

The Optimizer's Curse

Soheil

1 Main Ideas

In decision analysis, we typically identify a set of alternatives, calculate their expected values, and choose the ones with highest expected values.

When we are a decision maker, if we only choose alternatives based on the estimated values, we will often soon to be disappointed. This phenomenon is called optimizer's curse. This is not happening because of any inherent bias in the estimates, but because of the optimization based selection.

Being conditionally unbiased means:

$$\mathbb{E}[V_i | \mu_1, \dots, \mu_n] = \mu_i$$

2 Questions

1. Why is this paper relevant to reinforcement learning?

In reinforcement learning, we do the exact same thing by using bellman equations. First, we find the best expected value, then we pick it. But, as said in this paper, this method is prone to pick an alternative which almost always has a lower value than it's expected value.

2. How would you beat the optimizer's curse without using Bayesian estimates?

By splitting the samples into training and test samples. We use the training part to pick the best actions, and use the test part to evaluate our pick.

3. What are the pro's and con's of using Bayesian estimates?