```
Basic setting:[T, rep_times, sd_O, sd_D, sd_R, sd_u_O, w_O, w_A,
one, 96, 10, 10, None, 0.3, 0.5, 1, [True, False, True, False, 10


--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 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] 0:[tmux]*
```

```
-------------------------------------
Value of Behaviour policy:60.786
O_threshold = 80
MC for this TARGET:[70.884, 0.141]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.71, -4.34, -1.01]][[1.0, -36689.38, -10.1]]
std:[[0.39, 0.41, 0.38]][[0.35, 230353.71, 0.23]]
MSE:[[2.74, 4.36, 1.08]][[1.06, 233257.24, 10.1]]
MSE(-DR):[[0.0, 1.62, -1.66]][[-1.68, 233254.5, 7.36]]
==============


O_threshold = 90
MC for this TARGET:[69.371, 0.133]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.33, -3.78, -1.07]][[-0.68, -26825.5, -8.59]]
std:[[0.44, 0.45, 0.41]][[0.33, 184055.81, 0.23]]
```

```
MSE:[[2.37, 3.81, 1.15]][[0.76, 186000.4, 8.59]]
MSE(-DR):[[0.0, 1.44, -1.22]][[-1.61, 185998.03, 6.22]]
==============


O_threshold = 100
MC for this TARGET:[68.94, 0.132]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.93, -4.96, -3.34]][[-4.8, -56800.13, -8.15]]
std:[[0.43, 0.45, 0.39]][[0.32, 295712.03, 0.23]]
MSE:[[3.95, 4.98, 3.36]][[4.81, 301117.68, 8.15]]
MSE(-DR):[[0.0, 1.03, -0.59]][[0.86, 301113.73, 4.2]]
**
==============


O_threshold = 110
MC for this TARGET:[70.484, 0.135]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.46, -8.44, -6.74]][[-8.7, -93159.03, -9.7]]
std:[[0.53, 0.56, 0.43]][[0.32, 688277.61, 0.23]]
MSE:[[7.48, 8.46, 6.75]][[8.71, 694553.58, 9.7]]
MSE(-DR):[[0.0, 0.98, -0.73]][[1.23, 694546.1, 2.22]]
**
==============


[[2.7400e+00 4.3600e+00 1.0800e+00 1.0600e+00 2.3326e+05 1.0100e+0
 [2.3700e+00 3.8100e+00 1.1500e+00 7.6000e-01 1.8600e+05 8.5900e+0
 [3.9500e+00 4.9800e+00 3.3600e+00 4.8100e+00 3.0112e+05 8.1500e+0
 [7.4800e+00 8.4600e+00 6.7500e+00 8.7100e+00 6.9455e+05 9.7000e+0


time spent until now: 219.3 mins
```

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

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

87.8 97.8 52.4 162.7 58.1
```

```
--------------------------------------
Value of Behaviour policy:60.792
O_threshold = 80
MC for this TARGET:[70.887, 0.092]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.73, -4.36, -1.07]][[1.06, 52184.68, -10.09]]
std:[[0.45, 0.5, 0.3]][[0.27, 357274.38, 0.21]]
MSE:[[2.77, 4.39, 1.11]][[1.09, 361065.4, 10.09]]
MSE(-DR):[[0.0, 1.62, -1.66]][[-1.68, 361062.63, 7.32]]
==============


O_threshold = 90
MC for this TARGET:[69.373, 0.094]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.26, -3.73, -1.05]][[-0.6, -13503.74, -8.58]]
std:[[0.38, 0.4, 0.31]][[0.29, 399490.87, 0.21]]
MSE:[[2.29, 3.75, 1.09]][[0.67, 399719.03, 8.58]]
MSE(-DR):[[0.0, 1.46, -1.2]][[-1.62, 399716.74, 6.29]]
==============


O_threshold = 100
MC for this TARGET:[68.936, 0.097]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.91, -4.95, -3.31]][[-4.72, -65705.02, -8.14]]
std:[[0.38, 0.41, 0.33]][[0.3, 983876.31, 0.21]]
MSE:[[3.93, 4.97, 3.33]][[4.73, 986067.82, 8.14]]
MSE(-DR):[[0.0, 1.04, -0.6]][[0.8, 986063.89, 4.21]]
**
==============
```

```
V_threshold = 110
MC for this TARGET:[70.474, 0.102]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.35, -8.34, -6.67]][[-8.65, 23268.17, -9.68]]
std:[[0.44, 0.46, 0.4]][[0.3, 187838.64, 0.21]]
MSE:[[7.36, 8.35, 6.68]][[8.66, 189274.3, 9.68]]
MSE(-DR):[[0.0, 0.99, -0.68]][[1.3, 189266.94, 2.32]]
**
==============


[[2.7400e+00 4.3600e+00 1.0800e+00 1.0600e+00 2.3326e+05 1.0100e+
 [2.3700e+00 3.8100e+00 1.1500e+00 7.6000e-01 1.8600e+05 8.5900e+
 [3.9500e+00 4.9800e+00 3.3600e+00 4.8100e+00 3.0112e+05 8.1500e+
 [7.4800e+00 8.4600e+00 6.7500e+00 8.7100e+00 6.9455e+05 9.7000e+


[[2.7700e+00 4.3900e+00 1.1100e+00 1.0900e+00 3.6107e+05 1.0090e+
 [2.2900e+00 3.7500e+00 1.0900e+00 6.7000e-01 3.9972e+05 8.5800e+
 [3.9300e+00 4.9700e+00 3.3300e+00 4.7300e+00 9.8607e+05 8.1400e+
 [7.3600e+00 8.3500e+00 6.6800e+00 8.6600e+00 1.8927e+05 9.6800e+


time spent until now: 477.4 mins


------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 10]
```