```
Last login: Thu Apr  9 23:43:20 on ttys002
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-18-205-6-67.compute-1.amazonaws.com
ssh: connect to host ec2-18-205-6-67.compute-1.amazonaws.com port 22: Connection refused
Run-Mac:.ssh mac$
Run-Mac:.ssh mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-18-205-6-67.compute-1.amazonaws.com
The authenticity of host 'ec2-18-205-6-67.compute-1.amazonaws.com (18.205.6.67)' can't be established.
ECDSA key fingerprint is SHA256:cPoGmVo4ZjhGS2jvz+lYbbne8m/eP8aaOmHYk9jWEAw.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-18-205-6-67.compute-1.amazonaws.com,18.205.6.67' (ECDSA) 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 Fri Apr 10 12:16:12 UTC 2020

  System load:  1.06               Processes:           857
  Usage of /:   28.0% of 30.96GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.6.121
  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

89 packages can be updated.
39 updates are security updates.


Last login: Fri Apr  3 19:45:17 2020 from 107.13.161.147
ubuntu@ip-172-31-6-121:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
08:19, 04/10; num of cores:96

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, [5, 15, 25], None
]


--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 5]

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:64.872
O_threshold = 100
MC for this TARGET:[70.784, 0.113]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.35, -0.45, -0.94]][[-1.33, -70.78, -5.91]]
std:[[0.46, 0.47, 0.32]][[0.27, 0.0, 0.23]]
MSE:[[0.58, 0.65, 0.99]][[1.36, 70.78, 5.91]]
MSE(-DR):[[0.0, 0.07, 0.41]][[0.78, 70.2, 5.33]]
***
==============
O_threshold = 105
MC for this TARGET:[71.787, 0.107]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.68, -2.77, -3.37]][[-4.04, -71.79, -6.91]]
std:[[0.58, 0.58, 0.34]][[0.27, 0.0, 0.23]]
MSE:[[2.74, 2.83, 3.39]][[4.05, 71.79, 6.91]]
MSE(-DR):[[0.0, 0.09, 0.65]][[1.31, 69.05, 4.17]]
***
==============
O_threshold = 110
MC for this TARGET:[70.881, 0.104]
```

```
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.7, -2.77, -3.16]][[-4.75, -70.88, -6.01]]
std:[[0.59, 0.59, 0.36]][[0.27, 0.0, 0.23]]
MSE:[[2.76, 2.83, 3.18]][[4.76, 70.88, 6.01]]
MSE(-DR):[[0.0, 0.07, 0.42]][[2.0, 68.12, 3.25]]
***
==============
O_threshold = 115
MC for this TARGET:[71.805, 0.093]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.8, -5.82, -5.82]][[-9.63, -71.81, -6.93]]
std:[[0.77, 0.78, 0.45]][[0.25, 0.0, 0.23]]
MSE:[[5.85, 5.87, 5.84]][[9.63, 71.81, 6.93]]
MSE(-DR):[[0.0, 0.02, -0.01]][[3.78, 65.96, 1.08]]
**
==============
[[ 0.58  0.65  0.99  1.36 70.78  5.91]
 [ 2.74  2.83  3.39  4.05 71.79  6.91]
 [ 2.76  2.83  3.18  4.76 70.88  6.01]
 [ 5.85  5.87  5.84  9.63 71.81  6.93]]


time spent until now: 65.5 mins

09:24, 04/10

--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 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:64.885
O_threshold = 100
MC for this TARGET:[70.8, 0.2]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.33, -0.43, -0.95]][[-1.31, -70.8, -5.91]]
std:[[0.74, 0.75, 0.51]][[0.42, 0.0, 0.27]]
MSE:[[0.81, 0.86, 1.08]][[1.38, 70.8, 5.92]]
MSE(-DR):[[0.0, 0.05, 0.27]][[0.57, 69.99, 5.11]]
