

```
Last login: Thu Apr 16 11:52:38 on ttys001
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-34-201-49-219.compute-1.amazonaws.com
Welcome to Ubuntu 18.04.3 LTS (GNU/Linux 4.15.0-1060-aws x86_64)
```

```
* Documentation:  https://help.ubuntu.com
* Management:    https://landscape.canonical.com
* Support:        https://ubuntu.com/advantage
```

System information disabled due to load higher than 96.0

```
* 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
```

```
51 packages can be updated.
0 updates are security updates.
```

```
*** System restart required ***
```

```
Last login: Thu Apr 16 15:52:41 2020 from 107.13.161.147
```

```
ubuntu@ip-172-31-0-227:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
15:12, 04/16; num of cores:96
u_0_u_D_range_15_sd
```

```
Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D_range, t_func] = [16, None, None, 20, 0.5, 1.5, [15], N
one]
```

```
[thre_range, sd_R_range, day_range, penalty_range]:  [[100, 100.5, 105, 110, 110.5, 111], [0, 15, 30], [7], [[0.0003, 0
.0001, 5e-05], [0.0003, 0.0001, 5e-05]]]
```

```
-----
[pattern_seed, day, sd_R, u_0_u_D] = [2, 7, 0, 15]
```

```
max(u_0) = 145.8
0_threshold = 100
means of Order:
```

```
91.7 98.9 57.3 132.8 64.1
```

```
83.2 110.1 75.1 78.8 81.8
```

```
111.0 145.8 100.8 77.6 110.8
```

```
88.1 99.6 123.5 85.0 100.2
```

```
82.4 96.9 105.1 80.2 93.2
```

```
target policy:
```

```
0 0 0 1 0
```

```
0 1 0 0 0
```

```
1 1 1 0 1
```

```
0 0 1 0 1
```

```
0 0 1 0 0
```

```
number of reward locations: 9
```

```
0_threshold = 100.5
```

```
number of reward locations: 8
```

```
0_threshold = 105
```

```
number of reward locations: 7
```

```
0_threshold = 110
```

```
number of reward locations: 6
```

```
0_threshold = 110.5
```

```
number of reward locations: 5
0_threshold = 111
number of reward locations: 4
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
```

```
-----
Value of Behaviour policy:51.308
0_threshold = 100
MC for this TARGET:[56.694, 0.075]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.85, -1.02, -1.44]][[-2.06, -56.69, -5.39]]
std:[[0.6, 0.58, 0.41]][[0.31, 0.0, 0.21]]
MSE:[1.04, 1.17, 1.5][[2.08, 56.69, 5.39]]
MSE(-DR):[[0.0, 0.13, 0.46]][[1.04, 55.65, 4.35]]
***
```

```
=====
0_threshold = 100.5
MC for this TARGET:[53.941, 0.07]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.75, 0.6, -0.03]][[-1.24, -53.94, -2.63]]
std:[[0.5, 0.47, 0.42]][[0.3, 0.0, 0.21]]
MSE:[0.9, 0.76, 0.42][[1.28, 53.94, 2.64]]
MSE(-DR):[[0.0, -0.14, -0.48]][[0.38, 53.04, 1.74]]
**
```

```
=====
0_threshold = 105
MC for this TARGET:[55.699, 0.072]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.62, -2.78, -3.15]][[-4.58, -55.7, -4.39]]
std:[[0.7, 0.69, 0.4]][[0.32, 0.0, 0.21]]
MSE:[2.71, 2.86, 3.18][[4.59, 55.7, 4.4]]
MSE(-DR):[[0.0, 0.15, 0.47]][[1.88, 52.99, 1.69]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[54.532, 0.064]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.35, -2.45, -3.0]][[-5.52, -54.53, -3.22]]
std:[[0.58, 0.56, 0.44]][[0.32, 0.0, 0.21]]
MSE:[2.42, 2.51, 3.03][[5.53, 54.53, 3.23]]
MSE(-DR):[[0.0, 0.09, 0.61]][[3.11, 52.11, 0.81]]
***
```

```
=====
0_threshold = 110.5
MC for this TARGET:[57.162, 0.058]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.27, -5.38, -5.57]][[-9.9, -57.16, -5.85]]
std:[[0.66, 0.62, 0.53]][[0.3, 0.0, 0.21]]
MSE:[5.31, 5.42, 5.6][[9.9, 57.16, 5.85]]
MSE(-DR):[[0.0, 0.11, 0.29]][[4.59, 51.85, 0.54]]
***
```

```
=====
0_threshold = 111
MC for this TARGET:[55.171, 0.056]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-4.97, -5.02, -5.31]][[-10.05, -55.17, -3.86]]
std:[[0.6, 0.61, 0.5]][[0.28, 0.0, 0.21]]
MSE:[5.01, 5.06, 5.33][[10.05, 55.17, 3.87]]
MSE(-DR):[[0.0, 0.05, 0.32]][[5.04, 50.16, -1.14]]
***
```

```
=====
[[ 1.04  1.17  1.5  2.08 56.69  5.39]
 [ 0.9   0.76  0.42  1.28 53.94  2.64]
 [ 2.71  2.86  3.18  4.59 55.7   4.4 ]
 [ 2.42  2.51  3.03  5.53 54.53  3.23]
 [ 5.31  5.42  5.6   9.9  57.16  5.85]
 [ 5.01  5.06  5.33 10.05 55.17  3.87]]
```

