

```
Last login: Sun Apr 12 19:44:36 on ttys001
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-35-170-52-91.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 as of Sun Apr 12 23:48:02 UTC 2020

```
System load:  1.57          Processes:      213
Usage of /:   28.6% of 30.96GB Users logged in:  0
Memory usage: 1%           IP address for ens5: 172.31.7.151
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>

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

*** System restart required ***

```
Last login: Sun Apr 12 23:44:39 2020 from 107.13.161.147
ubuntu@ip-172-31-7-151:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:49, 04/12; num of cores:16
median_u_0_u_D_with_T
```

Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, t_func] = [16, None, None, 20, 0.5, 1.5, 10, None]

[thre_range, sd_R_range, day_range, penalty_range]: [[100, 105, 110, 115], [15], [3, 6, 9], [[0.0001, 5e-05], [0.0001, 5e-05]]]

[pattern_seed, day, sd_R] = [2, 3, 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:52.766
0_threshold = 100
MC for this TARGET:[58.177, 0.301]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.74, -0.84, -1.49]][[-1.69, -58.18, -5.41]]
std:[1.36, 1.35, 1.03][[0.92, 0.0, 0.38]]
MSE:[1.55, 1.59, 1.81][[1.92, 58.18, 5.42]]
MSE(-DR):[[0.0, 0.04, 0.26]][[0.37, 56.63, 3.87]]
```

```
***
=====
0_threshold = 105
MC for this TARGET:[57.733, 0.296]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.7, -2.79, -3.27]][[-4.45, -57.73, -4.97]]
std:[1.53, 1.56, 0.96][[0.78, 0.0, 0.38]]
MSE:[3.1, 3.2, 3.41][[4.52, 57.73, 4.98]]
MSE(-DR):[[0.0, 0.1, 0.31]][[1.42, 54.63, 1.88]]
***
```

```
=====
0_threshold = 110
MC for this TARGET:[56.725, 0.297]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.6, -2.62, -3.4]][[-5.62, -56.72, -3.96]]
std:[1.82, 1.81, 0.89][[0.81, 0.0, 0.38]]
MSE:[3.17, 3.18, 3.51][[5.68, 56.72, 3.98]]
MSE(-DR):[[0.0, 0.01, 0.34]][[2.51, 53.55, 0.81]]
***
```

```

=====
0_threshold = 115
MC for this TARGET:[58.664, 0.3]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.77, -6.78, -7.88]][[-13.48, -58.66, -5.9]]
std:[2.28, 2.25, 1.49]][[0.72, 0.0, 0.38]]
MSE:[7.14, 7.14, 8.02]][[13.5, 58.66, 5.91]]
MSE(-DR):[[0.0, 0.0, 0.88]][[6.36, 51.52, -1.23]]
***
=====
***** THIS SETTING IS GOOD *****
[[ 1.55  1.59  1.81  1.92 58.18  5.42]
 [ 3.1   3.2   3.41  4.52 57.73  4.98]
 [ 3.17  3.18  3.51  5.68 56.72  3.98]
 [ 7.14  7.14  8.02 13.5  58.66  5.91]]

