CSCI 145 Problem 23: Discount a Persistent Shock

August 27, 2026

Submission Instructions

Please upload your work to Gradescope by **11:59pm on DATE**.

- Turn in a single PDF of *handwritten* solutions. Typed (including LaTeX) solutions will not be accepted.
- Problems are grouped into three-week **cycles**. Once per cycle, you will meet with the grader and present the solution to *one* problem from that cycle, selected at random when you arrive; bring your work and be ready to explain any of them.

Problem 23: Discount a Persistent Shock 2 points

The scientific target is the far-sighted value $V^\star = \mu/(1-\gamma^\star)$ with $\gamma^\star = 0.99$. In one episode the rewards follow

$$r_t = \mu + Z + \varepsilon_t,$$

where Z is an episode-wide mean-zero shock with variance τ^2, while the ε_t are independent mean-zero noises with variance σ^2 and are independent of Z. You estimate $V^\star$ using the deliberately shorter return $G_\gamma = \sum_{t=0}^{\infty} \gamma^t r_t$, with $0 \le \gamma \le \gamma^\star$.

Choose γ by minimizing $\mathbb{E}[(G_\gamma - V^\star)^2]$. Derive the objective from the stochastic model, expose the contribution that averaging over time cannot remove, and report numerical minimizers to three decimal places for

$$\mu = 0.02, \quad \sigma = 0.30, \qquad \text{(i) } \tau = 0, \quad \text{(ii) } \tau = 0.05.$$

A calculator or short numerical search may be used after the exact objective is established.

Use the two answers to assess the statement "a longer effective horizon always gives a more accurate value estimate." Explain how this problem is the temporal analogue of correlated examples in a minibatch.

Results you may use. You may use the infinite geometric-series identities and independence rules for variance. The effective horizon is approximately $1/(1-\gamma)$.