```
Last login: Tue Apr  7 10:27:24 on ttys001
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-18-207-240-82.compute-1.amazonaws.com
The authenticity of host 'ec2-18-207-240-82.compute-1.amazonaws.com (18.207.240.82)' can't be established.
ECDSA key fingerprint is SHA256:YHSFC6Jd6hM4VEZmZbEOz+Oy/aT0IfULCxsCJl84eBs.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-18-207-240-82.compute-1.amazonaws.com,18.207.240.82' (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 Tue Apr  7 14:33:40 UTC 2020

  System load:  0.68               Processes:           796
  Usage of /:   57.0% of 15.45GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.64.145
  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
export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.pyubuntu@ip-172-31-64-145:~$ export openblas_num_threads=1; export O
MP_NUM_THREADS=1; python EC2.py
10:35, 04/07; num of cores:96

final sd_R trend for[10, 15] the same

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, t_func] = [None,
96, None, None, None, 30, 0.5, 1, [True, False, True, 10], [10, 15], None]


--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 10]

max(u_O) =  168.8
O_threshold = 80
number of reward locations:  15
O_threshold = 85
number of reward locations:  14
O_threshold = 90
number of reward locations:  12
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
O_threshold = 115
number of reward locations:  6
O_threshold = 120
number of reward locations:  3
^CProcess Process-84:
Process Process-63:
Process Process-73:
Traceback (most recent call last):
Process Process-32:
  File "EC2.py", line 86, in <module>
Process Process-65:
Process Process-31:
Process Process-74:
Process Process-33:
Process Process-95:
Process Process-82:
Process Process-10:
Process Process-55:
    with_MF = with_MF,
  File "/home/ubuntu/simu_funs.py", line 62, in simu
    value_reps = parmap(once, range(OPE_rep_times), n_cores)
  File "/home/ubuntu/_uti_basic.py", line 80, in parmap
```

```
  File "/home/ubuntu/simu_funs.py", line 212, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 149, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 149, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 108, in getOneRegionValue
    spatial = False)
  File "/home/ubuntu/main.py", line 253, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1142, in _run
    np_val = np.asarray(subfeed_val, dtype=subfeed_dtype)
  File "/home/ubuntu/anaconda3/lib/python3.7/site-packages/numpy/core/numeric.py", line 538, in asarray
    return array(a, dtype, copy=False, order=order)
KeyboardInterrupt
Traceback (most recent call last):
ubuntu@ip-172-31-64-145:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
10:36, 04/07; num of cores:96

final sd_R trend for[10, 15] the same

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,
96, None, None, None, 30, 0.5, 1, [True, False, True, 10], [10, 15], None]


---------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 10]

max(u_O) =  168.8
O_threshold = 75
number of reward locations:  16
O_threshold = 80
number of reward locations:  15
O_threshold = 85
number of reward locations:  14
O_threshold = 90
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
O_threshold = 120
number of reward locations:  3
target 1 in 8 DONE!
target 2 in 8 DONE!
target 3 in 8 DONE!
target 4 in 8 DONE!
target 5 in 8 DONE!
target 6 in 8 DONE!
target 7 in 8 DONE!
target 8 in 8 DONE!

