REINFORCE L-T-V ($\sigma=0.4$)
action hist for p0 < 0.2