← Back

Policy Gradient Methods

Practice Assignment • 45 min

1. Which of the following is true about policy-based methods? (**Select all that apply**) — 1 point

- [x] Policy-based methods can be applied to continuous action space domains.
- [x] Policy-based methods allow smooth improvement in the policy without drastic changes.
- [x] Policy-based methods are useful in problems where the policy is easier to approximate than action-value functions.
- [x] Policy-based methods can learn an optimal policy that is stochastic.

2. Which of the following statements about parameterized policies are true? (**Select all that apply**) — 1 point

- [x] The probability of selecting any action must be greater than or equal to zero.
- [x] For each state, the sum of all the action probabilities must equal to one.
- [ ] The function used for representing the policy must be a softmax function.
- [ ] The policy must be approximated using linear function approximation.

3. Assume you're given the following preferences $h_1 = 44, h_2 = 42$, and $h_3 = 38$, corresponding to three different actions (a_1, a_2, a_3), respectively. Under a softmax policy, what is the probability of choosing a_2, rounded to three decimal numbers? — 1 point

- ( ) 0.42
- ( ) 0.879
- ( ) 0.002
- (•) 0.119