

REINFORCE P-T ($\sigma=0.6$)

action hist for p0 in]0.6, 0.4]

