```
Last login: Thu Apr  2 12:01:06 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-3-80-231-120.compute-1.amazonaws.com
The authenticity of host 'ec2-3-80-231-120.compute-1.amazonaws.com (3.80.231.120)' can't be established.
ECDSA key fingerprint is SHA256:HiXSVfBJrClFi2fn5hAZbj6LIbvhLVvqsytNRn5uGZ4.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-3-80-231-120.compute-1.amazonaws.com,3.80.231.120' (ECDSA) 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 Thu Apr  2 18:14:47 UTC 2020

  System load:  0.85               Processes:           822
  Usage of /:   56.9% of 15.45GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.77.13
  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-172-31-77-13:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:16, 04/02; num of cores:96

Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, [M_in_R, mean_reversion, pois0, simple, u_O_u_D]] = [None, 96, 10, 10,
None, 0.3, 0.5, 1, [True, False, True, False, 10]]


--------------------------------------
[pattern_seed, day, sd_R] = [2, 5, 10]

max(u_O) =  197.9
O_threshold = 80
means of Order:

87.8 97.8 52.4 162.7 58.1

77.3 115.7 68.5 72.4 75.7

117.4 197.9 100.7 71.1 116.9

83.2 98.9 141.5 79.5 99.8

76.4 94.9 107.4 73.9 89.9

target policy:

1 1 0 1 0

0 1 0 0 0

1 1 1 0 1

1 1 1 0 1

0 1 1 0 1

number of reward locations:  15
O_threshold = 90
target policy:

0 1 0 1 0

0 1 0 0 0

1 1 1 0 1

0 1 1 0 1

0 1 1 0 0
```

```
number of reward locations:  12
O_threshold = 100
target policy:

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 0

0 0 1 0 0

number of reward locations:  8
O_threshold = 110
target policy:

0 0 0 1 0

0 1 0 0 0

1 1 0 0 1

0 0 1 0 0

0 0 0 0 0

number of reward locations:  6
1 -th region DONE!
6 -th region DONE!
11 -th region DONE!
16 -th region DONE!
21 -th region DONE!
1 -th region DONE!
6 -th region DONE!
11 -th region DONE!
16 -th region DONE!
21 -th region DONE!
1 -th region DONE!
6 -th region DONE!
11 -th region DONE!
16 -th region DONE!
21 -th region DONE!
1 -th region DONE!
6 -th region DONE!
11 -th region DONE!
16 -th region DONE!
21 -th region DONE!

---------------------------------------
Value of Behaviour policy:60.815
O_threshold = 80
MC for this TARGET:[70.883, 0.132]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.0, -0.14, -0.96]][[1.33, -2320.15, -10.07]]
std:[[0.86, 0.86, 0.49]][[0.39, 185297.99, 0.25]]
MSE:[[0.86, 0.87, 1.08]][[1.39, 185312.51, 10.07]]
MSE(-DR):[[0.0, 0.01, 0.22]][[0.53, 185311.65, 9.21]]
***
==============


O_threshold = 90
MC for this TARGET:[69.367, 0.133]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.07, -0.2, -1.01]][[-0.43, 8129.07, -8.55]]
std:[[0.74, 0.74, 0.42]][[0.37, 285813.07, 0.25]]
MSE:[[0.74, 0.77, 1.09]][[0.57, 285928.65, 8.55]]
MSE(-DR):[[0.0, 0.03, 0.35]][[-0.17, 285927.91, 7.81]]
==============


O_threshold = 100
MC for this TARGET:[68.93, 0.133]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.56, -2.66, -3.32]][[-4.73, 103361.73, -8.12]]
std:[[0.73, 0.73, 0.42]][[0.39, 1207533.13, 0.25]]
MSE:[[2.66, 2.76, 3.35]][[4.75, 1211948.81, 8.12]]
MSE(-DR):[[0.0, 0.1, 0.69]][[2.09, 1211946.15, 5.46]]
***
==============


O_threshold = 110
MC for this TARGET:[70.482, 0.138]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.31, -6.39, -6.81]][[-8.71, 42422.63, -9.67]]
std:[[0.92, 0.92, 0.46]][[0.4, 323850.08, 0.25]]
```

MSE:[[6.38, 6.46, 6.83]][[8.72, 326616.83, 9.67]]
MSE(-DR):[[0.0, 0.08, 0.45]][[2.34, 326610.45, 3.29]]
***
==========

[[8.6000e-01 8.7000e-01 1.0800e+00 1.3900e+00 1.8531e+05 1.0070e+01]
 [7.4000e-01 7.7000e-01 1.0900e+00 5.7000e-01 2.8593e+05 8.5500e+00]
 [2.6600e+00 2.7600e+00 3.3500e+00 4.7500e+00 1.2119e+06 8.1200e+00]
 [6.3800e+00 6.4600e+00 6.8300e+00 8.7200e+00 3.2662e+05 9.6700e+00]]

time spent until now: 115.0 mins

[pattern_seed, day, sd_R] = [2, 12, 10]

max(u_O) = 197.9
O_threshold = 80
means of Order:

87.8 97.8 52.4 162.7 58.1

77.3 115.7 68.5 72.4 75.7

117.4 197.9 100.7 71.1 116.9

83.2 98.9 141.5 79.5 99.8

76.4 94.9 107.4 73.9 89.9

target policy:

1 1 0 1 0

0 1 0 0 0

1 1 1 0 1

1 1 1 0 1

0 1 1 0 1

number of reward locations: 15
O_threshold = 90
target policy:

0 1 0 1 0

0 1 0 0 0

1 1 1 0 1

0 1 1 0 1

0 1 1 0 0

number of reward locations: 12
O_threshold = 100
target policy:

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 0

0 0 1 0 0

number of reward locations: 8
O_threshold = 110
target policy:

0 0 0 1 0

0 1 0 0 0

1 1 0 0 1

0 0 1 0 0

0 0 0 0 0

number of reward locations: 6