time spent until now: 33.2 mins

20:22, 04/12

-----
[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:52.876
0_threshold = 100
MC for this TARGET:[58.156, 0.207]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.45, -0.5, -1.12]][[-1.39, -58.16, -5.28]]
std:[0.69, 0.67, 0.63]][[0.49, 0.0, 0.33]]
MSE:[0.82, 0.84, 1.29]][[1.47, 58.16, 5.29]]
MSE(-DR):[[0.0, 0.02, 0.47]][[0.65, 57.34, 4.47]]
***
=====
0_threshold = 105
MC for this TARGET:[57.712, 0.199]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.27, -2.32, -3.04]][[-4.17, -57.71, -4.84]]
std:[0.94, 0.96, 0.69]][[0.51, 0.0, 0.33]]
MSE:[2.46, 2.51, 3.12]][[4.2, 57.71, 4.85]]
MSE(-DR):[[0.0, 0.05, 0.66]][[1.74, 55.25, 2.39]]
***
=====
0_threshold = 110
MC for this TARGET:[56.703, 0.195]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.14, -2.1, -3.01]][[-5.3, -56.7, -3.83]]
std:[1.14, 1.15, 0.58]][[0.53, 0.0, 0.33]]
MSE:[2.42, 2.39, 3.07]][[5.33, 56.7, 3.84]]
MSE(-DR):[[0.0, -0.03, 0.65]][[2.91, 54.28, 1.42]]
***
=====
0_threshold = 115
MC for this TARGET:[58.647, 0.194]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-6.95, -6.87, -7.79]][[-13.15, -58.65, -5.77]]
std:[1.34, 1.29, 0.92]][[0.57, 0.0, 0.33]]
MSE:[7.08, 6.99, 7.84]][[13.16, 58.65, 5.78]]
MSE(-DR):[[0.0, -0.09, 0.76]][[6.08, 51.57, -1.3]]
***
=====
***** THIS SETTING IS GOOD *****
[[ 1.55  1.59  1.81  1.92 58.18  5.42]
 [ 3.1   3.2   3.41  4.52 57.73  4.98]
 [ 3.17  3.18  3.51  5.68 56.72  3.98]
 [ 7.14  7.14  8.02 13.5  58.66  5.91]]

[[ 0.82  0.84  1.29  1.47 58.16  5.29]
 [ 2.46  2.51  3.12  4.2  57.71  4.85]
 [ 2.42  2.39  3.07  5.33 56.7   3.84]
 [ 7.08  6.99  7.84 13.16 58.65  5.78]]

```

time spent until now: 70.1 mins

20:59, 04/12

[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:52.875

0_threshold = 100
MC for this TARGET:[58.161, 0.158]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.54, -0.65, -1.14]][[-1.51, -58.16, -5.29]]
std:[1.03, 0.99, 0.61][0.45, 0.0, 0.19]
MSE:[1.16, 1.18, 1.29][1.58, 58.16, 5.29]
MSE(-DR):[0.0, 0.02, 0.13][0.42, 57.0, 4.13]

=====

0_threshold = 105
MC for this TARGET:[57.712, 0.153]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.55, -2.64, -3.14]][[-4.29, -57.71, -4.84]]
std:[1.37, 1.39, 0.63][0.4, 0.0, 0.19]
MSE:[2.89, 2.98, 3.2][4.31, 57.71, 4.84]
MSE(-DR):[0.0, 0.09, 0.31][1.42, 54.82, 1.95]

=====

0_threshold = 110
MC for this TARGET:[56.703, 0.151]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.74, -2.79, -3.04]][[-5.43, -56.7, -3.83]]
std:[1.27, 1.3, 0.62][0.41, 0.0, 0.19]
MSE:[3.02, 3.08, 3.1][5.45, 56.7, 3.83]
MSE(-DR):[0.0, 0.06, 0.08][2.43, 53.68, 0.81]

=====

0_threshold = 115
MC for this TARGET:[58.659, 0.158]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-8.4, -8.41, -7.91]][[-13.47, -58.66, -5.78]]
std:[0.98, 1.02, 0.71][0.41, 0.0, 0.19]
MSE:[8.46, 8.47, 7.94][13.48, 58.66, 5.78]
MSE(-DR):[0.0, 0.01, -0.52][5.02, 50.2, -2.68]
**

=====

[1.55	1.59	1.81	1.92	58.18	5.42]
[3.1	3.2	3.41	4.52	57.73	4.98]
[3.17	3.18	3.51	5.68	56.72	3.98]
[7.14	7.14	8.02	13.5	58.66	5.91]

[0.82	0.84	1.29	1.47	58.16	5.29]
[2.46	2.51	3.12	4.2	57.71	4.85]
[2.42	2.39	3.07	5.33	56.7	3.84]
[7.08	6.99	7.84	13.16	58.65	5.78]

[1.16	1.18	1.29	1.58	58.16	5.29]
[2.89	2.98	3.2	4.31	57.71	4.84]
[3.02	3.08	3.1	5.45	56.7	3.83]
[8.46	8.47	7.94	13.48	58.66	5.78]

time spent until now: 114.6 mins

21:43, 04/12

ubuntu@ip-172-31-7-151:~\$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py

22:00, 04/12; num of cores:16

median_u_0_u_D_with_T_78

Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, t_func] = [16, None, None, 20, 0.5, 1.5, 10, None]

[thre_range, sd_R_range, day_range, penalty_range]: [[100, 105, 110, 115], [15], [7, 8], [[0.0001, 5e-05], [0.0001, 5e-05]]]

