Average reward using actor-critic IRL
$\sigma_x = 1, \sigma_j = 1, \mu_j^* = 5, \gamma = 0.9, \alpha_\theta = 0.001, \alpha_\upsilon = 0.0001$