SI252 Reinforcement Learning 2020/02/27

Homework 1

Professor: Ziyu Shao *Due: 2020/03/22 11:59am*

1. Sampling from probability distributions. Show histograms and compare them to corresponding PDF.
(a) Sampling from the Logistic distribution by using Unif(0,1).
(b) Sampling from the Rayleigh distribution by using Unif (0,1).
(c) Sampling from the standard Normal distribution with both the Box-Muller method and the Acceptance-Rejection method. Compare the pros and cons of both methods.
(d) Sampling from the Beta distribution (you can use any method introduced in our class).

2. Given a random variable $X \sim N(0,1)$, evaluate the tail probability $P(X > 8)$
(a) Use the standard sample average method.
(b) Use the importance sampling method.

3. A coin with probability p of landing Heads is flipped repeatedly. Let N denote the number of flips until the pattern HH is observed.
(a) Suppose that p is a known constant, with $0 < p < 1$. Find $E(N)$
(b) Now suppose that p is unknown, and that we use a $Beta(a, b)$ prior to reflect our uncertainty about p (where a and b are known constants and are greater than 2). What is the expected number of flips until the pattern HH is observed.

4. Instead of predicting a single value for the parameter, we given an interval that is likely to contain the parameter: A $1 - \delta$ confidence interval for a parameter p is an interval $[\hat{p} - \epsilon, \hat{p} + \epsilon]$ such that $Pr\left(p \in [\hat{p} - \epsilon, \hat{p} + \epsilon]\right) \geq 1 - \delta$. Now we toss a coin with probability p landing heads and probability $1 - p$ landing tails. The parameter p is unknown and we need to estimate its value from experiments results. We toss such coin N times, Let $X_i = 1$ if the ith result is head, otherwise 0. We estimate p by using $\hat{p} = \frac{X_1 + \ldots + X_N}{N}$. Find the confidence interval for p, then discuss the impacts of δ and N.

5. We know that the MMSE of X given Y is given by $g(Y) = E[X|Y]$. We also know that the Linear Least Square Estimate (LLSE) of X given Y, denoted by $L[X|Y]$, is shown as follows:

$$L[X|Y] = E(X) + \frac{Cov(X,Y)}{Var(Y)}(Y - E(Y)).$$

Now we wish to estimate the probability of landing heads, denoted by θ, of a biased coin. We model θ as the value of a random variable Θ with a known prior PDF

$f_\Theta \sim \text{Unif}(0,1)$. We consider n independent tosses and let X be the number of heads observed. Find the MMSE $E[\Theta|X]$ and the LLSE $L[\Theta|X]$.

6. Given a coin with the probability p of landing heads. p is unknown and we need to estimate its value through data. In our data collection model, we have n independent tosses, result of each toss is either Head or Tail. Let X denote the number of heads in the total n tosses. Now we conduct experiments to collect data and find $X = k$. Then we need to find $\hat{p}$, the estimation of p.
(a) Assume p is a random variable with a prior distribution $p \sim Beta(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MAP (Maximum a Posterior Probability) rule.
(b) Assume p is an unknown constant. Find $\hat{p}$ through the MLE (Maximum Likelihood Estimation) rule.
(c) Assume p is a random variable with a prior distribution $p \sim Beta(a, b)$, where a and b are known constants. Find $\hat{p}$ through the MMSE (Minimal Mean Squared Error) rule.

7. Assume a random person's birthday is uniformly distributed on the 365 days of the year. People enter the room one by one. How many people are in the room the first time that two people share the same birthday? Let K be the desired number. Find $E(K)$ (integral form).

8. Suppose buses arrive at a bus stop according to a Poisson process N_t with parameter λ. Given a fixed $t > 0$. The time of the last bus before t is S_{N_t}, and the time of the next bus after t is S_{N_t+1}. Show the following identity:
$$E(S_{N_t+1} - S_{N_t}) = \frac{2 - e^{-\lambda t}}{\lambda}$$

9. Given k skill levels, we define a reward function $H(\cdot) : \{1, \ldots, k\} \to \mathcal{R}$. Then for skill levels $x \in \{1, \ldots, k\}$ and $y \in \{1, \ldots, k\}$, we define a soft-max function
$$\pi(x) = \frac{e^{H(x)}}{\sum_{y=1}^{k} e^{H(y)}}.$$
Please show the following result: for any skill level $a \in \{1, \ldots, k\}$, we have
$$\frac{\partial \pi(x)}{\partial H(a)} = \pi(x) \left(1_{\{x=a\}} - \pi(a)\right),$$
where 1_A is an index function of events, being 1 when event A is true and being 0 otherwise.