```
Last login: Fri Apr 10 13:37:18 on ttys002
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-35-174-61-41.compute-1.amazonaws.com
Warning: Permanently added the ED25519 host key for IP address '35.174.61.41' to the list of known hosts.
Welcome to Ubuntu 18.04.3 LTS (GNU/Linux 4.15.0-1063-aws x86_64)

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

  System information as of Fri Apr 10 22:10:03 UTC 2020

  System load:  0.86               Processes:           809
  Usage of /:   57.0% of 15.45GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.70.99
  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

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


Last login: Wed Apr  1 20:30:39 2020 from 107.13.161.147
ubuntu@ip-172-31-70-99:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
18:12, 04/10; num of cores:96
vary_T_sd_15_small

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, sd_R_range, t_func] = [96, None, None, 20, 0.5, 1, 0, [15], None]


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

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

--------------------------------------
Value of Behaviour policy:64.828
O_threshold = 100
MC for this TARGET:[70.797, 0.332]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.81, -0.87, -1.22]][[-1.61, -70.8, -5.97]]
std:[[1.59, 1.59, 0.99]][[0.83, 0.0, 0.48]]
MSE:[[1.78, 1.81, 1.57]][[1.81, 70.8, 5.99]]
MSE(-DR):[[0.0, 0.03, -0.21]][[0.03, 69.02, 4.21]]
**
==============
O_threshold = 105
MC for this TARGET:[71.795, 0.328]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.31, -3.37, -3.68]][[-4.26, -71.8, -6.97]]
std:[[1.93, 1.93, 1.03]][[0.79, 0.0, 0.48]]
MSE:[[3.83, 3.88, 3.82]][[4.33, 71.8, 6.99]]
MSE(-DR):[[0.0, 0.05, -0.01]][[0.5, 67.97, 3.16]]
**
==============
O_threshold = 110
MC for this TARGET:[70.893, 0.321]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.33, -3.38, -3.44]][[-4.99, -70.89, -6.06]]
std:[[1.86, 1.84, 0.91]][[0.82, 0.0, 0.48]]
MSE:[[3.81, 3.85, 3.56]][[5.06, 70.89, 6.08]]
MSE(-DR):[[0.0, 0.04, -0.25]][[1.25, 67.08, 2.27]]
**
==============
```

```
O_threshold = 115
MC for this TARGET:[71.81, 0.32]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[nan, -6.19, nan]][[-9.84, -71.81, -6.98]]
std:[[nan, 2.38, nan]][[0.76, 0.0, 0.48]]
MSE:[[nan, 6.63, nan]][[9.87, 71.81, 7.0]]
MSE(-DR):[[nan, nan, nan]][[nan, nan, nan]]
==============
[[ 1.78  1.81  1.57  1.81 70.8   5.99]
 [ 3.83  3.88  3.82  4.33 71.8   6.99]
 [ 3.81  3.85  3.56  5.06 70.89  6.08]
 [  nan  6.63   nan  9.87 71.81  7.  ]]


time spent until now: 55.6 mins

19:07, 04/10

--------------------------------------
[pattern_seed, day, sd_R] = [2, 3, 15]

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

--------------------------------------
Value of Behaviour policy:64.891
O_threshold = 100