***
==============
O_threshold = 105
MC for this TARGET:[71.803, 0.194]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.71, -2.82, -3.41]][[-4.03, -71.8, -6.92]]
std:[[1.01, 1.0, 0.58]][[0.4, 0.0, 0.27]]
MSE:[[2.89, 2.99, 3.46]][[4.05, 71.8, 6.93]]
MSE(-DR):[[0.0, 0.1, 0.57]][[1.16, 68.91, 4.04]]
***
==============
O_threshold = 110
MC for this TARGET:[70.897, 0.193]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.78, -2.85, -3.2]][[-4.74, -70.9, -6.01]]
std:[[1.01, 1.0, 0.57]][[0.39, 0.0, 0.27]]
MSE:[[2.96, 3.02, 3.25]][[4.76, 70.9, 6.02]]
MSE(-DR):[[0.0, 0.06, 0.29]][[1.8, 67.94, 3.06]]
***
==============
O_threshold = 115
MC for this TARGET:[71.821, 0.187]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.83, -5.86, -5.87]][[-9.61, -71.82, -6.94]]
std:[[1.24, 1.23, 0.75]][[0.39, 0.0, 0.27]]
MSE:[[5.96, 5.99, 5.92]][[9.62, 71.82, 6.95]]
MSE(-DR):[[0.0, 0.03, -0.04]][[3.66, 65.86, 0.99]]
**
==============
[[ 0.58  0.65  0.99  1.36 70.78  5.91]
 [ 2.74  2.83  3.39  4.05 71.79  6.91]
 [ 2.76  2.83  3.18  4.76 70.88  6.01]
 [ 5.85  5.87  5.84  9.63 71.81  6.93]]


[[ 0.81  0.86  1.08  1.38 70.8   5.92]
 [ 2.89  2.99  3.46  4.05 71.8   6.93]
```

```
 [ 2.96  3.02  3.25  4.76 70.9   6.02]
 [ 5.96  5.99  5.92  9.62 71.82  6.95]]


time spent until now: 130.9 mins

10:30, 04/10

---------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 25]

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:64.898
O_threshold = 100
MC for this TARGET:[70.815, 0.306]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.3, -0.41, -0.95]][[-1.29, -70.82, -5.92]]
std:[[1.13, 1.13, 0.77]][[0.62, 0.0, 0.34]]
MSE:[[1.17, 1.2, 1.22]][[1.43, 70.82, 5.93]]
MSE(-DR):[[0.0, 0.03, 0.05]][[0.26, 69.65, 4.76]]
***
==============
O_threshold = 105
MC for this TARGET:[71.819, 0.301]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.76, -2.88, -3.45]][[-4.02, -71.82, -6.92]]
std:[[1.56, 1.53, 0.88]][[0.6, 0.0, 0.34]]
MSE:[[3.17, 3.26, 3.56]][[4.06, 71.82, 6.93]]
MSE(-DR):[[0.0, 0.09, 0.39]][[0.89, 68.65, 3.76]]
***
==============
O_threshold = 110
MC for this TARGET:[70.913, 0.301]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.84, -2.93, -3.21]][[-4.73, -70.91, -6.01]]
std:[[1.53, 1.52, 0.88]][[0.58, 0.0, 0.34]]
MSE:[[3.23, 3.3, 3.33]][[4.77, 70.91, 6.02]]
MSE(-DR):[[0.0, 0.07, 0.1]][[1.54, 67.68, 2.79]]
***
==============
O_threshold = 115
MC for this TARGET:[71.837, 0.297]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.89, -5.94, -5.92]][[-9.6, -71.84, -6.94]]
std:[[1.84, 1.82, 1.16]][[0.6, 0.0, 0.34]]
MSE:[[6.17, 6.21, 6.03]][[9.62, 71.84, 6.95]]
MSE(-DR):[[0.0, 0.04, -0.14]][[3.45, 65.67, 0.78]]
**
==============
[[ 0.58  0.65  0.99  1.36 70.78  5.91]
 [ 2.74  2.83  3.39  4.05 71.79  6.91]
 [ 2.76  2.83  3.18  4.76 70.88  6.01]
 [ 5.85  5.87  5.84  9.63 71.81  6.93]]


[[ 0.81  0.86  1.08  1.38 70.8   5.92]
 [ 2.89  2.99  3.46  4.05 71.8   6.93]
 [ 2.96  3.02  3.25  4.76 70.9   6.02]
 [ 5.96  5.99  5.92  9.62 71.82  6.95]]


[[ 1.17  1.2   1.22  1.43 70.82  5.93]
 [ 3.17  3.26  3.56  4.06 71.82  6.93]
 [ 3.23  3.3   3.33  4.77 70.91  6.02]
 [ 6.17  6.21  6.03  9.62 71.84  6.95]]


time spent until now: 196.3 mins

11:35, 04/10
ubuntu@ip-172-31-6-121:~$
```