```
Last login: Thu Apr  9 13:29:52 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh mac$ ssh -i "Runzhe.pem" ubuntu@ec2-34-200-218-20.compute-1.amazonaws.com
Warning: Permanently added the ED25519 host key for IP address '34.200.218.20' 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  9 20:31:44 UTC 2020

  System load:  1.49               Processes:           387
  Usage of /:   28.0% of 30.96GB   Users logged in:     0
  Memory usage: 0%                 IP address for ens5: 172.31.8.0
  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
ubuntu@ip-172-31-8-0:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
Traceback (most recent call last):
  File "EC2.py", line 4, in <module>
    rep_times = n_cores
NameError: name 'n_cores' is not defined
ubuntu@ip-172-31-8-0:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
16:34, 04/09; num of cores:36

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, sd_R_range, t_func] = [36, None, None, 30, 0.5, 1, 0, [0, 10, 20, 30],
None]

Traceback (most recent call last):
  File "EC2.py", line 73, in <module>
    for day in day_range:
TypeError: 'int' object is not iterable
ubuntu@ip-172-31-8-0:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
16:34, 04/09; num of cores:36

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, sd_R_range, t_func] = [36, None, None, 30, 0.5, 1, 0, [0, 10, 20, 30],
None]


--------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 0]

max(u_O) =  168.8
O_threshold = 70
number of reward locations:  20
O_threshold = 80
number of reward locations:  15
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  6
target 1 in 5 DONE!
target 2 in 5 DONE!
target 3 in 5 DONE!
target 4 in 5 DONE!
target 5 in 5 DONE!

--------------------------------------
Value of Behaviour policy:61.656
O_threshold = 70
MC for this TARGET:[74.07, 0.058]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.23, 1.0, -0.06]][[1.49, -74.07, -12.41]]
std:[[0.43, 0.44, 0.24]][[0.17, 0.0, 0.13]]
MSE:[[1.3, 1.09, 0.25]][[1.5, 74.07, 12.41]]
MSE(-DR):[[0.0, -0.21, -1.05]][[0.2, 72.77, 11.11]]
**
==============
```

```
O_threshold = 80
MC for this TARGET:[73.429, 0.065]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.23, 0.0, -1.31]][[0.47, -73.43, -11.77]]
std:[[0.35, 0.37, 0.22]][[0.17, 0.0, 0.13]]
MSE:[[0.42, 0.37, 1.33]][[0.5, 73.43, 11.77]]
MSE(-DR):[[0.0, -0.05, 0.91]][[0.08, 73.01, 11.35]]
***
==============
O_threshold = 90
MC for this TARGET:[70.82, 0.07]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[0.21, 0.04, -0.79]][[-0.14, -70.82, -9.16]]
std:[[0.26, 0.26, 0.2]][[0.16, 0.0, 0.13]]
MSE:[[0.33, 0.26, 0.81]][[0.21, 70.82, 9.16]]
MSE(-DR):[[0.0, -0.07, 0.48]][[-0.12, 70.49, 8.83]]
==============
O_threshold = 100
MC for this TARGET:[70.166, 0.071]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.01, -1.16, -1.46]][[-2.17, -70.17, -8.51]]
std:[[0.28, 0.29, 0.2]][[0.17, 0.0, 0.13]]
MSE:[[1.05, 1.2, 1.47]][[2.18, 70.17, 8.51]]
MSE(-DR):[[0.0, 0.15, 0.42]][[1.13, 69.12, 7.46]]
***
==============
O_threshold = 110
MC for this TARGET:[69.509, 0.063]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-3.26, -3.36, -3.48]][[-5.07, -69.51, -7.85]]
std:[[0.4, 0.4, 0.21]][[0.16, 0.0, 0.13]]
MSE:[[3.28, 3.38, 3.49]][[5.07, 69.51, 7.85]]
MSE(-DR):[[0.0, 0.1, 0.21]][[1.79, 66.23, 4.57]]
***
==============
[[ 1.3   1.09  0.25  1.5  74.07 12.41]
 [ 0.42  0.37  1.33  0.5  73.43 11.77]
 [ 0.33  0.26  0.81  0.21 70.82  9.16]
 [ 1.05  1.2   1.47  2.18 70.17  8.51]
 [ 3.28  3.38  3.49  5.07 69.51  7.85]]


time spent until now: 125.6 mins

18:40, 04/09

--------------------------------------
[pattern_seed, day, sd_R] = [2, 14, 10]

max(u_O) =  168.8
O_threshold = 70
number of reward locations:  20
O_threshold = 80
number of reward locations:  15
