

```
Last login: Wed Apr  1 20:42:36 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~.ssh mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-3-233-226-188.compute-1.amazonaws.com
The authenticity of host 'ec2-3-233-226-188.compute-1.amazonaws.com (3.233.226.188)' can't be established.
ECDSA key fingerprint is SHA256:p8Vitif9oG9g9s0JwjJBH2/72BnAacXN45Br1DGN8jKg.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-3-233-226-188.compute-1.amazonaws.com,3.233.226.188' (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 Thu Apr  2 00:46:26 UTC 2020
```

```
Last login: Wed Apr  1 20:30:39 2020 from 107.13.161.147
ubuntu@ip-172-31-75-213:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
20:47, 04/01; num of cores:36
```

```
Basic setting:[T, sd_0, sd_D, sd_R, sd_u_0, w_0, w_A, [M_in_R, mean_reversion, pois0, simple, u_0_u_D]] = [None, 10, 10, 10, 0.3, 0.5, 1, [True, False, True, False, 10]]
```

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

```
max(u_0) = 197.9
0_threshold = 70
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
```

```
1 1 0 1 1
```

```
1 1 1 1 1
```

```
1 1 1 1 1
```

```
1 1 1 1 1
```

```
number of reward locations: 22
```

```
0_threshold = 80
```

```
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

Value of Behaviour policy:60.775

0_threshold = 70

MC for this TARGET:[69.312, 0.091]

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]

bias:[[3.38, 3.25, 1.38]][[5.14, -69.31, -8.54]]

std:[[0.7, 0.7, 0.33]][[0.26, 0.0, 0.14]]

MSE:[[3.45, 3.32, 1.42]][[5.15, 69.31, 8.54]]

MSE(-DR):[[0.0, -0.13, -2.03]][[1.7, 65.86, 5.09]]

=====

0_threshold = 80

MC for this TARGET:[70.894, 0.091]

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]

bias:[[0.73, 0.58, -0.98]][[1.25, -70.89, -10.12]]

std:[[0.54, 0.54, 0.29]][[0.25, 0.0, 0.14]]

MSE:[[0.91, 0.79, 1.02]][[1.27, 70.89, 10.12]]

MSE(-DR):[[0.0, -0.12, 0.11]][[0.36, 69.98, 9.21]]

=====

0_threshold = 90

MC for this TARGET:[69.377, 0.097]

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]

bias:[[0.13, -0.0, -0.97]][[-0.51, -69.38, -8.6]]

std:[[0.38, 0.38, 0.27]][[0.22, 0.0, 0.14]]

MSE:[[0.4, 0.38, 1.01]][[0.56, 69.38, 8.6]]

MSE(-DR):[[0.0, -0.02, 0.61]][[0.16, 68.98, 8.2]]

=====

0_threshold = 100

MC for this TARGET:[68.925, 0.09]

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]

bias:[[-2.94, -3.04, -3.25]][[-4.74, -68.92, -8.15]]

std:[[0.49, 0.5, 0.25]][[0.23, 0.0, 0.14]]

MSE:[[2.98, 3.08, 3.26]][[4.75, 68.92, 8.15]]

MSE(-DR):[[0.0, 0.1, 0.28]][[1.77, 65.94, 5.17]]

=====

0_threshold = 110

MC for this TARGET:[70.467, 0.083]

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]

bias:[[-6.54, -6.61, -6.51]][[-8.73, -70.47, -9.69]]

std:[[0.55, 0.55, 0.32]][[0.21, 0.0, 0.14]]

MSE:[[6.56, 6.63, 6.52]][[8.73, 70.47, 9.69]]

MSE(-DR):[[0.0, 0.07, -0.04]][[2.17, 63.91, 3.13]]

=====

[[3.45 3.32 1.42 5.15 69.31 8.54]

[0.91 0.79 1.02 1.27 70.89 10.12]

[0.4 0.38 1.01 0.56 69.38 8.6]

[2.98 3.08 3.26 4.75 68.92 8.15]

[6.56 6.63 6.52 8.73 70.47 9.69]]

time spent until now: 73.2 mins

ubuntu@ip-172-31-75-213:~\$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py

22:05, 04/01; num of cores:36

Basic setting:[T, sd_0, sd_D, sd_R, sd_u_0, w_0, w_A, [M_in_R, mean_reversion, pois0, simple, u_0_u_D]] = [None, 10, 10, 10, 0.3, 0.5, 1, [True, False, True, False, 10]]

[pattern_seed, day] = [2, 10]

max(u_0) = 197.9

0_threshold = 70

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

1 1 0 1 1

1 1 1 1 1

1 1 1 1 1

1 1 1 1 1

number of reward locations: 22

0_threshold = 80

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

```
-----
Value of Behaviour policy:60.798
0_threshold = 70
MC for this TARGET:[69.315, 0.095]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[2.74, 2.61, 1.37]][[5.18, -69.32, -8.52]]
std:[[0.98, 0.97, 0.43]][[0.26, 0.0, 0.23]]
MSE:[[2.91, 2.78, 1.44]][[5.19, 69.32, 8.52]]
MSE(-DR):[[0.0, -0.13, -1.47]][[2.28, 66.41, 5.61]]
***
=====
```

```
0_threshold = 80
MC for this TARGET:[70.887, 0.092]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.58, 0.43, -1.09]][[1.32, -70.89, -10.09]]
std:[[0.66, 0.65, 0.33]][[0.31, 0.0, 0.23]]
MSE:[[0.88, 0.78, 1.14]][[1.36, 70.89, 10.09]]
MSE(-DR):[[0.0, -0.1, 0.26]][[0.48, 70.01, 9.21]]
***
=====
```

```
0_threshold = 90
MC for this TARGET:[69.373, 0.094]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.12, -0.02, -1.01]][[-0.45, -69.37, -8.58]]
std:[[0.6, 0.59, 0.34]][[0.31, 0.0, 0.23]]
MSE:[[0.61, 0.59, 1.07]][[0.55, 69.37, 8.58]]
MSE(-DR):[[0.0, -0.02, 0.46]][[-0.06, 68.76, 7.97]]
=====
```

```
0_threshold = 100
MC for this TARGET:[68.936, 0.097]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.75, -2.86, -3.32]][[-4.7, -68.94, -8.14]]
std:[[0.65, 0.65, 0.39]][[0.33, 0.0, 0.23]]
MSE:[[2.83, 2.93, 3.34]][[4.71, 68.94, 8.14]]
MSE(-DR):[[0.0, 0.1, 0.51]][[1.88, 66.11, 5.31]]
***
=====
```

```
0_threshold = 110
MC for this TARGET:[70.474, 0.102]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.49, -6.56, -6.74]][[-8.72, -70.47, -9.68]]
std:[[0.72, 0.72, 0.43]][[0.32, 0.0, 0.23]]
MSE:[[6.53, 6.6, 6.75]][[8.73, 70.47, 9.68]]
MSE(-DR):[[0.0, 0.07, 0.22]][[2.2, 63.94, 3.15]]
***
=====
```

```
[ [ 2.91  2.78  1.44  5.19 69.32  8.52]
  [ 0.88  0.78  1.14  1.36 70.89 10.09]
  [ 0.61  0.59  1.07  0.55 69.37  8.58]
  [ 2.83  2.93  3.34  4.71 68.94  8.14]
  [ 6.53  6.6   6.75  8.73 70.47  9.68]]
```

time spent until now: 70.6 mins

ubuntu@ip-172-31-75-213:~\$