

```
Last login: Thu Apr 16 09:57:11 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:~ mac$ ssh -i "Runzhe.pem" ubuntu@ec2-34-201-49-219.compute-1.amazonaws.com
The authenticity of host 'ec2-34-201-49-219.compute-1.amazonaws.com (34.201.49.219)' can't be established.
ECDSA key fingerprint is SHA256:Kn4FsTU85RdKkJt0A0WpE81rGyh286NjZF7DXq0pW4Y.
Are you sure you want to continue connecting (yes/no)? yes
Warning: Permanently added 'ec2-34-201-49-219.compute-1.amazonaws.com,34.201.49.219' (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 Thu Apr 16 14:00:12 UTC 2020

```
System load:  0.87          Processes:      837
Usage of /:   28.0% of 30.96GB Users logged in: 0
Memory usage: 0%          IP address for ens5: 172.31.0.227
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/
```

```
Last login: Fri Apr  3 19:45:17 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
10:03, 04/16; num of cores:96
sd_u_0_20_full_sigma
```

```
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, [20, 10], None]
```

```
[thre_range, sd_R_range, day_range, penalty_range]: [[100, 100.5, 105, 110, 110.5, 111], [0, 15, 30, 45], [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, 20]
```

```
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:49.429
```

```

0_threshold = 100
MC for this TARGET:[54.842, 0.069]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.45, -1.62, -1.87]][[-2.91, -54.84, -5.41]]
std:[[0.64, 0.61, 0.42]][[0.3, 0.0, 0.2]]
MSE:[[1.58, 1.73, 1.92]][[2.93, 54.84, 5.41]]
MSE(-DR):[[0.0, 0.15, 0.34]][[1.35, 53.26, 3.83]]
***
=====
0_threshold = 100.5
MC for this TARGET:[52.113, 0.068]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.04, -0.09, -0.66]][[-2.08, -52.11, -2.68]]
std:[[0.52, 0.49, 0.43]][[0.29, 0.0, 0.2]]
MSE:[[0.52, 0.5, 0.79]][[2.1, 52.11, 2.69]]
MSE(-DR):[[0.0, -0.02, 0.27]][[1.58, 51.59, 2.17]]
***
=====
0_threshold = 105
MC for this TARGET:[53.573, 0.07]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.18, -3.33, -3.54]][[-5.34, -53.57, -4.14]]
std:[[0.71, 0.7, 0.42]][[0.32, 0.0, 0.2]]
MSE:[[3.26, 3.4, 3.56]][[5.35, 53.57, 4.14]]
MSE(-DR):[[0.0, 0.14, 0.3]][[2.09, 50.31, 0.88]]
***
=====
0_threshold = 110
MC for this TARGET:[52.191, 0.064]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.66, -2.75, -3.27]][[-6.0, -52.19, -2.76]]
std:[[0.61, 0.6, 0.46]][[0.31, 0.0, 0.2]]
MSE:[[2.73, 2.81, 3.3]][[6.01, 52.19, 2.77]]
MSE(-DR):[[0.0, 0.08, 0.57]][[3.28, 49.46, 0.04]]
***
=====
0_threshold = 110.5
MC for this TARGET:[54.58, 0.058]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-5.41, -5.5, -5.66]][[-10.33, -54.58, -5.15]]
std:[[0.61, 0.57, 0.5]][[0.29, 0.0, 0.2]]
MSE:[[5.44, 5.53, 5.68]][[10.33, 54.58, 5.15]]
MSE(-DR):[[0.0, 0.09, 0.24]][[4.89, 49.14, -0.29]]
***
=====
0_threshold = 111
MC for this TARGET:[52.528, 0.053]
  [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-4.99, -5.02, -5.37]][[-10.31, -52.53, -3.11]]
std:[[0.56, 0.55, 0.48]][[0.26, 0.0, 0.2]]
MSE:[[5.02, 5.05, 5.39]][[10.31, 52.53, 3.11]]
MSE(-DR):[[0.0, 0.03, 0.37]][[5.29, 47.51, -1.91]]
***
=====
***** THIS SETTING IS GOOD *****
[[ 1.58  1.73  1.92  2.93 54.84  5.41]
 [ 0.52  0.5   0.79  2.1  52.11  2.69]
 [ 3.26  3.4   3.56  5.35 53.57  4.14]
 [ 2.73  2.81  3.3   6.01 52.19  2.77]
 [ 5.44  5.53  5.68 10.33 54.58  5.15]
 [ 5.02  5.05  5.39 10.31 52.53  3.11]]

```

time spent until now: 20.8 mins

10:24, 04/16

[pattern_seed, day, sd_R, u_0_u_D] = [2, 7, 15, 20]

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

$O_{\text{threshold}} = 100.5$

number of reward locations: 8

$O_{\text{threshold}} = 105$

number of reward locations: 7

$O_{\text{threshold}} = 110$

number of reward locations: 6

$O_{\text{threshold}} = 110.5$

number of reward locations: 5

$O_{\text{threshold}} = 111$

number of reward locations: 4