

```
Last login: Fri Apr 10 18:09:50 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~ mac$ ssh -i "Runzhe_Song_0110.pem" ubuntu@ec2-18-232-186-156.compute-1.amazonaws.com
Warning: Permanently added the ED25519 host key for IP address '18.232.186.156' 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 Fri Apr 10 22:13:00 UTC 2020

```
System load:  6.85          Processes:      881
Usage of /:   57.2% of 15.45GB Users logged in: 0
Memory usage: 0%          IP address for ens5: 172.31.73.191
Swap usage:   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
```

```
62 packages can be updated.
11 updates are security updates.
```

```
Last login: Wed Apr 1 20:30:39 2020 from 107.13.161.147
ubuntu@ip-172-31-73-191:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
18:14, 04/10; num of cores:96
vary_T_sd_15_large
```

```
Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, sd_R_range, t_func] = [96, None, None, 20, 0.5, 1, 0, [15], None]
```

```
-----
[pattern_seed, day, sd_R] = [2, 6, 15]
```

```
max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!
```

```
-----
Value of Behaviour policy:64.871
0_threshold = 100
MC for this TARGET:[70.774, 0.219]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.39, -0.47, -1.02]][[-1.36, -70.77, -5.9]]
std:[[0.81, 0.8, 0.5]][[0.42, 0.0, 0.29]]
MSE:[[0.9, 0.93, 1.14]][[1.42, 70.77, 5.91]]
MSE(-DR):[[0.0, 0.03, 0.24]][[0.52, 69.87, 5.01]]
***
```

```
=====
0_threshold = 105
MC for this TARGET:[71.777, 0.212]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.73, -2.81, -3.45]][[-4.08, -71.78, -6.91]]
std:[[1.08, 1.08, 0.62]][[0.42, 0.0, 0.29]]
MSE:[[2.94, 3.01, 3.51]][[4.1, 71.78, 6.92]]
MSE(-DR):[[0.0, 0.07, 0.57]][[1.16, 68.84, 3.98]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[70.874, 0.212]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.74, -2.79, -3.24]][[-4.8, -70.87, -6.0]]
std:[[1.21, 1.2, 0.64]][[0.44, 0.0, 0.29]]
MSE:[[3.0, 3.04, 3.3]][[4.82, 70.87, 6.01]]
MSE(-DR):[[0.0, 0.04, 0.3]][[1.82, 67.87, 3.01]]
***
=====
```

```

0_threshold = 115
MC for this TARGET:[71.793, 0.205]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.63, -5.64, -5.87]][[-9.66, -71.79, -6.92]]
std:[[1.27, 1.29, 0.79]][[0.48, 0.0, 0.29]]
MSE:[[5.77, 5.79, 5.92]][[9.67, 71.79, 6.93]]
MSE(-DR):[[0.0, 0.02, 0.15]][[3.9, 66.02, 1.16]]
***
=====
***** THIS SETTING IS GOOD *****
[[ 0.9  0.93  1.14  1.42 70.77  5.91]
 [ 2.94 3.01 3.51  4.1 71.78  6.92]
 [ 3.   3.04 3.3  4.82 70.87  6.01]
 [ 5.77 5.79 5.92  9.67 71.79  6.93]]

```

time spent until now: 61.7 mins

19:16, 04/10

```

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

```

```

max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

```

```

-----
Value of Behaviour policy:64.885

```

```

0_threshold = 100
MC for this TARGET:[70.8, 0.2]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.33, -0.43, -0.96]][[-1.31, -70.8, -5.91]]
std:[[0.74, 0.75, 0.51]][[0.43, 0.0, 0.27]]
MSE:[[0.81, 0.86, 1.09]][[1.38, 70.8, 5.92]]
MSE(-DR):[[0.0, 0.05, 0.28]][[0.57, 69.99, 5.11]]
***

```

```

=====
0_threshold = 105
MC for this TARGET:[71.803, 0.194]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.71, -2.82, -3.41]][[-4.03, -71.8, -6.92]]
std:[[1.01, 1.0, 0.58]][[0.4, 0.0, 0.27]]
MSE:[[2.89, 2.99, 3.46]][[4.05, 71.8, 6.93]]
MSE(-DR):[[0.0, 0.1, 0.57]][[1.16, 68.91, 4.04]]
***

```

