

REINFORCE P-H-D ($\sigma=0.6$)

action hist for $p_0 \geq 0.8$

