```
Last login: Thu Apr 16 15:55:00 on ttys000
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 19:55:03 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
15:56, 04/16; num of cores:96
sd_u_O_25_sd

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, 25, 0.5, 1.5, [15], N
one]

[thre_range, sd_R_range, day_range, penalty_range]:  [[100, 101, 102, 110, 113, 113.5], [10, 20], [7], [[0.0003, 0.0001
, 5e-05], [0.0003, 0.0001, 5e-05]]]

----------------------------------------
[pattern_seed, day, sd_R, u_O_u_D] = [2, 7, 10, 15]

max(u_O) =  157.3
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 = 102
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 113
```

```
number of reward locations:  5
O_threshold = 113.5
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:50.113
O_threshold = 100
MC for this TARGET:[57.56, 0.121]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.86, -2.08, -2.42]][[-3.21, -57.56, -7.45]]
std:[[0.54, 0.55, 0.4]][[0.41, 0.0, 0.21]]
MSE:[[1.94, 2.15, 2.45]][[3.24, 57.56, 7.45]]
MSE(-DR):[[0.0, 0.21, 0.51]][[1.3, 55.62, 5.51]]
***
==============
O_threshold = 101
MC for this TARGET:[54.728, 0.12]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.11, -0.29, -0.86]][[-2.34, -54.73, -4.62]]
std:[[0.54, 0.53, 0.35]][[0.36, 0.0, 0.21]]
MSE:[[0.55, 0.6, 0.93]][[2.37, 54.73, 4.62]]
MSE(-DR):[[0.0, 0.05, 0.38]][[1.82, 54.18, 4.07]]
***
==============
O_threshold = 102
MC for this TARGET:[56.441, 0.119]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.6, -3.78, -4.01]][[-5.67, -56.44, -6.33]]
std:[[0.88, 0.83, 0.49]][[0.38, 0.0, 0.21]]
MSE:[[3.71, 3.87, 4.04]][[5.68, 56.44, 6.33]]
MSE(-DR):[[0.0, 0.16, 0.33]][[1.97, 52.73, 2.62]]
***
==============
O_threshold = 110
MC for this TARGET:[55.145, 0.117]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.29, -3.41, -3.84]][[-6.5, -55.14, -5.03]]
std:[[1.02, 0.96, 0.55]][[0.37, 0.0, 0.21]]
MSE:[[3.44, 3.54, 3.88]][[6.51, 55.14, 5.03]]
MSE(-DR):[[0.0, 0.1, 0.44]][[3.07, 51.7, 1.59]]
***
==============
O_threshold = 113
MC for this TARGET:[57.769, 0.12]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.32, -6.43, -6.36]][[-11.01, -57.77, -7.66]]
std:[[0.9, 0.85, 0.49]][[0.3, 0.0, 0.21]]
MSE:[[6.38, 6.49, 6.38]][[11.01, 57.77, 7.66]]
MSE(-DR):[[0.0, 0.11, 0.0]][[4.63, 51.39, 1.28]]
***
==============
O_threshold = 113.5
MC for this TARGET:[55.747, 0.119]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.94, -6.0, -6.13]][[-11.09, -55.75, -5.63]]
std:[[0.82, 0.78, 0.56]][[0.32, 0.0, 0.21]]
MSE:[[6.0, 6.05, 6.16]][[11.09, 55.75, 5.63]]
MSE(-DR):[[0.0, 0.05, 0.16]][[5.09, 49.75, -0.37]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 1.94  2.15  2.45  3.24 57.56  7.45]
 [ 0.55  0.6   0.93  2.37 54.73  4.62]
 [ 3.71  3.87  4.04  5.68 56.44  6.33]
 [ 3.44  3.54  3.88  6.51 55.14  5.03]
 [ 6.38  6.49  6.38 11.01 57.77  7.66]
