```
Last login: Sun Apr  5 11:14:35 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-3-235-75-45.compute-1.amazonaws.com
The authenticity of host 'ec2-3-235-75-45.compute-1.amazonaws.com (3.235.75.45)' can't be established.
ECDSA key fingerprint is SHA256:+UKHub6R3DETk7/34C/QknCPoaU7SuMvJOWXWVDYwNQ.
Are you sure you want to continue connecting (yes/no)? ^[[A^[[A^[[B
Please type 'yes' or 'no': yes
Warning: Permanently added 'ec2-3-235-75-45.compute-1.amazonaws.com,3.235.75.45' (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 Mon Apr  6 16:20:10 UTC 2020

  System load:  0.65               Processes:           227
  Usage of /:   57.0% of 15.45GB   Users logged in:     0
  Memory usage: 1%                 IP address for ens5: 172.31.74.33
  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-172-31-74-33:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
12:21, 04/06; num of cores:16

final sd_R trend for [0, 5, 10]

Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, [M_in_R, mean_reversion, poisO, u_O_u_D], sd_R_range] = [None, 16, None
, None, None, 20, 0.5, 1, [True, False, True, 10], [0, 5]]


---------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 0]

max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
target 1 DONE!
target 2 DONE!
target 3 DONE!
target 4 DONE!

---------------------------------------
Value of Behaviour policy:61.039
O_threshold = 95
MC for this TARGET:[67.357, 0.07]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.19, 1.04, 0.25]][[0.86, -21797.77, -6.32]]
std:[[0.43, 0.42, 0.32]][[0.26, 132365.76, 0.19]]
MSE:[[1.27, 1.12, 0.41]][[0.9, 134148.56, 6.32]]
MSE(-DR):[[0.0, -0.15, -0.86]][[-0.37, 134147.29, 5.05]]
==============
O_threshold = 100
MC for this TARGET:[67.473, 0.073]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.49, -1.58, -1.98]][[-2.76, 59338.25, -6.43]]
std:[[0.52, 0.52, 0.34]][[0.31, 2202183.29, 0.19]]
MSE:[[1.58, 1.66, 2.01]][[2.78, 2202982.59, 6.43]]
MSE(-DR):[[0.0, 0.08, 0.43]][[1.2, 2202981.01, 4.85]]
***
==============
O_threshold = 105
MC for this TARGET:[67.91, 0.075]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.9, -3.95, -4.38]][[-5.6, 19550.18, -6.87]]
```

```
std:[[0.48, 0.47, 0.35]][[0.33, 50181.03, 0.19]]
MSE:[[3.93, 3.98, 4.39]][[5.61, 53854.85, 6.87]]
MSE(-DR):[[0.0, 0.05, 0.46]][[1.68, 53850.92, 2.94]]
***
==============
O_threshold = 110
MC for this TARGET:[66.847, 0.073]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.76, -3.81, -4.17]][[-6.22, -1126.04, -5.81]]
std:[[0.4, 0.38, 0.4]][[0.32, 72143.38, 0.19]]
MSE:[[3.78, 3.83, 4.19]][[6.23, 72152.17, 5.81]]
MSE(-DR):[[0.0, 0.05, 0.41]][[2.45, 72148.39, 2.03]]
***
==============
[[1.2700e+00 1.1200e+00 4.1000e-01 9.0000e-01 1.3415e+05 6.3200e+00]
 [1.5800e+00 1.6600e+00 2.0100e+00 2.7800e+00 2.2030e+06 6.4300e+00]
 [3.9300e+00 3.9800e+00 4.3900e+00 5.6100e+00 5.3855e+04 6.8700e+00]
 [3.7800e+00 3.8300e+00 4.1900e+00 6.2300e+00 7.2152e+04 5.8100e+00]]


time spent until now: 63.8 mins


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

max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
target 1 DONE!
target 2 DONE!
target 3 DONE!
target 4 DONE!