time spent until now: 21.0 mins

15:33, 04/16

```
-----  
[pattern_seed, day, sd_R, u_0_u_D] = [2, 7, 15, 15]
```

```
max(u_0) = 145.8  
0_threshold = 100  
means of Order:
```

```
91.7 98.9 57.3 132.8 64.1  
83.2 110.1 75.1 78.8 81.8  
111.0 145.8 100.8 77.6 110.8  
88.1 99.6 123.5 85.0 100.2  
82.4 96.9 105.1 80.2 93.2
```

```
target policy:
```

```
0 0 0 1 0  
0 1 0 0 0  
1 1 1 0 1  
0 0 1 0 1  
0 0 1 0 0
```

```
number of reward locations: 9  
0_threshold = 100.5  
number of reward locations: 8  
0_threshold = 105  
number of reward locations: 7  
0_threshold = 110  
number of reward locations: 6  
0_threshold = 110.5  
number of reward locations: 5  
0_threshold = 111  
number of reward locations: 4  
target 1 in 1 DONE!  
target 1 in 1 DONE!  
target 1 in 1 DONE!  
target 1 in 1 DONE!  
target 1 in 1 DONE!  
target 1 in 1 DONE!
```

```
-----  
Value of Behaviour policy:51.287  
0_threshold = 100  
MC for this TARGET:[56.718, 0.189]  
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]  
bias:[[-0.77, -0.94, -1.55]][[-2.12, -56.72, -5.43]]  
std:[[0.73, 0.75, 0.56]][[0.31, 0.0, 0.21]]  
MSE:[[1.06, 1.2, 1.65]][[2.14, 56.72, 5.43]]  
MSE(-DR):[[0.0, 0.14, 0.59]][[1.08, 55.66, 4.37]]
```

```
***  
=====
```

```
0_threshold = 100.5  
MC for this TARGET:[53.965, 0.182]  
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]  
bias:[[0.92, 0.78, -0.09]][[-1.3, -53.96, -2.68]]  
std:[[0.66, 0.69, 0.46]][[0.3, 0.0, 0.21]]  
MSE:[[1.13, 1.04, 0.47]][[1.33, 53.96, 2.69]]  
MSE(-DR):[[0.0, -0.09, -0.66]][[0.2, 52.83, 1.56]]
```

```
***  
=====
```

```
0_threshold = 105  
MC for this TARGET:[55.722, 0.182]  
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]  
bias:[[-2.44, -2.62, -3.24]][[-4.67, -55.72, -4.44]]  
std:[[0.76, 0.74, 0.63]][[0.32, 0.0, 0.21]]  
MSE:[[2.56, 2.72, 3.3]][[4.68, 55.72, 4.44]]  
MSE(-DR):[[0.0, 0.16, 0.74]][[2.12, 53.16, 1.88]]
```

```
***  
=====
```

```
0_threshold = 110  
MC for this TARGET:[54.556, 0.18]
```

```