```

=====
0_threshold = 110
MC for this TARGET:[70.897, 0.193]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.77, -2.85, -3.19]][[-4.74, -70.9, -6.01]]
std:[[1.0, 1.0, 0.57]][[0.39, 0.0, 0.27]]
MSE:[[2.94, 3.02, 3.24]][[4.76, 70.9, 6.02]]
MSE(-DR):[[0.0, 0.08, 0.3]][[1.82, 67.96, 3.08]]
***

```

```

=====
0_threshold = 115
MC for this TARGET:[71.821, 0.187]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.82, -5.86, -5.85]][[-9.62, -71.82, -6.94]]
std:[[1.23, 1.23, 0.73]][[0.39, 0.0, 0.27]]
MSE:[[5.95, 5.99, 5.9]][[9.63, 71.82, 6.95]]
MSE(-DR):[[0.0, 0.04, -0.05]][[3.68, 65.87, 1.0]]
**

```

```

=====
[[ 0.9  0.93  1.14  1.42 70.77  5.91]
 [ 2.94 3.01 3.51  4.1 71.78  6.92]
 [ 3.   3.04 3.3  4.82 70.87  6.01]
 [ 5.77 5.79 5.92  9.67 71.79  6.93]]

```

```

[[ 0.81 0.86 1.09  1.38 70.8  5.92]
 [ 2.89 2.99 3.46  4.05 71.8  6.93]
 [ 2.94 3.02 3.24  4.76 70.9  6.02]
 [ 5.95 5.99 5.9  9.63 71.82  6.95]]

```

time spent until now: 126.5 mins

20:20, 04/10

-----  
[pattern\_seed, day, sd\_R] = [2, 8, 15]

```
max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!
```

-----  
Value of Behaviour policy:64.895

```
0_threshold = 100
MC for this TARGET:[70.777, 0.2]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.33, -0.42, -0.96]][[-1.37, -70.78, -5.88]]
std:[0.79, 0.78, 0.51][0.41, 0.0, 0.26]]
MSE:[0.86, 0.89, 1.08][1.43, 70.78, 5.89]]
MSE(-DR):[0.0, 0.03, 0.22][0.57, 69.92, 5.03]]
***
```

```
=====
0_threshold = 105
MC for this TARGET:[71.778, 0.194]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.74, -2.82, -3.36]][[-4.07, -71.78, -6.88]]
std:[0.96, 0.97, 0.53][0.41, 0.0, 0.26]]
MSE:[2.9, 2.98, 3.4][4.09, 71.78, 6.88]]
MSE(-DR):[0.0, 0.08, 0.5][1.19, 68.88, 3.98]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[70.871, 0.193]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.77, -2.83, -3.15]][[-4.76, -70.87, -5.98]]
std:[1.06, 1.05, 0.53][0.41, 0.0, 0.26]]
MSE:[2.97, 3.02, 3.19][4.78, 70.87, 5.99]]
MSE(-DR):[0.0, 0.05, 0.22][1.81, 67.9, 3.02]]
***
```

```
=====
0_threshold = 115
MC for this TARGET:[71.798, 0.191]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.97, -5.99, -5.8]][[-9.7, -71.8, -6.9]]
std:[1.3, 1.29, 0.67][0.4, 0.0, 0.26]]
MSE:[6.11, 6.13, 5.84][9.71, 71.8, 6.9]]
MSE(-DR):[0.0, 0.02, -0.27][3.6, 65.69, 0.79]]
**
```

```
=====
[[ 0.9  0.93  1.14  1.42 70.77  5.91]
 [ 2.94  3.01  3.51  4.1  71.78  6.92]
 [ 3.    3.04  3.3   4.82 70.87  6.01]
 [ 5.77  5.79  5.92  9.67 71.79  6.93]]
```