O_threshold = 90
number of reward locations:  12
O_threshold = 100
number of reward locations:  9
O_threshold = 110
number of reward locations:  6
target 1 in 5 DONE!
^CProcess Process-59:
Process Process-39:
Process Process-43:
Process Process-53:
Traceback (most recent call last):
  File "EC2.py", line 90, in <module>
Process Process-68:
Process Process-40:
Process Process-48:
Process Process-67:
Process Process-63:
    dim_S_plus_Ts = 3 + 3, epsilon = 1e-6, # Fixed
  File "/home/ubuntu/simu_funs.py", line 64, in simu
Process Process-55:
    value_reps = parmap(once, range(OPE_rep_times), n_cores)
  File "/home/ubuntu/_uti_basic.py", line 80, in parmap
Process Process-62:
    [q_in.put((None, None)) for _ in range(nprocs)]
  File "/home/ubuntu/_uti_basic.py", line 80, in <listcomp>
    [q_in.put((None, None)) for _ in range(nprocs)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
Process Process-64:
Process Process-61:
Process Process-46:
Process Process-72:
Process Process-50:
Process Process-54:
Process Process-60:
```

```
Process Process-57:
Process Process-56:
Process Process-38:
Process Process-37:
Process Process-51:
Process Process-71:
Process Process-52:
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-8-0:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:17, 04/09; num of cores:36

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, sd_R_range, t_func] = [16, None, None, 20, 0.5, 1, 0, [0, 10, 20, 30, 40], None]


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

max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
O_threshold = 115
number of reward locations:  3
target 1 in 3 DONE!
^CProcess Process-53:
Process Process-43:
Process Process-51:
Process Process-59:
Process Process-55:
Process Process-47:
Process Process-49:
Process Process-61:
Process Process-44:
Process Process-57:
Process Process-60:
Process Process-37:
Process Process-45:
Process Process-72:
Process Process-48:
Traceback (most recent call last):
Process Process-46:
  File "EC2.py", line 91, in <module>
    with_MF = with_MF, with_NO_MARL = with_NO_MARL, with_IS = with_IS,
  File "/home/ubuntu/simu_funs.py", line 66, in simu
    value_reps = rep_seeds(once, OPE_rep_times)
  File "/home/ubuntu/_uti_basic.py", line 124, in rep_seeds
    return list(map(fun, range(rep_times)))
  File "/home/ubuntu/simu_funs.py", line 62, in once
Traceback (most recent call last):
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 214, in simu_once
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 99, in run
    self._target(*self._args, **self._kwargs)
  File "/home/ubuntu/_uti_basic.py", line 67, in fun
    q_out.put((i, f(x)))
  File "/home/ubuntu/main.py", line 86, in getOneRegionValue
    epsilon = epsilon)
  File "/home/ubuntu/main.py", line 262, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 196, in train
    pi0 = den_b_disc(Ta_tl, n_neigh) * 0.5
  File "/home/ubuntu/_utility.py", line 56, in den_b_disc
    den += binom.pmf(i, N_neigh, 0.5)
KeyboardInterrupt
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 156, in V_DR
    r = arr(parmap(getOneRegionValue, range(N)))
  File "/home/ubuntu/_uti_basic.py", line 80, in parmap
    [q_in.put((None, None)) for _ in range(nprocs)]
  File "/home/ubuntu/_uti_basic.py", line 80, in <listcomp>
    [q_in.put((None, None)) for _ in range(nprocs)]
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/queues.py", line 82, in put
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-8-0:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:18, 04/09; num of cores:36

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, sd_R_range, t_func] = [16, None, None, 20, 0.5, 1, 0, [0, 10, 20, 30, 40], None]


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