[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.35, -2.45, -3.1]][[-5.58, -54.56, -3.27]]
std:[0.67, 0.64, 0.68][0.36, 0.0, 0.21]
MSE:[2.44, 2.53, 3.17][5.59, 54.56, 3.28]
MSE(-DR):[0.0, 0.09, 0.73][3.15, 52.12, 0.84]
***
=====
0_threshold = 110.5
MC for this TARGET:[57.185, 0.178]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.48, -5.56, -5.71]][[-9.93, -57.18, -5.9]]
std:[0.87, 0.81, 0.77][0.36, 0.0, 0.21]
MSE:[5.55, 5.62, 5.76][9.94, 57.18, 5.9]
MSE(-DR):[0.0, 0.07, 0.21][4.39, 51.63, 0.35]
***
=====
0_threshold = 111
MC for this TARGET:[55.194, 0.18]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.01, -5.04, -5.32]][[-10.07, -55.19, -3.91]]
std:[0.94, 0.9, 0.76][0.36, 0.0, 0.21]
MSE:[5.1, 5.12, 5.37][10.08, 55.19, 3.92]
MSE(-DR):[0.0, 0.02, 0.27][4.98, 50.09, -1.18]
***
=====
[[ 1.04  1.17  1.5   2.08 56.69  5.39]
 [ 0.9   0.76  0.42  1.28 53.94  2.64]
 [ 2.71  2.86  3.18  4.59 55.7   4.4 ]
 [ 2.42  2.51  3.03  5.53 54.53  3.23]
 [ 5.31  5.42  5.6   9.9  57.16  5.85]
 [ 5.01  5.06  5.33 10.05 55.17  3.87]]

[[ 1.06  1.2   1.65  2.14 56.72  5.43]
 [ 1.13  1.04  0.47  1.33 53.96  2.69]
 [ 2.56  2.72  3.3   4.68 55.72  4.44]
 [ 2.44  2.53  3.17  5.59 54.56  3.28]
 [ 5.55  5.62  5.76  9.94 57.18  5.9 ]
 [ 5.1   5.12  5.37 10.08 55.19  3.92]]

```

time spent until now: 42.1 mins

15:54, 04/16

```

[pattern_seed, day, sd_R, u_0_u_D] = [2, 7, 30, 15]

```

```

max(u_0) = 145.8
0_threshold = 100
means of Order:

```

```

91.7 98.9 57.3 132.8 64.1

83.2 110.1 75.1 78.8 81.8

111.0 145.8 100.8 77.6 110.8

88.1 99.6 123.5 85.0 100.2

82.4 96.9 105.1 80.2 93.2

```

target policy:

```

0 0 0 1 0

0 1 0 0 0

1 1 1 0 1

0 0 1 0 1

0 0 1 0 0

```

```

number of reward locations: 9
0_threshold = 100.5
number of reward locations: 8
0_threshold = 105

```

```

number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 110.5
number of reward locations: 5
0_threshold = 111
number of reward locations: 4
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!
target 1 in 1 DONE!

