

REINFORCE L-H-D ($\sigma=0.6$)
action hist for p0 in]0.4, 0.2]

