Bayesian learning and information foraging

A simple foraging task. The task modeled in `modellearning03.m` is a simple one-dimensional foraging task. Pellets are distributed according to one of several Gaussian sources. For each trial, the agent selects sample locations (along the x-axis) until the reward is found. Each sampling location can be sampled once per trial.

The agent learns the probability of reward at each sample location. We assume that the agent posits a set of possible models (latent causes) that produce Gaussian distributed reward. The agent uses a histogram over outcomes for previous sampling experience in order to predict the distribution of Bernoulli parameters at each sample location and, consequently, expected reward. The different models give different distributions of Bernoulli parameters. Using the Bayesian update associated with learning, we can compute the expected information gain at each sampling location.

Sampling behavior can be driven either by the expected reward (value) or the expected information gain. The former sampling strategy exploits current reward predictions while the latter explores potentially informative locations.

The code outputs two figures. You may need to change where the figures are printed on your screen. This can be done on lines 37 and 73 of `modellearning03.m`. The right figure shows the distribution for $p(\mu|D)$, the expected value and expected information gain, and the sampling distribution. The left figure shows the probability distributions $p(\mu|m_i)$ for each model the probability of the model given the data $p(m|D)$.

The purpose of the following exercises is to develop intuitions about differences between information-based exploration and reward-based exploitation.

Exercises

- The code is initialized to have three potential sources. Examine the expected information gain curve and explain why it has the shape it does.

- Examine the difference between information foraging and reward foraging. The action selection method is given on line 63 of `modellearning03.m`. Set the softmax parameter as b = 20 and run the code 5-10 times for each method. You can run the code by typing in `modellearning03(20)` at the command line. What happens to the probability of the model $p(m|D)$? *Plot the volatility or variance of $p(m|D)$ for each of the foraging strategies.* How does the sampling distribution change for each foraging strategy? How many samples are required to complete 20 trials (this is a measure of efficient reward collection)? *Plot the average number of trails required to complete 20 trials for each foraging strategy.*

- The action selection method can be switched from expected information gain-based exploration to exploiting expected value information by setting `a_method = 'switch'` on line 63. This switch from exploration to exploitation is based on the maximum expected information gain. Determine whether this switch increases model certainty and the rate of reward receipt (as given by the number of samples drawn to complete 20 trials).

- The information gain (in bits) for a given observation o is given by the change in the posterior

for the models relative to the prior

$$Infogain = D_{KL}\left(p(M|o,\mathbf{O}) \;\|\; p(M|\mathbf{O})\right) = \sum_M p(M|o,\mathbf{O})\log\frac{p(M|o,\mathbf{O})}{p(M|\mathbf{O})}$$

The expected information for a sampling position x is the sum of the information gain associated with a particular observation weighted by the probability of making that observation at position x.

On lines 14 and 15 of `modellearning03.m` you can change the number potential Gaussians (and their parameters) that account for the observations. Change the number of Guassian to two – with centers at $\mu = 4, 6$ and $\sigma = 2$. Explain why the expected information gain at $x = 5$ is zero. Imagine that we have pellets distributed according to two potential radial Gaussians on a two-dimensional plane. Where would the zero expected information gain sampling locations be? [Hint: Think about the geometry.] Where would the zero expected information gain sampling locations be if we added a third pellet dispenser?

- Vary the precision of the model distributions $p(\mu|m_i)$ on line 16. How does the uncertainty in these model distributions affect reward foraging and information foraging?