```
Last login: Thu Apr 16 10:27:01 on ttys001
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-3-231-3-46.compute-1.amazonaws.com
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 disabled due to load higher than 96.0

 * 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

51 packages can be updated.
0 updates are security updates.


*** System restart required ***
Last login: Thu Apr 16 14:01:14 2020 from 107.13.161.147
ubuntu@ip-172-31-0-7:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
10:28, 04/16; num of cores:96
sd_u_O_20_full_T

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D_range, t_func] = [16, None, None, 20, 0.5, 1.5, [10], None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[100, 100.5, 105, 110, 110.5, 111], [15], [2, 4, 6], [[0.0003, 0.0001, 5e-05], [0.0003, 0.0001, 5e-05]]]

---------------------------------------
[pattern_seed, day, sd_R, u_O_u_D] = [2, 2, 15, 10]

max(u_O) =  145.8
O_threshold = 100
means of Order:

91.7 98.9 57.3 132.8 64.1

83.2 110.1 75.1 78.8 81.8

111.0 145.8 100.8 77.6 110.8

88.1 99.6 123.5 85.0 100.2

82.4 96.9 105.1 80.2 93.2

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 = 100.5
number of reward locations:  8
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 110.5
number of reward locations:  5
O_threshold = 111
number of reward locations:  4
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

---------------------------------------
Value of Behaviour policy:52.833
```

O_threshold = 100
MC for this TARGET:[58.167, 0.321]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.46, -1.63, -1.57]][[-1.9, -58.17, -5.33]]
std:[[1.22, 1.26, 1.26]][[0.73, 0.0, 0.53]]
MSE:[[1.9, 2.06, 2.01]][[2.04, 58.17, 5.36]]
MSE(-DR):[[0.0, 0.16, 0.11]][[0.14, 56.27, 3.46]]
***
==============
O_threshold = 100.5
MC for this TARGET:[55.653, 0.312]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.14, -0.02, -0.28]][[-1.28, -55.65, -2.82]]
std:[[1.33, 1.39, 1.03]][[0.73, 0.0, 0.53]]
MSE:[[1.34, 1.39, 1.07]][[1.47, 55.65, 2.87]]
MSE(-DR):[[0.0, 0.05, -0.27]][[0.13, 54.31, 1.53]]
**
==============
O_threshold = 105
MC for this TARGET:[57.721, 0.315]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.49, -3.59, -3.62]][[-4.67, -57.72, -4.89]]
std:[[1.87, 1.85, 1.08]][[0.71, 0.0, 0.53]]
MSE:[[3.96, 4.04, 3.78]][[4.72, 57.72, 4.92]]
MSE(-DR):[[0.0, 0.08, -0.18]][[0.76, 53.76, 0.96]]
**
==============
O_threshold = 110
MC for this TARGET:[56.709, 0.307]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.64, -2.71, -3.52]][[-5.83, -56.71, -3.88]]
std:[[1.79, 1.79, 1.17]][[0.83, 0.0, 0.53]]
MSE:[[3.19, 3.25, 3.71]][[5.89, 56.71, 3.92]]
MSE(-DR):[[0.0, 0.06, 0.52]][[2.7, 53.52, 0.73]]
***
==============
O_threshold = 110.5
MC for this TARGET:[59.443, 0.307]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.14, -6.22, -6.34]][[-10.09, -59.44, -6.61]]
std:[[1.85, 1.8, 1.42]][[0.86, 0.0, 0.53]]
MSE:[[6.41, 6.48, 6.5]][[10.13, 59.44, 6.63]]
MSE(-DR):[[0.0, 0.07, 0.09]][[3.72, 53.03, 0.22]]
***
==============
O_threshold = 111
MC for this TARGET:[57.594, 0.305]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.35, -5.39, -6.19]][[-10.44, -57.59, -4.76]]
std:[[2.04, 1.98, 1.26]][[0.8, 0.0, 0.53]]
MSE:[[5.73, 5.74, 6.32]][[10.47, 57.59, 4.79]]
MSE(-DR):[[0.0, 0.01, 0.59]][[4.74, 51.86, -0.94]]
***
==============
[[1.9   2.06 2.01 2.04 58.17 5.36]
 [1.34 1.39 1.07 1.47 55.65 2.87]
 [3.96 4.04 3.78 4.72 57.72 4.92]
 [3.19 3.25 3.71 5.89 56.71 3.92]
 [6.41 6.48 6.5 10.13 59.44 6.63]
 [5.73 5.74 6.32 10.47 57.59 4.79]]

time spent until now: 15.8 mins

10:44, 04/16

[pattern_seed, day, sd_R, u_O_u_D] = [2, 4, 15, 10]

max(u_O) = 145.8
O_threshold = 100
means of Order:

