```
Last login: Thu Apr  9 20:43:22 on ttys002
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-3-89-81-40.compute-1.amazonaws.com
ssh: connect to host ec2-3-89-81-40.compute-1.amazonaws.com port 22: Operation timed out
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-3-89-81-40.compute-1.amazonaws.com
The authenticity of host 'ec2-3-89-81-40.compute-1.amazonaws.com (3.89.81.40)' can't be established.
ECDSA key fingerprint is SHA256:ox9SGn/DUcebG28IK/5COhC9Q4C8xnlVtUMvEA8T7GE.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-3-89-81-40.compute-1.amazonaws.com,3.89.81.40' (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 00:52:40 UTC 2020

  System load:  2.22               Processes:              962
  Usage of /:   28.9% of 30.96GB   Users logged in:        0
  Memory usage: 0%                 IP address for enp125s0: 172.31.12.239
  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

78 packages can be updated.
27 updates are security updates.


*** System restart required ***
Last login: Fri Apr  3 19:45:17 2020 from 107.13.161.147
ubuntu@ip-172-31-12-239:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
20:54, 04/09; 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, [30], None]


---------------------------------------
[pattern_seed, day, sd_R] = [2, 4, 30]

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:64.884
O_threshold = 100
MC for this TARGET:[70.787, 0.479]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.62, -0.69, -1.29]][[-1.32, -70.79, -5.9]]
std:[[1.87, 1.82, 1.24]][[0.84, 0.0, 0.54]]
MSE:[[1.97, 1.95, 1.79]][[1.56, 70.79, 5.92]]
MSE(-DR):[[0.0, -0.02, -0.18]][[-0.41, 68.82, 3.95]]
==============
O_threshold = 105
MC for this TARGET:[71.788, 0.471]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.7, -2.77, -3.61]][[-4.07, -71.79, -6.9]]
std:[[2.18, 2.18, 1.42]][[0.89, 0.0, 0.54]]
MSE:[[3.47, 3.52, 3.88]][[4.17, 71.79, 6.92]]
MSE(-DR):[[0.0, 0.05, 0.41]][[0.7, 68.32, 3.45]]
***
==============
O_threshold = 110
MC for this TARGET:[70.889, 0.473]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.46, -2.51, -3.37]][[-4.71, -70.89, -6.0]]
std:[[2.28, 2.29, 1.4]][[0.92, 0.0, 0.54]]
```

```
MSE:[[3.35, 3.4, 3.65]][[4.8, 70.89, 6.02]]
MSE(-DR):[[0.0, 0.05, 0.3]][[1.45, 67.54, 2.67]]
***
==============
O_threshold = 115
MC for this TARGET:[71.803, 0.476]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.49, -5.48, -5.85]][[-9.56, -71.8, -6.92]]
std:[[2.8, 2.78, 1.58]][[1.01, 0.0, 0.54]]
MSE:[[6.16, 6.14, 6.06]][[9.61, 71.8, 6.94]]
MSE(-DR):[[0.0, -0.02, -0.1]][[3.45, 65.64, 0.78]]
**
==============
[[ 1.97  1.95  1.79  1.56 70.79  5.92]
 [ 3.47  3.52  3.88  4.17 71.79  6.92]
 [ 3.35  3.4   3.65  4.8  70.89  6.02]
 [ 6.16  6.14  6.06  9.61 71.8   6.94]]


time spent until now: 50.5 mins

21:45, 04/09

---------------------------------------
[pattern_seed, day, sd_R] = [2, 8, 30]

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:64.896
O_threshold = 100
MC for this TARGET:[70.783, 0.361]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.32, -0.41, -0.97]][[-1.39, -70.78, -5.89]]
std:[[1.4, 1.38, 0.84]][[0.7, 0.0, 0.39]]
MSE:[[1.44, 1.44, 1.28]][[1.56, 70.78, 5.9]]
MSE(-DR):[[0.0, 0.0, -0.16]][[0.12, 69.34, 4.46]]
**
==============
O_threshold = 105
MC for this TARGET:[71.783, 0.355]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.79, -2.87, -3.35]][[-4.08, -71.78, -6.89]]
std:[[1.68, 1.67, 0.95]][[0.7, 0.0, 0.39]]
MSE:[[3.26, 3.32, 3.48]][[4.14, 71.78, 6.9]]
MSE(-DR):[[0.0, 0.06, 0.22]][[0.88, 68.52, 3.64]]
***
==============
O_threshold = 110
MC for this TARGET:[70.877, 0.355]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.83, -2.89, -3.14]][[-4.77, -70.88, -5.98]]
std:[[1.82, 1.81, 0.96]][[0.72, 0.0, 0.39]]
MSE:[[3.36, 3.41, 3.28]][[4.82, 70.88, 5.99]]
MSE(-DR):[[0.0, 0.05, -0.08]][[1.46, 67.52, 2.63]]
**
==============
O_threshold = 115
MC for this TARGET:[71.804, 0.355]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.12, -6.13, -5.81]][[-9.72, -71.8, -6.91]]
std:[[2.3, 2.28, 1.14]][[0.67, 0.0, 0.39]]
MSE:[[6.54, 6.54, 5.92]][[9.74, 71.8, 6.92]]
MSE(-DR):[[0.0, 0.0, -0.62]][[3.2, 65.26, 0.38]]
**
==============
[[ 1.97  1.95  1.79  1.56 70.79  5.92]
 [ 3.47  3.52  3.88  4.17 71.79  6.92]
 [ 3.35  3.4   3.65  4.8  70.89  6.02]
 [ 6.16  6.14  6.06  9.61 71.8   6.94]]


[[ 1.44  1.44  1.28  1.56 70.78  5.9 ]
 [ 3.26  3.32  3.48  4.14 71.78  6.9 ]
 [ 3.36  3.41  3.28  4.82 70.88  5.99]
 [ 6.54  6.54  5.92  9.74 71.8   6.92]]
```

time spent until now: 110.5 mins

22:45, 04/09

[pattern_seed, day, sd_R] = [2, 12, 30]

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