SI252 Reinforcement Learning 2020/03/06

Homework 1

Professor: Ziyu Shao *Due: 2020/03/26 11:59am*

1. For a random variable X whose moment of order $r > 0$ is finite, we define the following norm

$$||X_{|}|_r = (\mathbb{E}(|X|^r))^{\frac{1}{r}}.$$

Show the following norm inequalities hold.

- **The Holder Inequality.** Let $\frac{1}{p} + \frac{1}{q} = 1$. If $\mathbb{E}(|X|^p), \mathbb{E}(|Y|^q) < \infty$, then $|\mathbb{E}(XY)| \leq \mathbb{E}|XY| \leq ||X||_p \cdot ||Y||_q$.
- **The Lyapunov Inequality.** For $0 < r \leq p$, $||X||_r \leq ||X||_p$.
- **The Minkowski Inequality**. Let $p \geq 1$, $\mathbb{E}(|X|^p), \mathbb{E}(|Y|^p) < \infty$, then $||X + Y||_p \leq ||X||_p + ||Y||_p$.

2. Let the random variables $X_1, X_2, \ldots, X_n$ be independent with $E(X_i) = \mu$, $a \leq X_i \leq b$ for each $i = 1, \ldots, n$, where a, b are constants. Then for any $\epsilon \geq 0$, show the following inequality hold (Hoeffding Bound):

$$\mathbb{P}(|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu| \geq \epsilon) \leq 2e^{-\frac{2n\epsilon^2}{(b-a)^2}}.$$

3. 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.

4. 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.

5. Show the following theorem hold (Orthogonality Property of MMSE).
   (a) For any function $\phi(\cdot)$, one has

$$E[(Y - E[Y|X])\phi(X)] = 0$$

   (b) Moreover, if the function $g(X)$ is such that

$$E[(Y - g(X))\phi(X)] = 0, \forall \phi(\cdot).$$

   then $g(X) = E(Y|X)$

6. The Linear Least Square Estimate (LLSE) of Y given X, denoted by $L[Y|X]$, is the linear function $a + bX$ that minimizes $E[(Y - a - bX)^2]$. Show the following equation hold:

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

7. 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 $L[\Theta|X]$.

8. 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.