

```
Last login: Sat Apr 11 10:42:07 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~ mac$ ssh -i "Runzhe.pem" ubuntu@ec2-18-209-241-162.compute-1.amazonaws.com
The authenticity of host 'ec2-18-209-241-162.compute-1.amazonaws.com (18.209.241.162)' can't be established.
ECDSA key fingerprint is SHA256:GtI8K0FlBMmSKcvc46ZE6pbllXPSi6czjWQGVHzEov4.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-18-209-241-162.compute-1.amazonaws.com,18.209.241.162' (ECDSA) to the list of known hosts.
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 Sat Apr 11 23:26:20 UTC 2020

```
System load:  0.88      Processes:      225
Usage of /:   28.0% of 30.96GB   Users logged in:  0
Memory usage: 1%      IP address for ens5: 172.31.13.166
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
```

```
89 packages can be updated.
39 updates are security updates.
```

```
Last login: Fri Apr  3 19:45:17 2020 from 107.13.161.147
export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
ubuntu@ip-172-31-13-166:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:28, 04/11; num of cores:16
large_A
```

```
Basic setting:[rep_times, sd_0, sd_D, sd_u_0, w_0, w_A, u_0_u_D, sd_R_range, t_func] = [16, None, None, 20, 1, 2, 0, [0, 10, 20], None]
```

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

```
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:51.225
0_threshold = 100
MC for this TARGET:[57.031, 0.083]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.03, -0.08, -1.47]][[-0.3, -57.03, -5.81]]
std:[[0.46, 0.47, 0.42]][[0.38, 0.0, 0.25]]
MSE:[[0.46, 0.48, 1.53]][[0.48, 57.03, 5.82]]
MSE(-DR):[[0.0, 0.02, 1.07]][[0.02, 56.57, 5.36]]
***
=====
0_threshold = 105
MC for this TARGET:[57.121, 0.067]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.55, -2.63, -3.87]][[-3.31, -57.12, -5.9]]
std:[[0.59, 0.59, 0.45]][[0.38, 0.0, 0.25]]
MSE:[[2.62, 2.7, 3.9]][[3.33, 57.12, 5.91]]
MSE(-DR):[[0.0, 0.08, 1.28]][[0.71, 54.5, 3.29]]
***
=====
0_threshold = 110
MC for this TARGET:[56.119, 0.059]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.66, -2.71, -3.86]][[-4.35, -56.12, -4.89]]
std:[[0.62, 0.59, 0.55]][[0.39, 0.0, 0.25]]
```

```

MSE:[2.73, 2.77, 3.9]][[4.37, 56.12, 4.9]]
MSE(-DR):[[0.0, 0.04, 1.17]][[1.64, 53.39, 2.17]]
***
=====
0_threshold = 115
MC for this TARGET:[59.529, 0.055]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-9.78, -9.77, -10.34]][[-13.41, -59.53, -8.3]]
std:[[0.68, 0.67, 0.54]][[0.36, 0.0, 0.25]]
MSE:[9.8, 9.79, 10.35]][[13.41, 59.53, 8.3]]
MSE(-DR):[[0.0, -0.01, 0.55]][[3.61, 49.73, -1.5]]
***
=====
***** THIS SETTING IS GOOD *****
[[ 0.46  0.48  1.53  0.48 57.03  5.82]
 [ 2.62  2.7   3.9   3.33 57.12  5.91]
 [ 2.73  2.77  3.9   4.37 56.12  4.9 ]
 [ 9.8   9.79 10.35 13.41 59.53  8.3 ]]

time spent until now: 41.5 mins

20:10, 04/11

-----
[pattern_seed, day, sd_R] = [2, 7, 10]

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:51.21
0_threshold = 100
MC for this TARGET:[57.046, 0.138]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.08, -0.03, -1.52]][[-0.33, -57.05, -5.84]]
std:[[0.57, 0.58, 0.51]][[0.34, 0.0, 0.22]]
MSE:[0.58, 0.58, 1.6]][[0.47, 57.05, 5.84]]
MSE(-DR):[[0.0, 0.0, 1.02]][[-0.11, 56.47, 5.26]]
=====
0_threshold = 105
MC for this TARGET:[57.137, 0.127]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.45, -2.51, -3.96]][[-3.36, -57.14, -5.93]]
std:[[0.61, 0.61, 0.51]][[0.38, 0.0, 0.22]]
MSE:[2.52, 2.58, 3.99]][[3.38, 57.14, 5.93]]
MSE(-DR):[[0.0, 0.06, 1.47]][[0.86, 54.62, 3.41]]
***
=====
0_threshold = 110
MC for this TARGET:[56.135, 0.125]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.63, -2.67, -3.89]][[-4.38, -56.14, -4.92]]
std:[[0.56, 0.53, 0.62]][[0.41, 0.0, 0.22]]
MSE:[2.69, 2.72, 3.94]][[4.4, 56.14, 4.92]]
MSE(-DR):[[0.0, 0.03, 1.25]][[1.71, 53.45, 2.23]]
***
=====
0_threshold = 115
MC for this TARGET:[59.545, 0.13]
[DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-9.88, -9.85, -10.38]][[-13.44, -59.54, -8.33]]
std:[[0.88, 0.83, 0.63]][[0.38, 0.0, 0.22]]
MSE:[9.92, 9.88, 10.4]][[13.45, 59.54, 8.33]]
MSE(-DR):[[0.0, -0.04, 0.48]][[3.53, 49.62, -1.59]]
***
=====
[[ 0.46  0.48  1.53  0.48 57.03  5.82]
 [ 2.62  2.7   3.9   3.33 57.12  5.91]
 [ 2.73  2.77  3.9   4.37 56.12  4.9 ]
 [ 9.8   9.79 10.35 13.41 59.53  8.3 ]]

[[ 0.58  0.58  1.6   0.47 57.05  5.84]
 [ 2.52  2.58  3.99  3.38 57.14  5.93]
 [ 2.69  2.72  3.94  4.4  56.14  4.92]
 [ 9.92  9.88 10.4  13.45 59.54  8.33]]

```

time spent until now: 83.0 mins

20:51, 04/11

[pattern_seed, day, sd_R] = [2, 7, 20]

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