```
Last login: Sat Apr 11 10:42:07 on ttys000
Run-Mac:~ mac$ cd ~/.ssh
Run-Mac:.ssh 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_O, sd_D, sd_u_O, w_O, w_A, u_O_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_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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
O_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]]
***
==============
O_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]]
***
==============
O_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]]
***
==============
O_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_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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
O_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, 0.5]][[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]]
==============
O_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]]
***
==============
O_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]]
***
==============
O_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]]
```

```
    feed_dict_tensor, options, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
    options, feed_dict, fetch_list, target_list, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
    options, feed_dict, fetch_list, target_list, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1158, in _run
    self._graph, fetches, feed_dict_tensor, feed_handles=feed_handles)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1350, in _do_run
    run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
KeyboardInterrupt
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
    return fn(*args)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 474, in __init__
    self._fetch_mapper = _FetchMapper.for_fetch(fetches)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1356, in _do_call
    return fn(*args)
KeyboardInterrupt
KeyboardInterrupt
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
    options, feed_dict, fetch_list, target_list, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 264, in for_fetch
    return _ListFetchMapper(fetch)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1341, in _run_fn
    options, feed_dict, fetch_list, target_list, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 374, in __init__
    self._unique_fetches, self._value_indices = _uniquify_fetches(self._mappers)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_sessionrun
    run_metadata)
KeyboardInterrupt
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 357, in _uniquify_fetches
    unique_fetches.append(f)
KeyboardInterrupt
KeyboardInterrupt
Traceback (most recent call last):
  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/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 79, in getOneRegionValue
    CV_QV = CV_QV, penalty_range = penalty, spatial = True)
  File "/home/ubuntu/main.py", line 302, in computeQV
    validation_set = valid_tuples)
  File "/home/ubuntu/main.py", line 432, in computeQV_basic
    alpha_eta = np.linalg.lstsq(left, np.expand_dims(right,1))[0]
  File "/home/ubuntu/anaconda3/lib/python3.7/site-packages/numpy/linalg/linalg.py", line 2236, in lstsq
    x, resids, rank, s = gufunc(a, b, rcond, signature=signature, extobj=extobj)
KeyboardInterrupt
    if not self._sem.acquire(block, timeout):
KeyboardInterrupt
ubuntu@ip-172-31-13-166:~$ python simu0.py
python: can't open file 'simu0.py': [Errno 2] No such file or directory
ubuntu@ip-172-31-13-166:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
21:01, 04/11; num of cores:16
large_A

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, 1, 2, 0, [0, 10, 20], None]


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

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
```

```
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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:53.65
O_threshold = 100
MC for this TARGET:[53.709, 0.079]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[4.25, 4.16, 3.34]][[4.66, -53.71, -0.06]]
std:[[0.52, 0.51, 0.41]][[0.3, 0.0, 0.25]]
MSE:[[4.28, 4.19, 3.37]][[4.67, 53.71, 0.26]]
MSE(-DR):[[0.0, -0.09, -0.91]][[0.39, 49.43, -4.02]]
**
==============
O_threshold = 105
MC for this TARGET:[53.895, 0.067]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.55, 1.47, 1.11]][[1.76, -53.9, -0.25]]
std:[[0.46, 0.45, 0.43]][[0.27, 0.0, 0.25]]
MSE:[[1.62, 1.54, 1.19]][[1.78, 53.9, 0.35]]
MSE(-DR):[[0.0, -0.08, -0.43]][[0.16, 52.28, -1.27]]
**
==============
O_threshold = 110
MC for this TARGET:[54.39, 0.06]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.36, -0.43, -0.46]][[-0.43, -54.39, -0.74]]
std:[[0.53, 0.53, 0.53]][[0.3, 0.0, 0.25]]
MSE:[[0.64, 0.68, 0.7]][[0.52, 54.39, 0.78]]
MSE(-DR):[[0.0, 0.04, 0.06]][[-0.12, 53.75, 0.14]]
==============
O_threshold = 115
MC for this TARGET:[60.623, 0.057]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-10.21, -10.22, -9.15]][[-11.86, -60.62, -6.97]]
std:[[0.54, 0.51, 0.52]][[0.27, 0.0, 0.25]]
MSE:[[10.22, 10.23, 9.16]][[11.86, 60.62, 6.97]]
MSE(-DR):[[0.0, 0.01, -1.06]][[1.64, 50.4, -3.25]]
**
==============
[[ 4.28  4.19  3.37  4.67 53.71  0.26]
 [ 1.62  1.54  1.19  1.78 53.9   0.35]
 [ 0.64  0.68  0.7   0.52 54.39  0.78]
 [10.22 10.23  9.16 11.86 60.62  6.97]]


time spent until now: 41.2 mins

21:42, 04/11

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

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
^CProcess Process-18:
Process Process-24:
Traceback (most recent call last):
Process Process-32:
Process Process-23:
Process Process-22:
  File "EC2.py", line 80, in <module>
    with_MF = with_MF, with_NO_MARL = with_NO_MARL, with_IS = with_IS,
  File "/home/ubuntu/simu_funs.py", line 63, in simu
Process Process-25:
    value_reps = parmap(once, range(OPE_rep_times), n_cores)
  File "/home/ubuntu/_uti_basic.py", line 80, in parmap
Process Process-27:
    [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-31:
    if not self._sem.acquire(block, timeout):
```

```
KeyboardInterrupt
Process Process-29:
Traceback (most recent call last):
Traceback (most recent call last):
Process Process-30:
  File "/home/ubuntu/anaconda3/lib/python3.7/multiprocessing/process.py", line 297, in _bootstrap
    self.run()
Traceback (most recent call last):
  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/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
  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/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 86, in getOneRegionValue
    epsilon = epsilon)
  File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1173, in _run
    feed_dict_tensor, options, run_metadata)
  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/simu_funs.py", line 61, in once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/simu_funs.py", line 213, in simu_once
    inner_parallel = inner_parallel)
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
Process Process-17:
  File "/home/ubuntu/main.py", line 157, in V_DR
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 157, in <listcomp>
    r = arr([getOneRegionValue(i) for i in range(N)])
  File "/home/ubuntu/main.py", line 86, in getOneRegionValue
    epsilon = epsilon)
  File "/home/ubuntu/main.py", line 261, in getWeight
    epsilon = epsilon, spatial = spatial, mean_field = mean_field)
  File "/home/ubuntu/weight.py", line 301, in train
    self.policy_ratio2: policy_ratio2
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 950, in run
    run_metadata_ptr)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1158, in _run
    self._graph, fetches, feed_dict_tensor, feed_handles=feed_handles)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 474, in __init__
    self._fetch_mapper = _FetchMapper.for_fetch(fetches)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 264, in for_fetch
    return _ListFetchMapper(fetch)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 373, in __init__
    self._mappers = [_FetchMapper.for_fetch(fetch) for fetch in fetches]
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 373, in <listcomp>
    self._mappers = [_FetchMapper.for_fetch(fetch) for fetch in fetches]
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 265, in for_fetch
    elif isinstance(fetch, collections.Mapping):
ubuntu@ip-172-31-13-166:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
22:00, 04/11; num of cores:16
simple_old

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, 1.5, 0, [0, 10, 20], No
ne]


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

