```
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
KeyboardInterrupt
ubuntu@ip-172-31-68-86:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
20:11, 04/06; num of cores:16

final sd_R trend for[0, 10, 20], 130 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,
16, None, None, None, 30, 0.5, 1, [True, False, True, 10], [0, 10, 20], None]


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

max(u_O) =  168.8
O_threshold = 80
number of reward locations:  15
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  6
O_threshold = 120
number of reward locations:  3
target 1 in 5 DONE!
target 2 in 5 DONE!
target 3 in 5 DONE!
target 4 in 5 DONE!
target 5 in 5 DONE!

---------------------------------------
Value of Behaviour policy:57.696
O_threshold = 80
MC for this TARGET:[68.355, 0.062]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.21, -0.43, -1.08]][[1.12, -68.36, -10.66]]
std:[[0.39, 0.39, 0.27]][[0.31, 0.0, 0.22]]
MSE:[[0.44, 0.58, 1.11]][[1.16, 68.36, 10.66]]
MSE(-DR):[[0.0, 0.14, 0.67]][[0.72, 67.92, 10.22]]
***
==============
O_threshold = 90
MC for this TARGET:[66.714, 0.067]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.05, -0.23, -0.78]][[-0.5, -66.71, -9.02]]
std:[[0.36, 0.36, 0.23]][[0.3, 0.0, 0.22]]
MSE:[[0.36, 0.43, 0.81]][[0.58, 66.71, 9.02]]
MSE(-DR):[[0.0, 0.07, 0.45]][[0.22, 66.35, 8.66]]
***
==============
O_threshold = 100
MC for this TARGET:[66.948, 0.071]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.79, -2.92, -3.21]][[-4.26, -66.95, -9.25]]
std:[[0.38, 0.4, 0.26]][[0.3, 0.0, 0.22]]
MSE:[[2.82, 2.95, 3.22]][[4.27, 66.95, 9.25]]
MSE(-DR):[[0.0, 0.13, 0.4]][[1.45, 64.13, 6.43]]
***
==============
O_threshold = 110
MC for this TARGET:[65.968, 0.07]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.22, -5.29, -5.44]][[-7.44, -65.97, -8.27]]
std:[[0.34, 0.36, 0.47]][[0.33, 0.0, 0.22]]
MSE:[[5.23, 5.3, 5.46]][[7.45, 65.97, 8.27]]
MSE(-DR):[[0.0, 0.07, 0.23]][[2.22, 60.74, 3.04]]
***
==============
O_threshold = 120
MC for this TARGET:[65.336, 0.059]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.16, -7.16, -7.01]][[-11.48, -65.34, -7.64]]
std:[[0.54, 0.56, 0.46]][[0.32, 0.0, 0.22]]
MSE:[[7.18, 7.18, 7.03]][[11.48, 65.34, 7.64]]
MSE(-DR):[[0.0, 0.0, -0.15]][[4.3, 58.16, 0.46]]
**
==============
[[ 0.44  0.58  1.11  1.16 68.36 10.66]
 [ 0.36  0.43  0.81  0.58 66.71  9.02]
 [ 2.82  2.95  3.22  4.27 66.95  9.25]
 [ 5.23  5.3   5.46  7.45 65.97  8.27]
 [ 7.18  7.18  7.03 11.48 65.34  7.64]]


time spent until now: 48.8 mins
```

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

max(u_O) =  168.8
O_threshold = 80
number of reward locations:  15
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  6
O_threshold = 120
number of reward locations:  3
target 1 in 5 DONE!
target 2 in 5 DONE!
target 3 in 5 DONE!
target 4 in 5 DONE!
target 5 in 5 DONE!

--------------------------------------
Value of Behaviour policy:57.682
O_threshold = 80
MC for this TARGET:[68.371, 0.129]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.28, -0.51, -1.06]][[1.05, -68.37, -10.69]]
std:[[0.71, 0.73, 0.41]][[0.36, 0.0, 0.2]]
MSE:[[0.76, 0.89, 1.14]][[1.11, 68.37, 10.69]]
MSE(-DR):[[0.0, 0.13, 0.38]][[0.35, 67.61, 9.93]]
***
==============
O_threshold = 90
MC for this TARGET:[66.73, 0.139]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.23, -0.41, -0.86]][[-0.54, -66.73, -9.05]]
std:[[0.61, 0.61, 0.4]][[0.28, 0.0, 0.2]]
MSE:[[0.65, 0.73, 0.95]][[0.61, 66.73, 9.05]]
MSE(-DR):[[0.0, 0.08, 0.3]][[-0.04, 66.08, 8.4]]
==============
O_threshold = 100
MC for this TARGET:[66.964, 0.143]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.73, -2.86, -3.19]][[-4.29, -66.96, -9.28]]
std:[[0.64, 0.64, 0.37]][[0.29, 0.0, 0.2]]
MSE:[[2.8, 2.93, 3.21]][[4.3, 66.96, 9.28]]
MSE(-DR):[[0.0, 0.13, 0.41]][[1.5, 64.16, 6.48]]
***
==============
O_threshold = 110
MC for this TARGET:[65.984, 0.143]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.28, -5.35, -5.4]][[-7.48, -65.98, -8.3]]
std:[[0.63, 0.62, 0.55]][[0.36, 0.0, 0.2]]
MSE:[[5.32, 5.39, 5.43]][[7.49, 65.98, 8.3]]
MSE(-DR):[[0.0, 0.07, 0.11]][[2.17, 60.66, 2.98]]
***
==============
O_threshold = 120
MC for this TARGET:[65.352, 0.129]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.11, -7.13, -7.0]][[-11.51, -65.35, -7.67]]
std:[[0.8, 0.8, 0.54]][[0.38, 0.0, 0.2]]
MSE:[[7.15, 7.17, 7.02]][[11.52, 65.35, 7.67]]
MSE(-DR):[[0.0, 0.02, -0.13]][[4.37, 58.2, 0.52]]
**
==============
[[ 0.44  0.58  1.11  1.16 68.36 10.66]
 [ 0.36  0.43  0.81  0.58 66.71  9.02]
 [ 2.82  2.95  3.22  4.27 66.95  9.25]
 [ 5.23  5.3   5.46  7.45 65.97  8.27]
 [ 7.18  7.18  7.03 11.48 65.34  7.64]]


