

```
%  
% Demonstrate the technique of reinforcement comparison  
% We compare three methods  
%  
% M1) action-value method with constant update ( $\alpha = 0.1$ ) and  $\epsilon=0.1$  greedy  
% M2) action-value method with  $1/k$  update and  $\epsilon=0.1$  greedy  
% M3) the reinforcement comparison method  
%  
% Inputs:  
%   nB: the number of bandits  
%   nA: the number of arms  
%   nP: the number of plays (times we will pull a arm)  
%   sigmaReward: the standard deviation of the return from each of the arms  
%  
% Written by:  
% --  
% John L. Weatherwax                2007-11-13  
%  
% email: wax@alum.mit.edu  
%  
% Please send comments and especially bug reports to the  
% above email address.  
%  
%-----  
  
close all;  
  
% sample_discrete.m  
addpath( genpath( '../FullBNT-1.0.4/KPMstats/' ) );  
  
nB=2000; nA=10; nP=1000; sigmaReward=1.0;  
reinforcement_comparison_methods(nB,nA,nP,sigmaReward);
```