---------------------------------------
Value of Behaviour policy:61.032
O_threshold = 95
MC for this TARGET:[67.365, 0.092]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.26, 1.11, 0.29]][[0.88, -53144.57, -6.33]]
std:[[0.4, 0.4, 0.25]][[0.25, 246924.93, 0.16]]
MSE:[[1.32, 1.18, 0.38]][[0.91, 252579.23, 6.33]]
MSE(-DR):[[0.0, -0.14, -0.94]][[-0.41, 252577.91, 5.01]]
==============
O_threshold = 100
MC for this TARGET:[67.481, 0.097]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.47, -1.56, -2.0]][[-2.77, 56444.33, -6.45]]
std:[[0.49, 0.51, 0.35]][[0.27, 1561683.32, 0.16]]
MSE:[[1.55, 1.64, 2.03]][[2.78, 1562703.03, 6.45]]
MSE(-DR):[[0.0, 0.09, 0.48]][[1.23, 1562701.48, 4.9]]
***
==============
O_threshold = 105
MC for this TARGET:[67.917, 0.095]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.78, -3.86, -4.41]][[-5.64, 18349.76, -6.88]]
std:[[0.52, 0.51, 0.38]][[0.28, 49441.91, 0.16]]
MSE:[[3.82, 3.89, 4.43]][[5.65, 52737.24, 6.88]]
MSE(-DR):[[0.0, 0.07, 0.61]][[1.83, 52733.42, 3.06]]
***
==============
O_threshold = 110
MC for this TARGET:[66.854, 0.094]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.7, -3.74, -4.18]][[-6.24, -7410.1, -5.82]]
std:[[0.42, 0.42, 0.39]][[0.28, 87191.36, 0.16]]
MSE:[[3.72, 3.76, 4.2]][[6.25, 87505.67, 5.82]]
MSE(-DR):[[0.0, 0.04, 0.48]][[2.53, 87501.95, 2.1]]
***
==============
[[1.2700e+00 1.1200e+00 4.1000e-01 9.0000e-01 1.3415e+05 6.3200e+00]
 [1.5800e+00 1.6600e+00 2.0100e+00 2.7800e+00 2.2030e+06 6.4300e+00]
 [3.9300e+00 3.9800e+00 4.3900e+00 5.6100e+00 5.3855e+04 6.8700e+00]
 [3.7800e+00 3.8300e+00 4.1900e+00 6.2300e+00 7.2152e+04 5.8100e+00]]


[[1.3200e+00 1.1800e+00 3.8000e-01 9.1000e-01 2.5258e+05 6.3300e+00]
 [1.5500e+00 1.6400e+00 2.0300e+00 2.7800e+00 1.5627e+06 6.4500e+00]
 [3.8200e+00 3.8900e+00 4.4300e+00 5.6500e+00 5.2737e+04 6.8800e+00]
 [3.7200e+00 3.7600e+00 4.2000e+00 6.2500e+00 8.7506e+04 5.8200e+00]]
```

```
time spent until now: 127.5 mins

ubuntu@ip-172-31-74-33:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:39, 04/06; num of cores:16

final sd_R trend for[0, 5]

Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, [M_in_R, mean_reversion, pois0, u_O_u_D], sd_R_range, t_func] = [None,
16, None, None, None, 20, 0.5, 1, [True, False, True, 10], [0, 5], None]


----------------------------------------
[pattern_seed, day, sd_R] = [0, 7, 0]

Traceback (most recent call last):
  File "EC2.py", line 82, in <module>
    with_MF = with_MF,
TypeError: simu() got an unexpected keyword argument 'with_MF'
ubuntu@ip-172-31-74-33:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
14:46, 04/06; num of cores:16

final sd_R trend for[0, 5]

Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A, [M_in_R, mean_reversion, pois0, u_O_u_D], sd_R_range, t_func] = [None,
16, None, None, None, 20, 0.5, 1, [True, False, True, 10], [0, 5], None]


----------------------------------------
[pattern_seed, day, sd_R] = [0, 7, 0]

max(u_O) =  145.4
O_threshold = 100
number of reward locations:  18
O_threshold = 105
number of reward locations:  16
O_threshold = 110
number of reward locations:  11
O_threshold = 115
number of reward locations:  10
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

----------------------------------------
Value of Behaviour policy:69.864
O_threshold = 100