MC for this TARGET:[70.785, 0.313]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.35, -0.44, -1.14]][[-1.45, -70.78, -5.89]]
std:[[1.32, 1.31, 0.9]][[0.7, 0.0, 0.42]]
MSE:[[1.37, 1.38, 1.45]][[1.61, 70.78, 5.9]]
MSE(-DR):[[0.0, 0.01, 0.08]][[0.24, 69.41, 4.53]]
***
==============
O_threshold = 105
MC for this TARGET:[71.791, 0.308]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.8, -2.87, -3.49]][[-4.15, -71.79, -6.9]]
std:[[1.47, 1.45, 0.9]][[0.68, 0.0, 0.42]]
MSE:[[3.16, 3.22, 3.6]][[4.21, 71.79, 6.91]]
MSE(-DR):[[0.0, 0.06, 0.44]][[1.05, 68.63, 3.75]]
***
==============
O_threshold = 110
MC for this TARGET:[70.894, 0.307]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.63, -2.69, -3.23]][[-4.88, -70.89, -6.0]]
std:[[1.63, 1.62, 0.98]][[0.69, 0.0, 0.42]]
MSE:[[3.09, 3.14, 3.38]][[4.93, 70.89, 6.01]]
MSE(-DR):[[0.0, 0.05, 0.29]][[1.84, 67.8, 2.92]]
***
==============
O_threshold = 115
MC for this TARGET:[71.809, 0.291]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.4, -5.4, -5.68]][[-9.7, -71.81, -6.92]]
std:[[2.08, 2.06, 1.19]][[0.7, 0.0, 0.42]]
MSE:[[5.79, 5.78, 5.8]][[9.73, 71.81, 6.93]]
MSE(-DR):[[0.0, -0.01, 0.01]][[3.94, 66.02, 1.14]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 1.78  1.81  1.57  1.81 70.8   5.99]
 [ 3.83  3.88  3.82  4.33 71.8   6.99]
 [ 3.81  3.85  3.56  5.06 70.89  6.08]
 [  nan  6.63   nan  9.87 71.81  7.  ]]


[[ 1.37  1.38  1.45  1.61 70.78  5.9 ]
 [ 3.16  3.22  3.6   4.21 71.79  6.91]
 [ 3.09  3.14  3.38  4.93 70.89  6.01]
 [ 5.79  5.78  5.8   9.73 71.81  6.93]]


time spent until now: 112.4 mins

20:04, 04/10
```

```
--------------------------------------
[pattern_seed, day, sd_R] = [2, 4, 15]

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

--------------------------------------
Value of Behaviour policy:64.89
O_threshold = 100
MC for this TARGET:[70.785, 0.255]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.52, -0.6, -1.18]][[-1.36, -70.78, -5.9]]
std:[[1.01, 0.99, 0.71]][[0.5, 0.0, 0.36]]
MSE:[[1.14, 1.16, 1.38]][[1.45, 70.78, 5.91]]
MSE(-DR):[[0.0, 0.02, 0.24]][[0.31, 69.64, 4.77]]
***
==============
O_threshold = 105
MC for this TARGET:[71.786, 0.245]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.7, -2.77, -3.52]][[-4.07, -71.79, -6.9]]
std:[[1.23, 1.23, 0.81]][[0.53, 0.0, 0.36]]
MSE:[[2.97, 3.03, 3.61]][[4.1, 71.79, 6.91]]
MSE(-DR):[[0.0, 0.06, 0.64]][[1.13, 68.82, 3.94]]
***
==============
O_threshold = 110
MC for this TARGET:[70.887, 0.248]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.54, -2.59, -3.3]][[-4.76, -70.89, -6.0]]
std:[[1.29, 1.3, 0.79]][[0.55, 0.0, 0.36]]
MSE:[[2.85, 2.9, 3.39]][[4.79, 70.89, 6.01]]
MSE(-DR):[[0.0, 0.05, 0.54]][[1.94, 68.04, 3.16]]
***
==============
O_threshold = 115
MC for this TARGET:[71.801, 0.248]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.42, -5.42, -5.81]][[-9.62, -71.8, -6.91]]
std:[[1.57, 1.55, 0.93]][[0.6, 0.0, 0.36]]
