REINFORCE L-H-D-V (σ=0.4)
action hist for p0 in]0.8, 0.6]