```

```

-----
Value of Behaviour policy:51.265
0_threshold = 100
MC for this TARGET:[56.742, 0.354]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.78, -0.95, -1.63]][[-2.14, -56.74, -5.48]]
std:[1.26, 1.23, 0.91]][[0.53, 0.0, 0.31]]
MSE:[1.48, 1.55, 1.87]][[2.2, 56.74, 5.49]]
MSE(-DR):[[0.0, 0.07, 0.39]][[0.72, 55.26, 4.01]]
***
=====
0_threshold = 100.5
MC for this TARGET:[53.988, 0.348]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[1.03, 0.86, -0.09]][[-1.34, -53.99, -2.72]]
std:[1.24, 1.23, 0.78]][[0.53, 0.0, 0.31]]
MSE:[1.61, 1.5, 0.79]][[1.44, 53.99, 2.74]]
MSE(-DR):[[0.0, -0.11, -0.82]][[-0.17, 52.38, 1.13]]
=====
0_threshold = 105
MC for this TARGET:[55.746, 0.348]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.35, -2.54, -3.39]][[-4.75, -55.75, -4.48]]
std:[1.32, 1.32, 0.89]][[0.55, 0.0, 0.31]]
MSE:[2.7, 2.86, 3.5]][[4.78, 55.75, 4.49]]
MSE(-DR):[[0.0, 0.16, 0.8]][[2.08, 53.05, 1.79]]
***
=====
0_threshold = 110
MC for this TARGET:[54.58, 0.347]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.38, -2.53, -3.13]][[-5.63, -54.58, -3.31]]
std:[1.2, 1.15, 1.01]][[0.61, 0.0, 0.31]]
MSE:[2.67, 2.78, 3.29]][[5.66, 54.58, 3.32]]
MSE(-DR):[[0.0, 0.11, 0.62]][[2.99, 51.91, 0.65]]
***
=====
0_threshold = 110.5
MC for this TARGET:[57.209, 0.346]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.72, -5.83, -5.73]][[-9.97, -57.21, -5.94]]
std:[1.44, 1.33, 1.25]][[0.63, 0.0, 0.31]]
MSE:[5.9, 5.98, 5.86]][[9.99, 57.21, 5.95]]
MSE(-DR):[[0.0, 0.08, -0.04]][[4.09, 51.31, 0.05]]
**
=====
0_threshold = 111
MC for this TARGET:[55.218, 0.349]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.06, -5.11, -5.33]][[-10.09, -55.22, -3.95]]
std:[1.53, 1.5, 1.26]][[0.66, 0.0, 0.31]]
MSE:[5.29, 5.33, 5.48]][[10.11, 55.22, 3.96]]
MSE(-DR):[[0.0, 0.04, 0.19]][[4.82, 49.93, -1.33]]
***
=====
[[ 1.04  1.17  1.5   2.08 56.69  5.39]
 [ 0.9   0.76  0.42  1.28 53.94  2.64]
 [ 2.71  2.86  3.18  4.59 55.7   4.4 ]
 [ 2.42  2.51  3.03  5.53 54.53  3.23]
 [ 5.31  5.42  5.6   9.9   57.16  5.85]
 [ 5.01  5.06  5.33 10.05 55.17  3.87]]

[[ 1.06  1.2   1.65  2.14 56.72  5.43]

```

```
[ 1.13  1.04  0.47  1.33 53.96  2.69]
[ 2.56  2.72  3.3   4.68 55.72  4.44]
[ 2.44  2.53  3.17  5.59 54.56  3.28]
[ 5.55  5.62  5.76  9.94 57.18  5.9 ]
[ 5.1   5.12  5.37 10.08 55.19  3.92]]
```

```
[[ 1.48  1.55  1.87  2.2  56.74  5.49]
 [ 1.61  1.5   0.79  1.44 53.99  2.74]
 [ 2.7   2.86  3.5   4.78 55.75  4.49]
 [ 2.67  2.78  3.29  5.66 54.58  3.32]
 [ 5.9   5.98  5.86  9.99 57.21  5.95]
 [ 5.29  5.33  5.48 10.11 55.22  3.96]]
```

time spent until now: 62.2 mins

16:14, 04/16

ubuntu@ip-172-31-0-227:~\$