MSE:[[5.64, 5.64, 5.88]][[9.64, 71.8, 6.92]]
MSE(-DR):[[0.0, 0.0, 0.24]][[4.0, 66.16, 1.28]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 1.78  1.81  1.57  1.81 70.8   5.99]
 [ 3.83  3.88  3.82  4.33 71.8   6.99]
 [ 3.81  3.85  3.56  5.06 70.89  6.08]
 [  nan  6.63   nan  9.87 71.81  7.  ]]


[[ 1.37  1.38  1.45  1.61 70.78  5.9 ]
 [ 3.16  3.22  3.6   4.21 71.79  6.91]
 [ 3.09  3.14  3.38  4.93 70.89  6.01]
 [ 5.79  5.78  5.8   9.73 71.81  6.93]]


[[ 1.14  1.16  1.38  1.45 70.78  5.91]
 [ 2.97  3.03  3.61  4.1  71.79  6.91]
 [ 2.85  2.9   3.39  4.79 70.89  6.01]
 [ 5.64  5.64  5.88  9.64 71.8   6.92]]


time spent until now: 171.6 mins

21:03, 04/10

--------------------------------------
[pattern_seed, day, sd_R] = [2, 5, 15]

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
```

```
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

---------------------------------------
Value of Behaviour policy:64.926
O_threshold = 100
MC for this TARGET:[70.781, 0.213]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.19, -0.29, -1.0]][[-1.33, -70.78, -5.85]]
std:[[1.01, 1.0, 0.57]][[0.53, 0.0, 0.28]]
MSE:[[1.03, 1.04, 1.15]][[1.43, 70.78, 5.86]]
MSE(-DR):[[0.0, 0.01, 0.12]][[0.4, 69.75, 4.83]]
***
==============
O_threshold = 105
MC for this TARGET:[71.785, 0.207]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.4, -2.48, -3.38]][[-4.07, -71.78, -6.86]]
std:[[1.21, 1.21, 0.76]][[0.56, 0.0, 0.28]]
MSE:[[2.69, 2.76, 3.46]][[4.11, 71.78, 6.87]]
MSE(-DR):[[0.0, 0.07, 0.77]][[1.42, 69.09, 4.18]]
***
==============
O_threshold = 110
MC for this TARGET:[70.884, 0.209]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.37, -2.44, -3.15]][[-4.79, -70.88, -5.96]]
std:[[1.25, 1.27, 0.8]][[0.56, 0.0, 0.28]]
MSE:[[2.68, 2.75, 3.25]][[4.82, 70.88, 5.97]]
MSE(-DR):[[0.0, 0.07, 0.57]][[2.14, 68.2, 3.29]]
***
==============
O_threshold = 115
MC for this TARGET:[71.797, 0.204]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.44, -5.46, -5.7]][[-9.64, -71.8, -6.87]]
std:[[1.38, 1.38, 0.84]][[0.51, 0.0, 0.28]]
MSE:[[5.61, 5.63, 5.76]][[9.65, 71.8, 6.88]]
MSE(-DR):[[0.0, 0.02, 0.15]][[4.04, 66.19, 1.27]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 1.78  1.81  1.57  1.81 70.8   5.99]
 [ 3.83  3.88  3.82  4.33 71.8   6.99]
 [ 3.81  3.85  3.56  5.06 70.89  6.08]
 [  nan  6.63   nan  9.87 71.81  7.  ]]


[[ 1.37  1.38  1.45  1.61 70.78  5.9 ]
 [ 3.16  3.22  3.6   4.21 71.79  6.91]
 [ 3.09  3.14  3.38  4.93 70.89  6.01]
 [ 5.79  5.78  5.8   9.73 71.81  6.93]]


[[ 1.14  1.16  1.38  1.45 70.78  5.91]
 [ 2.97  3.03  3.61  4.1  71.79  6.91]
 [ 2.85  2.9   3.39  4.79 70.89  6.01]
 [ 5.64  5.64  5.88  9.64 71.8   6.92]]


[[ 1.03  1.04  1.15  1.43 70.78  5.86]
 [ 2.69  2.76  3.46  4.11 71.78  6.87]
 [ 2.68  2.75  3.25  4.82 70.88  5.97]
 [ 5.61  5.63  5.76  9.65 71.8   6.88]]


time spent until now: 230.6 mins

22:02, 04/10
ubuntu@ip-172-31-70-99:~$
```