.4000e-01 4.4900e+00 3.0448e+03 2.0690e+01]
 [1.2650e+01 1.3720e+01 1.6080e+01 2.9620e+01 3.2456e+03 4.0150e+01]
 [1.1690e+01 1.2360e+01 1.5280e+01 3.9660e+01 3.1047e+03 2.5880e+01]]


[[3.2200e+00 3.8800e+00 5.0100e+00 9.0300e+00 3.3628e+03 5.3790e+01]
 [6.3000e-01 6.2000e-01 7.5000e-01 4.5000e+00 3.0426e+03 2.0400e+01]
 [1.2250e+01 1.3410e+01 1.5530e+01 2.9360e+01 3.2444e+03 3.9750e+01]
 [1.0660e+01 1.1370e+01 1.4110e+01 3.9650e+01 3.1036e+03 2.5560e+01]]


time spent until now: 237.3 mins

14:50, 04/21
ubuntu@ip-172-31-7-244:~$
```