91.7 98.9 57.3 132.8 64.1

83.2 110.1 75.1 78.8 81.8

111.0 145.8 100.8 77.6 110.8

88.1 99.6 123.5 85.0 100.2

82.4 96.9 105.1 80.2 93.2

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 = 100.5
number of reward locations: 8
O_threshold = 105
number of reward locations: 7
O_threshold = 110
number of reward locations: 6
O_threshold = 110.5
number of reward locations: 5
O_threshold = 111
number of reward locations: 4
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

Value of Behaviour policy:52.766
O_threshold = 100
MC for this TARGET:[58.163, 0.249]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.91, -1.08, -1.5]][[-1.36, -58.16, -5.4]]
std:[[1.12, 1.12, 0.88]][[0.66, 0.0, 0.4]]
MSE:[[1.44, 1.56, 1.74]][[1.51, 58.16, 5.41]]
MSE(-DR):[[0.0, 0.12, 0.3]][[0.07, 56.72, 3.97]]
***
==============
O_threshold = 100.5
MC for this TARGET:[55.649, 0.243]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.56, 0.45, -0.18]][[-0.8, -55.65, -2.88]]
std:[[1.03, 1.05, 0.88]][[0.63, 0.0, 0.4]]
MSE:[[1.17, 1.14, 0.9]][[1.02, 55.65, 2.91]]
MSE(-DR):[[0.0, -0.03, -0.27]][[-0.15, 54.48, 1.74]]
==============
O_threshold = 105
MC for this TARGET:[57.717, 0.24]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.22, -3.33, -3.48]][[-4.19, -57.72, -4.95]]
std:[[1.14, 1.14, 1.03]][[0.58, 0.0, 0.4]]
MSE:[[3.42, 3.52, 3.63]][[4.23, 57.72, 4.97]]
MSE(-DR):[[0.0, 0.1, 0.21]][[0.81, 54.3, 1.55]]
***
==============
O_threshold = 110
MC for this TARGET:[56.709, 0.241]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.02, -3.11, -3.32]][[-5.3, -56.71, -3.94]]
std:[[1.19, 1.18, 1.08]][[0.57, 0.0, 0.4]]
MSE:[[3.25, 3.33, 3.49]][[5.33, 56.71, 3.96]]
MSE(-DR):[[0.0, 0.08, 0.24]][[2.08, 53.46, 0.71]]
***
==============
O_threshold = 110.5
MC for this TARGET:[59.452, 0.238]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.17, -6.19, -6.1]][[-9.57, -59.45, -6.69]]
std:[[1.28, 1.3, 1.05]][[0.58, 0.0, 0.4]]
MSE:[[6.3, 6.33, 6.19]][[9.59, 59.45, 6.7]]
MSE(-DR):[[0.0, 0.03, -0.11]][[3.29, 53.15, 0.4]]
**
==============
O_threshold = 111
MC for this TARGET:[57.589, 0.241]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.78, -5.78, -5.84]][[-9.95, -57.59, -4.82]]
std:[[1.4, 1.35, 1.2]][[0.62, 0.0, 0.4]]
MSE:[[5.95, 5.94, 5.96]][[9.97, 57.59, 4.84]]
MSE(-DR):[[0.0, -0.01, 0.01]][[4.02, 51.64, -1.11]]
***
==============
[[1.9   2.06 2.01 2.04 58.17 5.36]
 [1.34 1.39 1.07 1.47 55.65 2.87]
 [3.96 4.04 3.78 4.72 57.72 4.92]
 [3.19 3.25 3.71 5.89 56.71 3.92]
 [6.41 6.48 6.5 10.13 59.44 6.63]
 [5.73 5.74 6.32 10.47 57.59 4.79]]