MC for this TARGET:[81.342, 0.087]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.24, 0.04, -1.67]][[2.3, -81.34, -11.48]]
std:[[0.48, 0.48, 0.35]][[0.29, 0.0, 0.18]]
MSE:[[0.54, 0.48, 1.71]][[2.32, 81.34, 11.48]]
MSE(-DR):[[0.0, -0.06, 1.17]][[1.78, 80.8, 10.94]]
***
==============
O_threshold = 105
MC for this TARGET:[81.319, 0.078]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.75, -0.96, -3.09]][[0.51, -81.32, -11.45]]
std:[[0.36, 0.34, 0.34]][[0.32, 0.0, 0.18]]
MSE:[[0.83, 1.02, 3.11]][[0.6, 81.32, 11.45]]
MSE(-DR):[[0.0, 0.19, 2.28]][[-0.23, 80.49, 10.62]]
==============
O_threshold = 110
MC for this TARGET:[76.831, 0.07]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.95, -1.05, -1.49]][[-1.61, -76.83, -6.97]]
std:[[0.41, 0.41, 0.33]][[0.33, 0.0, 0.18]]
MSE:[[1.03, 1.13, 1.53]][[1.64, 76.83, 6.97]]
MSE(-DR):[[0.0, 0.1, 0.5]][[0.61, 75.8, 5.94]]
***
==============
O_threshold = 115
MC for this TARGET:[75.989, 0.065]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.9, -1.97, -2.08]][[-2.5, -75.99, -6.12]]
std:[[0.42, 0.42, 0.36]][[0.33, 0.0, 0.18]]
MSE:[[1.95, 2.01, 2.11]][[2.52, 75.99, 6.12]]
MSE(-DR):[[0.0, 0.06, 0.16]][[0.57, 74.04, 4.17]]
***
==============
[[ 0.54  0.48  1.71  2.32 81.34 11.48]
 [ 0.83  1.02  3.11  0.6  81.32 11.45]
 [ 1.03  1.13  1.53  1.64 76.83  6.97]
 [ 1.95  2.01  2.11  2.52 75.99  6.12]]


time spent until now: 40.3 mins
```

```
--------------------------------------
[pattern_seed, day, sd_R] = [0, 7, 5]

max(u_O) =  145.4
O_threshold = 100
number of reward locations:  18
O_threshold = 105
number of reward locations:  16
O_threshold = 110
number of reward locations:  11
O_threshold = 115
number of reward locations:  10
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

--------------------------------------
Value of Behaviour policy:69.857
O_threshold = 100
MC for this TARGET:[81.35, 0.107]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.28, 0.08, -1.67]][[2.3, -81.35, -11.49]]
std:[[0.52, 0.52, 0.35]][[0.32, 0.0, 0.17]]
MSE:[[0.59, 0.53, 1.71]][[2.32, 81.35, 11.49]]
MSE(-DR):[[0.0, -0.06, 1.12]][[1.73, 80.76, 10.9]]
***
==============
O_threshold = 105
MC for this TARGET:[81.327, 0.102]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.75, -0.97, -3.1]][[0.51, -81.33, -11.47]]
std:[[0.42, 0.41, 0.35]][[0.36, 0.0, 0.17]]
MSE:[[0.86, 1.05, 3.12]][[0.62, 81.33, 11.47]]
MSE(-DR):[[0.0, 0.19, 2.26]][[-0.24, 80.47, 10.61]]
==============
O_threshold = 110
MC for this TARGET:[76.838, 0.096]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.04, -1.15, -1.53]][[-1.65, -76.84, -6.98]]
std:[[0.45, 0.45, 0.34]][[0.32, 0.0, 0.17]]
MSE:[[1.13, 1.23, 1.57]][[1.68, 76.84, 6.98]]
MSE(-DR):[[0.0, 0.1, 0.44]][[0.55, 75.71, 5.85]]
***
==============
O_threshold = 115
MC for this TARGET:[75.997, 0.095]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.97, -2.03, -2.15]][[-2.54, -76.0, -6.14]]
std:[[0.5, 0.51, 0.37]][[0.34, 0.0, 0.17]]
MSE:[[2.03, 2.09, 2.18]][[2.56, 76.0, 6.14]]
MSE(-DR):[[0.0, 0.06, 0.15]][[0.53, 73.97, 4.11]]
***
==============
[[ 0.54  0.48  1.71  2.32 81.34 11.48]
 [ 0.83  1.02  3.11  0.6  81.32 11.45]
 [ 1.03  1.13  1.53  1.64 76.83  6.97]
 [ 1.95  2.01  2.11  2.52 75.99  6.12]]


[[ 0.59  0.53  1.71  2.32 81.35 11.49]
 [ 0.86  1.05  3.12  0.62 81.33 11.47]
 [ 1.13  1.23  1.57  1.68 76.84  6.98]
 [ 2.03  2.09  2.18  2.56 76.    6.14]]


time spent until now: 79.8 mins

ubuntu@ip-172-31-74-33:~$
```