COMP-597: Reinforcement Learning - Assignment 3

Prepared by Marc Bellemare

Posted Tuesday April 5, 2022
Due April 13, 2022 - can be turned in without penalty until April 19

The assignment can be carried out individually or in teams of two. Further instructions about how to submit will be provided by the TAs

1 Transforming distributions [40 points]

1.1 Shift and scale: $\Pr(R + \gamma Z)$

Consider a random variable Z with distribution ν. Suppose that $\gamma = {}^2\!/\!_3$. Write, in closed form, the probability distribution of $\Pr(R + \gamma Z)$ in the following cases:

- Z has a normal distribution with mean 1 and variance 2 ($Z \sim \mathcal{N}(1, 2)$), and R = 0.5;
- $Z \sim \mathcal{N}(1, 2)$ and R has a Bernoulli distribution with parameter ${}^3\!/\!_4$.
- $Z \sim \mathcal{N}(1, 2)$ and $R \sim \mathcal{N}(0, 1)$?

1.2 Random return

Consider the random return

$$G = \sum_{t=0}^{\infty} \gamma^t R_t.$$

1.3 Distribution of G

What is the distribution of G when

- $R_t \sim \mathcal{N}(0, 1)$, for all t?
- R_t has uniform distribution on $[1, 2]$ ($R_t \sim \mathcal{U}([1, 2])$)?
- $R_t \sim {}^1\!/\!_3\delta_{-1} + {}^1\!/\!_3\delta_0 + {}^1\!/\!_3\delta_1$?
- $R_t \sim \mathcal{N}(1, t)$.

1.4 Return with infinite variance

Find a sequence of normally-distributed random variables $(R_t)_{t \geq 0}$ such that

$$\mathbb{E}[G] = 0 \qquad \text{Var}(G) = \infty.$$

2 Categorical dynamic programming [50 points]

For a fixed set of locations $\{\theta_1, \ldots, \theta_m\}$, the categorical projection of ν', written $\Pi_C \nu'$, can be written using triangular and half-triangular kernels. If we write

$$\Pi_C \nu' = \sum_{i=1}^{m} q_i \delta_{\theta_i},$$

then the probabilities $(q_i)_{i=1}^m$ are given by the following expectation, where $Y \sim \nu'$:

$$p_i = \mathbb{E}[h_i(Y)]$$

with $h_i = \max\{0, 1 - |z|\}$ for $i = 2, \ldots, m-1$, and

$$h_1(z) = \begin{cases} 1 & z \leq 0 \\ \max\{0, 1-z\} & z > 0 \end{cases} \qquad h_m(z) = \begin{cases} \max\{0, 1-z\} & \leq 0 \\ 1 & z > 0 \end{cases}.$$

Suppose that you are given the parameters μ and σ^2 of a normal distribution, and that you want to approximate this distribution using the categorical representation. That is, let

$$\nu = \sum_{i=1}^{m} p_i \delta_{\theta_i}.$$

Provide an algorithm which, given μ, σ^2, and $m \geq 2$ as an input, returns a set of locations $\{\theta_1, \ldots, \theta_m\}$ and probabilities $\{p_1, \ldots, p_m\}$ such that ν approximates the categorical projection of $\mathcal{N}(\mu, \sigma^2)$:

$$\nu \approx \Pi_C \mathcal{N}(\mu, \sigma^2),$$

[Bonus 10 points]: Suppose that you are additionally given $\epsilon > 0$, and let $W \sim \Pi_C \mathcal{N}(\mu, \sigma^2)$. Modify your procedure to guarantee that the cumulative distribution of W and Z are sufficiently close. That is,

$$\big| \Pr(Z \leq z) - \Pr(W \leq z) \big| \leq \epsilon, \quad \forall\, z.$$