

Density histogram of TD RPE sign under the optimal policy

Stationary environments with varying reward probability p

