

Math: Pumpkin Spice Latte

Problem Story

Recently, the team bought a robot to help with one of the most important tasks in the company — brewing coffee.

The robot looks nice and is advertised to be "AI-powered". It can brew two types of coffee: **Cappuccino** and **Pumpkin Spice Latte**. The whole team loves Pumpkin Spice Latte, but unfortunately, the robot doesn't always brew it. The team needs to improve the robot's policy!

Alice is helping Jānis train the robot using **Reinforcement Learning (RL)**.

Reinforcement Learning (RL) — a machine learning approach where an agent learns to make decisions by taking actions in an environment and receiving feedback in the form of rewards. The agent's goal is to learn a policy (a set of decisions given the environment) that maximizes rewards over time.

Problem

The robot's decision-making is modeled as a "Policy Bag" containing tokens:

- **Pumpkin Spice tokens**
- **Cappuccino tokens**

The robot decides which coffee to brew by randomly selecting a token from the bag.

Due to a bug in the initialization code, they don't know which of two starting policies the robot loaded:

- **Policy A ("Advanced Robot"):** Started with 2 Pumpkin Spice tokens and 1 Cappuccino token.
- **Policy B ("Intern"):** Started with 1 Pumpkin Spice token and 2 Cappuccino tokens.

There's a 50/50 chance of either policy being loaded.

The Experiment:

Alice and Jānis ran a training experiment with two sequential episodes (Episode 2 occurred after the bag was updated in Episode 1):

Episode 1:

- The robot selected a token and brewed a Pumpkin Spice Latte!
- The team gave positive feedback.
- Following the RL algorithm, the robot returned the token to the bag and added one more Pumpkin Spice token.

Episode 2:

- The robot drew from the updated bag and brewed a Pumpkin Spice Latte again!

Task {1} Given that the robot brewed Pumpkin Spice Latte in both episodes, what is the probability that it started with **Policy A ("Advanced Robot")**?

Hint

This task can be solved using different approaches. The more advanced one involves **Bayes' Theorem**. It is not necessary to use it, and if you don't know it, you can still solve the problem. But if you wish to find out more about Bayes' Theorem (which is one of the most fundamental concepts in probability theory and Machine Learning), you can watch a nice video by 3b1b here: <https://www.youtube.com/watch?v=HZGCoVF3YvM>

Format

Submit the probability, without proof or explanation.
Give either a fraction or a decimal.

Grading

- 100% for correct task {1}