```
[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:52.843
```

```
0_threshold = 100
```

```
MC for this TARGET:[58.178, 0.188]
```

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

```
bias:[[-0.19, -0.31, -1.17]][[-1.37, -58.18, -5.33]]
```

```
std:[[0.74, 0.76, 0.56]][[0.31, 0.0, 0.21]]
```

```
MSE:[[0.76, 0.82, 1.3]][[1.4, 58.18, 5.33]]
```

```
MSE(-DR):[[0.0, 0.06, 0.54]][[0.64, 57.42, 4.57]]
```

```
***
```

```
=====
0_threshold = 105
```

```
MC for this TARGET:[57.732, 0.182]
```

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

```
bias:[[-1.93, -2.04, -3.17]][[-4.24, -57.73, -4.89]]
```

```
std:[[0.91, 0.91, 0.57]][[0.33, 0.0, 0.21]]
```

```
MSE:[[2.13, 2.23, 3.22]][[4.25, 57.73, 4.89]]
```

```
MSE(-DR):[[0.0, 0.1, 1.09]][[2.12, 55.6, 2.76]]
```

```
***
```

```
=====
0_threshold = 110
```

```
MC for this TARGET:[56.721, 0.18]
```

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

```
bias:[[-2.01, -2.08, -2.99]][[-5.33, -56.72, -3.88]]
```

```
std:[[0.72, 0.71, 0.65]][[0.37, 0.0, 0.21]]
```

```
MSE:[[2.14, 2.2, 3.06]][[5.34, 56.72, 3.89]]
```

```
MSE(-DR):[[0.0, 0.06, 0.92]][[3.2, 54.58, 1.75]]
```

```
***
```

```
=====
0_threshold = 115
```

```
MC for this TARGET:[58.671, 0.183]
```

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

```
bias:[[-7.21, -7.23, -7.52]][[-13.25, -58.67, -5.83]]
```

```
std:[[1.03, 0.98, 0.7]][[0.38, 0.0, 0.21]]
```

```
MSE:[[7.28, 7.3, 7.55]][[13.26, 58.67, 5.83]]
```

```
MSE(-DR):[[0.0, 0.02, 0.27]][[5.98, 51.39, -1.45]]
```

```
***
```

```
=====
***** THIS SETTING IS GOOD *****
```

```
[[ 0.76 0.82 1.3 1.4 58.18 5.33]
```

```
 [ 2.13 2.23 3.22 4.25 57.73 4.89]
```

```
 [ 2.14 2.2 3.06 5.34 56.72 3.89]
```

```
 [ 7.28 7.3 7.55 13.26 58.67 5.83]]
```

```
time spent until now: 38.8 mins
```

```
22:39, 04/12
```

```
-----
[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!
```

```
Connection to ec2-35-170-52-91.compute-1.amazonaws.com closed by remote host.
```

```
Connection to ec2-35-170-52-91.compute-1.amazonaws.com closed.
```

```
Run-Mac:~$ ssh mac$
```