```
max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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:58.0
O_threshold = 100
MC for this TARGET:[56.952, 0.081]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[4.85, 4.73, 4.47]][[5.13, -56.95, 1.05]]
std:[[0.57, 0.54, 0.3]][[0.26, 0.0, 0.23]]
MSE:[[4.88, 4.76, 4.48]][[5.14, 56.95, 1.07]]
MSE(-DR):[[0.0, -0.12, -0.4]][[0.26, 52.07, -3.81]]
**
==============
O_threshold = 105
MC for this TARGET:[58.216, 0.067]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.36, 1.28, 1.52]][[1.34, -58.22, -0.22]]
std:[[0.4, 0.37, 0.33]][[0.26, 0.0, 0.23]]
MSE:[[1.42, 1.33, 1.56]][[1.36, 58.22, 0.32]]
MSE(-DR):[[0.0, -0.09, 0.14]][[-0.06, 56.8, -1.1]]
==============
O_threshold = 110
MC for this TARGET:[59.025, 0.061]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.56, -0.61, -0.12]][[-1.15, -59.02, -1.03]]
std:[[0.36, 0.34, 0.4]][[0.31, 0.0, 0.23]]
MSE:[[0.67, 0.7, 0.42]][[1.19, 59.02, 1.06]]
MSE(-DR):[[0.0, 0.03, -0.25]][[0.52, 58.35, 0.39]]
**
==============
O_threshold = 115
MC for this TARGET:[64.283, 0.059]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-8.74, -8.72, -7.26]][[-11.53, -64.28, -6.28]]
std:[[0.62, 0.62, 0.44]][[0.32, 0.0, 0.23]]
MSE:[[8.76, 8.74, 7.27]][[11.53, 64.28, 6.28]]
MSE(-DR):[[0.0, -0.02, -1.49]][[2.77, 55.52, -2.48]]
**
==============
[[ 4.88  4.76  4.48  5.14 56.95  1.07]
 [ 1.42  1.33  1.56  1.36 58.22  0.32]
 [ 0.67  0.7   0.42  1.19 59.02  1.06]
 [ 8.76  8.74  7.27 11.53 64.28  6.28]]


time spent until now: 40.7 mins

22:41, 04/11

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

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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:57.985
O_threshold = 100
MC for this TARGET:[56.968, 0.135]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[4.8, 4.69, 4.32]][[5.11, -56.97, 1.02]]
std:[[0.66, 0.64, 0.33]][[0.31, 0.0, 0.22]]
MSE:[[4.85, 4.73, 4.33]][[5.12, 56.97, 1.04]]
MSE(-DR):[[0.0, -0.12, -0.52]][[0.27, 52.12, -3.81]]
**
```