```
max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
target 1 in 2 DONE!
target 2 in 2 DONE!
Rep 5 DONE
Rep 10 DONE
Rep 15 DONE

---------------------------------------
Value of Behaviour policy:64.845
O_threshold = 95
MC for this TARGET:[71.023, 0.087]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.05, 0.93, -0.02]][[1.09, -71.02, -6.18]]
std:[[0.33, 0.34, 0.28]][[0.28, 0.0, 0.2]]
MSE:[[1.1, 0.99, 0.28]][[1.13, 71.02, 6.18]]
MSE(-DR):[[0.0, -0.11, -0.82]][[0.03, 69.92, 5.08]]
**
==============
O_threshold = 105
MC for this TARGET:[71.779, 0.092]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.59, -2.66, -3.34]][[-3.99, -71.78, -6.93]]
std:[[0.5, 0.48, 0.36]][[0.32, 0.0, 0.2]]
MSE:[[2.64, 2.7, 3.36]][[4.0, 71.78, 6.93]]
MSE(-DR):[[0.0, 0.06, 0.72]][[1.36, 69.14, 4.29]]
***
==============
[[ 1.1   0.99  0.28  1.13 71.02  6.18]
 [ 2.64  2.7   3.36  4.   71.78  6.93]]


time spent until now: 11.3 mins

19:29, 04/09

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

max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
target 1 in 2 DONE!
target 2 in 2 DONE!
Rep 5 DONE
Rep 10 DONE
Rep 15 DONE

---------------------------------------
Value of Behaviour policy:64.831
O_threshold = 95
MC for this TARGET:[71.038, 0.144]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.19, 1.07, 0.05]][[1.13, -71.04, -6.21]]
std:[[0.42, 0.41, 0.31]][[0.33, 0.0, 0.15]]
MSE:[[1.26, 1.15, 0.31]][[1.18, 71.04, 6.21]]
MSE(-DR):[[0.0, -0.11, -0.95]][[-0.08, 69.78, 4.95]]
==============
O_threshold = 105
MC for this TARGET:[71.795, 0.146]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.41, -2.46, -3.44]][[-4.06, -71.8, -6.96]]
std:[[0.68, 0.69, 0.39]][[0.29, 0.0, 0.15]]
MSE:[[2.5, 2.55, 3.46]][[4.07, 71.8, 6.96]]
MSE(-DR):[[0.0, 0.05, 0.96]][[1.57, 69.3, 4.46]]
***
==============
[[ 1.1   0.99  0.28  1.13 71.02  6.18]
 [ 2.64  2.7   3.36  4.   71.78  6.93]]


[[ 1.26  1.15  0.31  1.18 71.04  6.21]
 [ 2.5   2.55  3.46  4.07 71.8   6.96]]


time spent until now: 22.6 mins

19:40, 04/09

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

max(u_O) =  145.8
O_threshold = 95
```

```
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
target 1 in 2 DONE!
target 2 in 2 DONE!
Rep 5 DONE
Rep 10 DONE
Rep 15 DONE

--------------------------------------
Value of Behaviour policy:64.817
O_threshold = 95
MC for this TARGET:[71.054, 0.246]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.31, 1.21, 0.11]][[1.17, -71.05, -6.24]]
std:[[0.72, 0.69, 0.49]][[0.48, 0.0, 0.18]]
MSE:[[1.49, 1.39, 0.5]][[1.26, 71.05, 6.24]]
MSE(-DR):[[0.0, -0.1, -0.99]][[-0.23, 69.56, 4.75]]
==============
O_threshold = 105
MC for this TARGET:[71.811, 0.247]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.17, -2.28, -3.47]][[-4.12, -71.81, -6.99]]
std:[[1.21, 1.26, 0.59]][[0.4, 0.0, 0.18]]
MSE:[[2.48, 2.6, 3.52]][[4.14, 71.81, 6.99]]
MSE(-DR):[[0.0, 0.12, 1.04]][[1.66, 69.33, 4.51]]
***
==============
[[ 1.1   0.99  0.28  1.13 71.02  6.18]
 [ 2.64  2.7   3.36  4.   71.78  6.93]]


[[ 1.26  1.15  0.31  1.18 71.04  6.21]
 [ 2.5   2.55  3.46  4.07 71.8   6.96]]