```
[[ 0.81  0.86  1.09  1.38 70.8   5.92]
 [ 2.89  2.99  3.46  4.05 71.8   6.93]
 [ 2.94  3.02  3.24  4.76 70.9   6.02]
 [ 5.95  5.99  5.9   9.63 71.82  6.95]]
```

```
[[ 0.86  0.89  1.08  1.43 70.78  5.89]
 [ 2.9   2.98  3.4   4.09 71.78  6.88]
 [ 2.97  3.02  3.19  4.78 70.87  5.99]
 [ 6.11  6.13  5.84  9.71 71.8   6.9 ]]
```

time spent until now: 195.2 mins

21:29, 04/10

-----  
[pattern\_seed, day, sd\_R] = [2, 9, 15]

```
max(u_0) = 145.8
0_threshold = 100
number of reward locations: 9
0_threshold = 105
number of reward locations: 7
0_threshold = 110
number of reward locations: 6
```

```
0_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!
```

```
-----
Value of Behaviour policy:64.887
```

```
0_threshold = 100
MC for this TARGET:[70.786, 0.164]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.41, -0.51, -1.02]][[-1.37, -70.79, -5.9]]
std:[[0.72, 0.71, 0.48]][[0.37, 0.0, 0.21]]
MSE:[[0.83, 0.87, 1.13]][[1.42, 70.79, 5.9]]
MSE(-DR):[[0.0, 0.04, 0.3]][[0.59, 69.96, 5.07]]
***
```

```
=====
0_threshold = 105
MC for this TARGET:[71.785, 0.163]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.79, -2.87, -3.42]][[-4.05, -71.78, -6.9]]
std:[[0.9, 0.89, 0.55]][[0.35, 0.0, 0.21]]
MSE:[[2.93, 3.0, 3.46]][[4.07, 71.78, 6.9]]
MSE(-DR):[[0.0, 0.07, 0.53]][[1.14, 68.85, 3.97]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[70.88, 0.16]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.88, -2.96, -3.19]][[-4.78, -70.88, -5.99]]
std:[[0.94, 0.93, 0.53]][[0.35, 0.0, 0.21]]
MSE:[[3.03, 3.1, 3.23]][[4.79, 70.88, 5.99]]
MSE(-DR):[[0.0, 0.07, 0.2]][[1.76, 67.85, 2.96]]
***
```

```
=====
0_threshold = 115
MC for this TARGET:[71.808, 0.165]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.12, -6.15, -5.81]][[-9.69, -71.81, -6.92]]
std:[[1.15, 1.14, 0.59]][[0.34, 0.0, 0.21]]
MSE:[[6.23, 6.25, 5.84]][[9.7, 71.81, 6.92]]
MSE(-DR):[[0.0, 0.02, -0.39]][[3.47, 65.58, 0.69]]
***
```

```
=====
[[ 0.9  0.93  1.14  1.42 70.77  5.91]
 [ 2.94  3.01  3.51  4.1  71.78  6.92]
 [ 3.    3.04  3.3  4.82 70.87  6.01]
 [ 5.77  5.79  5.92  9.67 71.79  6.93]]
```

```
[[ 0.81  0.86  1.09  1.38 70.8  5.92]
 [ 2.89  2.99  3.46  4.05 71.8  6.93]
 [ 2.94  3.02  3.24  4.76 70.9  6.02]
 [ 5.95  5.99  5.9  9.63 71.82  6.95]]
```

```
[[ 0.86  0.89  1.08  1.43 70.78  5.89]
 [ 2.9  2.98  3.4  4.09 71.78  6.88]
 [ 2.97  3.02  3.19  4.78 70.87  5.99]
 [ 6.11  6.13  5.84  9.71 71.8  6.9 ]]
```

```
[[ 0.83  0.87  1.13  1.42 70.79  5.9 ]
 [ 2.93  3.    3.46  4.07 71.78  6.9 ]
 [ 3.03  3.1  3.23  4.79 70.88  5.99]
 [ 6.23  6.25  5.84  9.7  71.81  6.92]]
```

```
time spent until now: 269.5 mins
```

```
22:43, 04/10
```

```
ubuntu@ip-172-31-73-191:~$ large
```