```
==============
O_threshold = 105
MC for this TARGET:[58.232, 0.125]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.38, 1.31, 1.4]][[1.29, -58.23, -0.25]]
std:[[0.6, 0.55, 0.32]][[0.32, 0.0, 0.22]]
MSE:[[1.5, 1.42, 1.44]][[1.33, 58.23, 0.33]]
MSE(-DR):[[0.0, -0.08, -0.06]][[-0.17, 56.73, -1.17]]
==============
O_threshold = 110
MC for this TARGET:[59.041, 0.125]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.51, -0.56, -0.14]][[-1.17, -59.04, -1.06]]
std:[[0.57, 0.51, 0.31]][[0.32, 0.0, 0.22]]
MSE:[[0.76, 0.76, 0.34]][[1.21, 59.04, 1.08]]
MSE(-DR):[[0.0, 0.0, -0.42]][[0.45, 58.28, 0.32]]
**
==============
O_threshold = 115
MC for this TARGET:[64.298, 0.13]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-8.61, -8.58, -7.22]][[-11.53, -64.3, -6.31]]
std:[[0.87, 0.88, 0.39]][[0.31, 0.0, 0.22]]
MSE:[[8.65, 8.63, 7.23]][[11.53, 64.3, 6.31]]
MSE(-DR):[[0.0, -0.02, -1.42]][[2.88, 55.65, -2.34]]
**
==============
[[ 4.88  4.76  4.48  5.14 56.95  1.07]
 [ 1.42  1.33  1.56  1.36 58.22  0.32]
 [ 0.67  0.7   0.42  1.19 59.02  1.06]
 [ 8.76  8.74  7.27 11.53 64.28  6.28]]


[[ 4.85  4.73  4.33  5.12 56.97  1.04]
 [ 1.5   1.42  1.44  1.33 58.23  0.33]
 [ 0.76  0.76  0.34  1.21 59.04  1.08]
 [ 8.65  8.63  7.23 11.53 64.3   6.31]]


time spent until now: 80.5 mins

23:21, 04/11

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

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_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:57.971
O_threshold = 100
MC for this TARGET:[56.984, 0.238]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[4.73, 4.61, 4.18]][[5.09, -56.98, 0.99]]
std:[[1.07, 1.04, 0.61]][[0.49, 0.0, 0.26]]
MSE:[[4.85, 4.73, 4.22]][[5.11, 56.98, 1.02]]
MSE(-DR):[[0.0, -0.12, -0.63]][[0.26, 52.13, -3.83]]
**
==============
O_threshold = 105
MC for this TARGET:[58.248, 0.231]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[1.41, 1.32, 1.38]][[1.24, -58.25, -0.28]]
std:[[0.97, 0.89, 0.45]][[0.52, 0.0, 0.26]]
MSE:[[1.71, 1.59, 1.45]][[1.34, 58.25, 0.38]]
MSE(-DR):[[0.0, -0.12, -0.26]][[-0.37, 56.54, -1.33]]
==============
O_threshold = 110
MC for this TARGET:[59.056, 0.233]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.49, -0.52, -0.2]][[-1.19, -59.06, -1.08]]
std:[[0.94, 0.87, 0.52]][[0.46, 0.0, 0.26]]
MSE:[[1.06, 1.01, 0.56]][[1.28, 59.06, 1.11]]
MSE(-DR):[[0.0, -0.05, -0.5]][[0.22, 58.0, 0.05]]
**
==============
O_threshold = 115
```

```
MC for this TARGET:[64.314, 0.239]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-8.52, -8.44, -7.31]][[-11.54, -64.31, -6.34]]
std:[[1.29, 1.28, 0.63]][[0.45, 0.0, 0.26]]
MSE:[[8.62, 8.54, 7.34]][[11.55, 64.31, 6.35]]
MSE(-DR):[[0.0, -0.08, -1.28]][[2.93, 55.69, -2.27]]
**
==============
[[ 4.88  4.76  4.48  5.14 56.95  1.07]
 [ 1.42  1.33  1.56  1.36 58.22  0.32]
 [ 0.67  0.7   0.42  1.19 59.02  1.06]
 [ 8.76  8.74  7.27 11.53 64.28  6.28]]


[[ 4.85  4.73  4.33  5.12 56.97  1.04]
 [ 1.5   1.42  1.44  1.33 58.23  0.33]
 [ 0.76  0.76  0.34  1.21 59.04  1.08]
 [ 8.65  8.63  7.23 11.53 64.3   6.31]]


[[ 4.85  4.73  4.22  5.11 56.98  1.02]
 [ 1.71  1.59  1.45  1.34 58.25  0.38]
 [ 1.06  1.01  0.56  1.28 59.06  1.11]
 [ 8.62  8.54  7.34 11.55 64.31  6.35]]


time spent until now: 119.5 mins

00:00, 04/12
ubuntu@ip-172-31-13-166:~$
```