[[ 1.49  1.39  0.5   1.26 71.05  6.24]
 [ 2.48  2.6   3.52  4.14 71.81  6.99]]


time spent until now: 33.9 mins

19:52, 04/09

--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 30]

max(u_O) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
target 1 in 2 DONE!
target 2 in 2 DONE!
Rep 5 DONE
Rep 10 DONE
Rep 15 DONE

--------------------------------------
Value of Behaviour policy:64.802
O_threshold = 95
MC for this TARGET:[71.07, 0.357]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.44, 1.35, 0.2]][[1.21, -71.07, -6.27]]
std:[[1.08, 1.03, 0.68]][[0.67, 0.0, 0.26]]
MSE:[[1.8, 1.7, 0.71]][[1.38, 71.07, 6.28]]
MSE(-DR):[[0.0, -0.1, -1.09]][[-0.42, 69.27, 4.48]]
==============
O_threshold = 105
MC for this TARGET:[71.827, 0.356]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.05, -2.15, -3.54]][[-4.19, -71.83, -7.02]]
std:[[1.86, 1.89, 0.79]][[0.56, 0.0, 0.26]]
MSE:[[2.77, 2.86, 3.63]][[4.23, 71.83, 7.02]]
MSE(-DR):[[0.0, 0.09, 0.86]][[1.46, 69.06, 4.25]]
***
==============
[[ 1.1   0.99  0.28  1.13 71.02  6.18]
 [ 2.64  2.7   3.36  4.   71.78  6.93]]


[[ 1.26  1.15  0.31  1.18 71.04  6.21]
 [ 2.5   2.55  3.46  4.07 71.8   6.96]]


[[ 1.49  1.39  0.5   1.26 71.05  6.24]
 [ 2.48  2.6   3.52  4.14 71.81  6.99]]
```

```
[[ 1.8   1.7   0.71  1.38 71.07  6.28]
 [ 2.77  2.86  3.63  4.23 71.83  7.02]]


time spent until now: 45.3 mins

20:03, 04/09

--------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 40]

max(u_0) =  145.8
O_threshold = 95
number of reward locations:  12
O_threshold = 105
number of reward locations:  7
target 1 in 2 DONE!
target 2 in 2 DONE!
Rep 5 DONE
Rep 10 DONE
Rep 15 DONE

--------------------------------------
Value of Behaviour policy:64.788
O_threshold = 95
MC for this TARGET:[71.086, 0.469]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.61, 1.49, 0.31]][[1.24, -71.09, -6.3]]
std:[[1.44, 1.38, 1.01]][[0.88, 0.0, 0.35]]
MSE:[[2.16, 2.03, 1.06]][[1.52, 71.09, 6.31]]
MSE(-DR):[[0.0, -0.13, -1.1]][[-0.64, 68.93, 4.15]]
==============
O_threshold = 105
MC for this TARGET:[71.843, 0.469]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.98, -2.06, -3.61]][[-4.27, -71.84, -7.05]]
std:[[2.48, 2.51, 1.04]][[0.76, 0.0, 0.35]]
MSE:[[3.17, 3.25, 3.76]][[4.34, 71.84, 7.06]]
MSE(-DR):[[0.0, 0.08, 0.59]][[1.17, 68.67, 3.89]]
***
==============
[[ 1.1   0.99  0.28  1.13 71.02  6.18]
 [ 2.64  2.7   3.36  4.   71.78  6.93]]


[[ 1.26  1.15  0.31  1.18 71.04  6.21]
 [ 2.5   2.55  3.46  4.07 71.8   6.96]]


[[ 1.49  1.39  0.5   1.26 71.05  6.24]
 [ 2.48  2.6   3.52  4.14 71.81  6.99]]


[[ 1.8   1.7   0.71  1.38 71.07  6.28]
 [ 2.77  2.86  3.63  4.23 71.83  7.02]]


[[ 2.16  2.03  1.06  1.52 71.09  6.31]
 [ 3.17  3.25  3.76  4.34 71.84  7.06]]


time spent until now: 56.6 mins

20:14, 04/09
ubuntu@ip-172-31-8-0:~$
```