[[ 0.76  0.89  1.14  1.11 68.37 10.69]
 [ 0.65  0.73  0.95  0.61 66.73  9.05]
 [ 2.8   2.93  3.21  4.3  66.96  9.28]
 [ 5.32  5.39  5.43  7.49 65.98  8.3 ]
 [ 7.15  7.17  7.02 11.52 65.35  7.67]]


time spent until now: 97.7 mins


--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 20]

max(u_O) =  168.8
O_threshold = 80
number of reward locations:  15
```

```
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  6
O_threshold = 120
number of reward locations:  3
target 1 in 5 DONE!
target 2 in 5 DONE!
target 3 in 5 DONE!
target 4 in 5 DONE!
target 5 in 5 DONE!

---------------------------------------
Value of Behaviour policy:57.668
O_threshold = 80
MC for this TARGET:[68.387, 0.236]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.35, -0.61, -1.01]][[0.99, -68.39, -10.72]]
std:[[1.17, 1.18, 0.59]][[0.51, 0.0, 0.24]]
MSE:[[1.22, 1.33, 1.17]][[1.11, 68.39, 10.72]]
MSE(-DR):[[0.0, 0.11, -0.05]][[-0.11, 67.17, 9.5]]
==============
O_threshold = 90
MC for this TARGET:[66.745, 0.246]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.4, -0.63, -0.89]][[-0.58, -66.74, -9.08]]
std:[[1.14, 1.12, 0.78]][[0.42, 0.0, 0.24]]
MSE:[[1.21, 1.29, 1.18]][[0.72, 66.74, 9.08]]
MSE(-DR):[[0.0, 0.08, -0.03]][[-0.49, 65.53, 7.87]]
==============
O_threshold = 100
MC for this TARGET:[66.98, 0.25]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.69, -2.81, -3.15]][[-4.32, -66.98, -9.31]]
std:[[1.1, 1.1, 0.64]][[0.44, 0.0, 0.24]]
MSE:[[2.91, 3.02, 3.21]][[4.34, 66.98, 9.31]]
MSE(-DR):[[0.0, 0.11, 0.3]][[1.43, 64.07, 6.4]]
***
==============
O_threshold = 110
MC for this TARGET:[65.999, 0.25]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.38, -5.45, -5.33]][[-7.53, -66.0, -8.33]]
std:[[1.17, 1.16, 0.78]][[0.5, 0.0, 0.24]]
MSE:[[5.51, 5.57, 5.39]][[7.55, 66.0, 8.33]]
MSE(-DR):[[0.0, 0.06, -0.12]][[2.04, 60.49, 2.82]]
**
==============
O_threshold = 120
MC for this TARGET:[65.368, 0.237]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.08, -7.14, -6.97]][[-11.55, -65.37, -7.7]]
std:[[1.21, 1.22, 0.77]][[0.54, 0.0, 0.24]]
MSE:[[7.18, 7.24, 7.01]][[11.56, 65.37, 7.7]]
MSE(-DR):[[0.0, 0.06, -0.17]][[4.38, 58.19, 0.52]]
**
==============
[[ 0.44  0.58  1.11  1.16 68.36 10.66]
 [ 0.36  0.43  0.81  0.58 66.71  9.02]
 [ 2.82  2.95  3.22  4.27 66.95  9.25]
 [ 5.23  5.3   5.46  7.45 65.97  8.27]
 [ 7.18  7.18  7.03 11.48 65.34  7.64]]


[[ 0.76  0.89  1.14  1.11 68.37 10.69]
 [ 0.65  0.73  0.95  0.61 66.73  9.05]
 [ 2.8   2.93  3.21  4.3  66.96  9.28]
 [ 5.32  5.39  5.43  7.49 65.98  8.3 ]
 [ 7.15  7.17  7.02 11.52 65.35  7.67]]


[[ 1.22  1.33  1.17  1.11 68.39 10.72]
 [ 1.21  1.29  1.18  0.72 66.74  9.08]
 [ 2.91  3.02  3.21  4.34 66.98  9.31]
 [ 5.51  5.57  5.39  7.55 66.    8.33]
 [ 7.18  7.24  7.01 11.56 65.37  7.7 ]]


time spent until now: 147.9 mins

ubuntu@ip-172-31-68-86:~$
```