---------------------------------------
Value of Behaviour policy:57.737
O_threshold = 75
MC for this TARGET:[66.871, 0.127]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.55, 0.35, -0.41]][[2.48, -66.87, -9.13]]
std:[[0.78, 0.79, 0.41]][[0.35, 0.0, 0.22]]
MSE:[[0.95, 0.86, 0.58]][[2.5, 66.87, 9.13]]
MSE(-DR):[[0.0, -0.09, -0.37]][[1.55, 65.92, 8.18]]
**
==============
O_threshold = 80
MC for this TARGET:[68.371, 0.129]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.05, -0.28, -1.04]][[1.16, -68.37, -10.63]]
std:[[0.73, 0.73, 0.41]][[0.36, 0.0, 0.22]]
MSE:[[0.73, 0.78, 1.12]][[1.21, 68.37, 10.63]]
MSE(-DR):[[0.0, 0.05, 0.39]][[0.48, 67.64, 9.9]]
***
==============
O_threshold = 85
MC for this TARGET:[68.386, 0.134]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.01, -0.22, -1.3]][[0.1, -68.39, -10.65]]
std:[[0.55, 0.54, 0.38]][[0.37, 0.0, 0.22]]
MSE:[[0.55, 0.58, 1.35]][[0.38, 68.39, 10.65]]
MSE(-DR):[[0.0, 0.03, 0.8]][[-0.17, 67.84, 10.1]]
==============
O_threshold = 90
MC for this TARGET:[66.73, 0.139]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
```

```
bias:[[0.01, -0.17, -0.76]][[-0.45, -66.73, -8.99]]
std:[[0.52, 0.52, 0.38]][[0.35, 0.0, 0.22]]
MSE:[[0.52, 0.55, 0.85]][[0.57, 66.73, 8.99]]
MSE(-DR):[[0.0, 0.03, 0.33]][[0.05, 66.21, 8.47]]
***
==============
O_threshold = 100
MC for this TARGET:[66.964, 0.143]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.78, -2.91, -3.16]][[-4.17, -66.96, -9.23]]
std:[[0.61, 0.61, 0.36]][[0.35, 0.0, 0.22]]
MSE:[[2.85, 2.97, 3.18]][[4.18, 66.96, 9.23]]
MSE(-DR):[[0.0, 0.12, 0.33]][[1.33, 64.11, 6.38]]
***
==============
O_threshold = 105
MC for this TARGET:[67.31, 0.146]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.37, -5.47, -5.71]][[-7.02, -67.31, -9.57]]
std:[[0.7, 0.7, 0.47]][[0.36, 0.0, 0.22]]
MSE:[[5.42, 5.51, 5.73]][[7.03, 67.31, 9.57]]
MSE(-DR):[[0.0, 0.09, 0.31]][[1.61, 61.89, 4.15]]
***
==============
O_threshold = 110
MC for this TARGET:[65.984, 0.143]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.12, -5.2, -5.41]][[-7.36, -65.98, -8.25]]
std:[[0.75, 0.75, 0.47]][[0.37, 0.0, 0.22]]
MSE:[[5.17, 5.25, 5.43]][[7.37, 65.98, 8.25]]
MSE(-DR):[[0.0, 0.08, 0.26]][[2.2, 60.81, 3.08]]
***
==============
O_threshold = 120
MC for this TARGET:[65.352, 0.129]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.08, -7.08, -6.94]][[-11.44, -65.35, -7.62]]
std:[[0.99, 0.99, 0.57]][[0.35, 0.0, 0.22]]
MSE:[[7.15, 7.15, 6.96]][[11.45, 65.35, 7.62]]
MSE(-DR):[[0.0, 0.0, -0.19]][[4.3, 58.2, 0.47]]
**
==============
[[ 0.95  0.86  0.58  2.5  66.87  9.13]
 [ 0.73  0.78  1.12  1.21 68.37 10.63]
 [ 0.55  0.58  1.35  0.38 68.39 10.65]
 [ 0.52  0.55  0.85  0.57 66.73  8.99]
 [ 2.85  2.97  3.18  4.18 66.96  9.23]
 [ 5.42  5.51  5.73  7.03 67.31  9.57]
 [ 5.17  5.25  5.43  7.37 65.98  8.25]
 [ 7.15  7.15  6.96 11.45 65.35  7.62]]


time spent until now: 128.1 mins


--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 15]

max(u_O) =  168.8
O_threshold = 75
number of reward locations:  16
O_threshold = 80
number of reward locations:  15
O_threshold = 85
number of reward locations:  14
O_threshold = 90
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
O_threshold = 120
number of reward locations:  3
target 1 in 8 DONE!
target 2 in 8 DONE!
target 3 in 8 DONE!
target 4 in 8 DONE!
target 5 in 8 DONE!
target 6 in 8 DONE!
target 7 in 8 DONE!
target 8 in 8 DONE!

