

time spent until now: 121.7 mins

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

max(u_0) = 197.9
0_threshold = 80
means of 0order:

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

0_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

0_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

0_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!

2 -th region DONE!

3 -th region DONE!

4 -th region DONE!

5 -th region DONE!

6 -th region DONE!

7 -th region DONE!

8 -th region DONE!

9 -th region DONE!

10 -th region DONE!

11 -th region DONE!

12 -th region DONE!

13 -th region DONE!

```

Value of Behaviour policy:60.792
0_threshold = 80
MC for this TARGET:[70.887, 0.092]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.52, 0.37, -1.06]][[1.28, 31372.66, -10.09]]
std:[[0.56, 0.56, 0.31]][[0.28, 442245.99, 0.21]]
MSE:[[0.76, 0.67, 1.1]][[1.31, 443357.37, 10.09]]
MSE(-DR):[[0.0, -0.09, 0.34]][[0.55, 443356.61, 9.33]]
***
=====

0_threshold = 90
MC for this TARGET:[69.373, 0.094]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.01, -0.15, -1.04]][[-0.47, -31361.52, -8.58]]
std:[[0.53, 0.52, 0.32]][[0.31, 437401.7, 0.21]]
MSE:[[0.53, 0.54, 1.09]][[0.56, 438524.56, 8.58]]
MSE(-DR):[[0.0, 0.01, 0.56]][[0.03, 438524.03, 8.05]]
***
=====

0_threshold = 100
MC for this TARGET:[68.936, 0.097]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.85, -2.96, -3.31]][[-4.72, -87311.31, -8.14]]
std:[[0.57, 0.57, 0.33]][[0.32, 1233155.49, 0.21]]
MSE:[[2.91, 3.01, 3.33]][[4.73, 1236242.58, 8.14]]
MSE(-DR):[[0.0, 0.1, 0.42]][[1.82, 1236239.67, 5.23]]
***
=====

0_threshold = 110
MC for this TARGET:[70.474, 0.102]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.53, -6.61, -6.69]][[-8.72, 30246.33, -9.68]]
std:[[0.64, 0.64, 0.4]][[0.32, 216590.1, 0.21]]
MSE:[[6.56, 6.64, 6.7]][[8.73, 218691.82, 9.68]]
MSE(-DR):[[0.0, 0.08, 0.14]][[2.17, 218685.26, 3.12]]
***
=====

***** THIS SETTING IS GOOD *****
[[6.9000e-01 6.4000e-01 1.0900e+00 1.2800e+00 2.4311e+05 1.0100e+01]
 [6.6000e-01 7.1000e-01 1.1500e+00 6.7000e-01 1.8835e+05 8.5900e+00]
 [2.8400e+00 2.9500e+00 3.3500e+00 4.8300e+00 3.0359e+05 8.1500e+00]
 [6.4700e+00 6.5600e+00 6.7600e+00 8.8100e+00 5.7988e+05 9.7000e+00]]

[[7.6000e-01 6.7000e-01 1.1000e+00 1.3100e+00 4.4336e+05 1.0090e+01]
 [5.3000e-01 5.4000e-01 1.0900e+00 5.6000e-01 4.3852e+05 8.5800e+00]
 [2.9100e+00 3.0100e+00 3.3300e+00 4.7300e+00 1.2362e+06 8.1400e+00]
 [6.5600e+00 6.6400e+00 6.7000e+00 8.7300e+00 2.1869e+05 9.6800e+00]]

time spent until now: 251.3 mins

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

max(u_0) = 197.9
0_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

```