[[1.44 1.56 1.74 1.51 58.16 5.41]
 [1.17 1.14 0.9   1.02 55.65 2.91]
```

```
 [ 3.42  3.52  3.63  4.23 57.72  4.97]
 [ 3.25  3.33  3.49  5.33 56.71  3.96]
 [ 6.3   6.33  6.19  9.59 59.45  6.7 ]
 [ 5.95  5.94  5.96  9.97 57.59  4.84]]
```

time spent until now: 32.6 mins

11:01, 04/16

[pattern_seed, day, sd_R, u_O_u_D] = [2, 6, 15, 10]

max(u_O) = 145.8
O_threshold = 100
means of Order:

91.7 98.9 57.3 132.8 64.1

83.2 110.1 75.1 78.8 81.8

111.0 145.8 100.8 77.6 110.8

88.1 99.6 123.5 85.0 100.2

82.4 96.9 105.1 80.2 93.2

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 = 100.5
number of reward locations: 8
O_threshold = 105
number of reward locations: 7
O_threshold = 110
number of reward locations: 6
O_threshold = 110.5
number of reward locations: 5
O_threshold = 111
number of reward locations: 4
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

Value of Behaviour policy:52.876
O_threshold = 100
MC for this TARGET:[58.156, 0.207]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.56, -0.71, -1.07]][[-1.37, -58.16, -5.28]]
std:[[0.61, 0.59, 0.62]][[0.49, 0.0, 0.33]]
MSE:[[0.83, 0.92, 1.24]][[1.45, 58.16, 5.29]]
MSE(-DR):[[0.0, 0.09, 0.41]][[0.62, 57.33, 4.46]]
***
==============
O_threshold = 100.5
MC for this TARGET:[55.643, 0.198]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.99, 0.89, 0.19]][[-0.8, -55.64, -2.77]]
std:[[0.69, 0.68, 0.63]][[0.47, 0.0, 0.33]]
MSE:[[1.21, 1.12, 0.66]][[0.93, 55.64, 2.79]]
MSE(-DR):[[0.0, -0.09, -0.55]][[-0.28, 54.43, 1.58]]
==============
O_threshold = 105
MC for this TARGET:[57.712, 0.199]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.45, -2.57, -3.02]][[-4.14, -57.71, -4.84]]
std:[[0.77, 0.79, 0.66]][[0.52, 0.0, 0.33]]
MSE:[[2.57, 2.69, 3.09]][[4.17, 57.71, 4.85]]
MSE(-DR):[[0.0, 0.12, 0.52]][[1.6, 55.14, 2.28]]
***
==============
O_threshold = 110
MC for this TARGET:[56.703, 0.195]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.25, -2.31, -2.96]][[-5.28, -56.7, -3.83]]
```

```
std:[[0.95, 0.94, 0.57]][[0.53, 0.0, 0.33]]
MSE:[[2.44, 2.49, 3.01]][[5.31, 56.7, 3.84]]
MSE(-DR):[[0.0, 0.05, 0.57]][[2.87, 54.26, 1.4]]
***
==============
O_threshold = 110.5
MC for this TARGET:[59.442, 0.195]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.63, -5.7, -5.92]][[-9.52, -59.44, -6.57]]
std:[[1.08, 1.04, 0.8]][[0.59, 0.0, 0.33]]
MSE:[[5.73, 5.79, 5.97]][[9.54, 59.44, 6.58]]
MSE(-DR):[[0.0, 0.06, 0.24]][[3.81, 53.71, 0.85]]
***
==============
O_threshold = 111
MC for this TARGET:[57.58, 0.192]
    [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.32, -5.32, -5.78]][[-9.81, -57.58, -4.7]]
std:[[1.06, 0.98, 0.92]][[0.57, 0.0, 0.33]]
MSE:[[5.42, 5.41, 5.85]][[9.83, 57.58, 4.71]]
MSE(-DR):[[0.0, -0.01, 0.43]][[4.41, 52.16, -0.71]]
***
==============
[[ 1.9   2.06  2.01  2.04 58.17  5.36]
 [ 1.34  1.39  1.07  1.47 55.65  2.87]
 [ 3.96  4.04  3.78  4.72 57.72  4.92]
 [ 3.19  3.25  3.71  5.89 56.71  3.92]
 [ 6.41  6.48  6.5  10.13 59.44  6.63]
 [ 5.73  5.74  6.32 10.47 57.59  4.79]]


[[ 1.44  1.56  1.74  1.51 58.16  5.41]
 [ 1.17  1.14  0.9   1.02 55.65  2.91]
 [ 3.42  3.52  3.63  4.23 57.72  4.97]
 [ 3.25  3.33  3.49  5.33 56.71  3.96]
 [ 6.3   6.33  6.19  9.59 59.45  6.7 ]
 [ 5.95  5.94  5.96  9.97 57.59  4.84]]


[[ 0.83  0.92  1.24  1.45 58.16  5.29]
 [ 1.21  1.12  0.66  0.93 55.64  2.79]
 [ 2.57  2.69  3.09  4.17 57.71  4.85]
 [ 2.44  2.49  3.01  5.31 56.7   3.84]
 [ 5.73  5.79  5.97  9.54 59.44  6.58]
 [ 5.42  5.41  5.85  9.83 57.58  4.71]]


time spent until now: 51.5 mins

11:20, 04/16
ubuntu@ip-172-31-0-7:~$
```