--------------------------------------
Value of Behaviour policy:57.743
O_threshold = 75
MC for this TARGET:[66.879, 0.179]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
```

```
bias:[[0.57, 0.37, -0.39]][[2.48, -66.88, -9.14]]
std:[[0.99, 1.0, 0.57]][[0.45, 0.0, 0.25]]
MSE:[[1.14, 1.07, 0.69]][[2.52, 66.88, 9.14]]
MSE(-DR):[[0.0, -0.07, -0.45]][[1.38, 65.74, 8.0]]
**
==============
O_threshold = 80
MC for this TARGET:[68.379, 0.181]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.05, -0.27, -1.01]][[1.15, -68.38, -10.64]]
std:[[0.94, 0.93, 0.54]][[0.46, 0.0, 0.25]]
MSE:[[0.94, 0.97, 1.15]][[1.24, 68.38, 10.64]]
MSE(-DR):[[0.0, 0.03, 0.21]][[0.3, 67.44, 9.7]]
***
==============
O_threshold = 85
MC for this TARGET:[68.394, 0.187]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.01, -0.22, -1.3]][[0.1, -68.39, -10.65]]
std:[[0.75, 0.73, 0.54]][[0.46, 0.0, 0.25]]
MSE:[[0.75, 0.76, 1.41]][[0.47, 68.39, 10.65]]
MSE(-DR):[[0.0, 0.01, 0.66]][[-0.28, 67.64, 9.9]]
==============
O_threshold = 90
MC for this TARGET:[66.737, 0.191]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.02, -0.17, -0.75]][[-0.46, -66.74, -8.99]]
std:[[0.7, 0.69, 0.53]][[0.45, 0.0, 0.25]]
MSE:[[0.7, 0.71, 0.92]][[0.64, 66.74, 8.99]]
MSE(-DR):[[0.0, 0.01, 0.22]][[-0.06, 66.04, 8.29]]
==============
O_threshold = 100
MC for this TARGET:[66.972, 0.196]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.79, -2.92, -3.15]][[-4.18, -66.97, -9.23]]
std:[[0.82, 0.81, 0.48]][[0.45, 0.0, 0.25]]
MSE:[[2.91, 3.03, 3.19]][[4.2, 66.97, 9.23]]
MSE(-DR):[[0.0, 0.12, 0.28]][[1.29, 64.06, 6.32]]
***
==============
O_threshold = 105
MC for this TARGET:[67.318, 0.198]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.41, -5.51, -5.73]][[-7.03, -67.32, -9.57]]
std:[[0.95, 0.93, 0.62]][[0.45, 0.0, 0.25]]
MSE:[[5.49, 5.59, 5.76]][[7.04, 67.32, 9.57]]
MSE(-DR):[[0.0, 0.1, 0.27]][[1.55, 61.83, 4.08]]
***
==============
O_threshold = 110
MC for this TARGET:[65.991, 0.195]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.15, -5.23, -5.42]][[-7.36, -65.99, -8.25]]
std:[[1.01, 0.99, 0.62]][[0.46, 0.0, 0.25]]
MSE:[[5.25, 5.32, 5.46]][[7.37, 65.99, 8.25]]
MSE(-DR):[[0.0, 0.07, 0.21]][[2.12, 60.74, 3.0]]
***
==============
O_threshold = 120
MC for this TARGET:[65.36, 0.182]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.11, -7.1, -6.96]][[-11.46, -65.36, -7.62]]
std:[[1.29, 1.29, 0.71]][[0.44, 0.0, 0.25]]
MSE:[[7.23, 7.22, 7.0]][[11.47, 65.36, 7.62]]
MSE(-DR):[[0.0, -0.01, -0.23]][[4.24, 58.13, 0.39]]
**
==============
[[ 0.95  0.86  0.58  2.5  66.87  9.13]
 [ 0.73  0.78  1.12  1.21 68.37 10.63]
 [ 0.55  0.58  1.35  0.38 68.39 10.65]
 [ 0.52  0.55  0.85  0.57 66.73  8.99]
 [ 2.85  2.97  3.18  4.18 66.96  9.23]
 [ 5.42  5.51  5.73  7.03 67.31  9.57]
 [ 5.17  5.25  5.43  7.37 65.98  8.25]
 [ 7.15  7.15  6.96 11.45 65.35  7.62]]


[[ 1.14  1.07  0.69  2.52 66.88  9.14]
 [ 0.94  0.97  1.15  1.24 68.38 10.64]
 [ 0.75  0.76  1.41  0.47 68.39 10.65]
 [ 0.7   0.71  0.92  0.64 66.74  8.99]
 [ 2.91  3.03  3.19  4.2  66.97  9.23]
 [ 5.49  5.59  5.76  7.04 67.32  9.57]
 [ 5.25  5.32  5.46  7.37 65.99  8.25]
 [ 7.23  7.22  7.   11.47 65.36  7.62]]


time spent until now: 254.8 mins
```