 [ 6.    6.05  6.16 11.09 55.75  5.63]]


time spent until now: 20.9 mins

16:17, 04/16
```

```
---------------------------------------
[pattern_seed, day, sd_R, u_O_u_D] = [2, 7, 20, 15]

max(u_O) =  157.3
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 = 102
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 113
number of reward locations:  5
O_threshold = 113.5
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:50.098
O_threshold = 100
MC for this TARGET:[57.575, 0.228]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.91, -2.14, -2.45]][[-3.25, -57.58, -7.48]]
std:[[0.83, 0.83, 0.56]][[0.61, 0.0, 0.25]]
MSE:[[2.08, 2.3, 2.51]][[3.31, 57.58, 7.48]]
MSE(-DR):[[0.0, 0.22, 0.43]][[1.23, 55.5, 5.4]]
***
==============
O_threshold = 101
MC for this TARGET:[54.744, 0.228]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.08, -0.28, -0.8]][[-2.4, -54.74, -4.65]]
std:[[0.87, 0.87, 0.4]][[0.52, 0.0, 0.25]]
MSE:[[0.87, 0.91, 0.89]][[2.46, 54.74, 4.66]]
MSE(-DR):[[0.0, 0.04, 0.02]][[1.59, 53.87, 3.79]]
***
==============
O_threshold = 102
MC for this TARGET:[56.457, 0.226]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.58, -3.76, -3.99]][[-5.73, -56.46, -6.36]]
std:[[1.17, 1.1, 0.63]][[0.51, 0.0, 0.25]]
MSE:[[3.77, 3.92, 4.04]][[5.75, 56.46, 6.36]]
MSE(-DR):[[0.0, 0.15, 0.27]][[1.98, 52.69, 2.59]]
***
==============
O_threshold = 110
```

```
MC for this TARGET:[55.161, 0.225]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.26, -3.41, -3.84]][[-6.56, -55.16, -5.06]]
std:[[1.49, 1.4, 0.7]][[0.5, 0.0, 0.25]]
MSE:[[3.58, 3.69, 3.9]][[6.58, 55.16, 5.07]]
MSE(-DR):[[0.0, 0.11, 0.32]][[3.0, 51.58, 1.49]]
***
==============
O_threshold = 113
MC for this TARGET:[57.785, 0.228]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.3, -6.4, -6.35]][[-11.08, -57.78, -7.69]]
std:[[1.37, 1.23, 0.8]][[0.41, 0.0, 0.25]]
MSE:[[6.45, 6.52, 6.4]][[11.09, 57.78, 7.69]]
MSE(-DR):[[0.0, 0.07, -0.05]][[4.64, 51.33, 1.24]]
**
==============
O_threshold = 113.5
MC for this TARGET:[55.762, 0.228]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.88, -5.92, -6.16]][[-11.14, -55.76, -5.66]]
std:[[1.3, 1.23, 0.86]][[0.42, 0.0, 0.25]]
MSE:[[6.02, 6.05, 6.22]][[11.15, 55.76, 5.67]]
MSE(-DR):[[0.0, 0.03, 0.2]][[5.13, 49.74, -0.35]]
***
==============
[[ 1.94  2.15  2.45  3.24 57.56  7.45]
 [ 0.55  0.6   0.93  2.37 54.73  4.62]
 [ 3.71  3.87  4.04  5.68 56.44  6.33]
 [ 3.44  3.54  3.88  6.51 55.14  5.03]
 [ 6.38  6.49  6.38 11.01 57.77  7.66]
 [ 6.    6.05  6.16 11.09 55.75  5.63]]


[[ 2.08  2.3   2.51  3.31 57.58  7.48]
 [ 0.87  0.91  0.89  2.46 54.74  4.66]
 [ 3.77  3.92  4.04  5.75 56.46  6.36]
 [ 3.58  3.69  3.9   6.58 55.16  5.07]
 [ 6.45  6.52  6.4  11.09 57.78  7.69]
 [ 6.02  6.05  6.22 11.15 55.76  5.67]]


time spent until now: 42.1 mins

16:38, 04/16
ubuntu@ip-172-31-0-7:~$
```