تمرین اول درس یادگیری ماشین آماری

نیمسال اول ۹۹-۹۸ ؛ مهلت ارسال : ۱۷ آبان ۱۳۹۸

Theoretical (part A)

1. Two computers are connected to a password-protected wireless network. When the password is temporarily removed, a virus can attack the first computer with probability 0.5, the computer with probability 0.7, and it can attack both computers with probability 0.4.
    a. The first computer appears infected with a virus. What is the probability that the second computer was also attacked?
    b. The first computer was not attacked at all. What is the probability that the second computer was attacked?

2. Let X have probability density function f_x:
    a. Find the CDF of X
    b. Let $Y = \frac{1}{X}$. Find the PDF of $f_Y(y)$ for Y.

$$f_x = \begin{cases} 0.25 & 0 < x < 1 \\ 0.375 & 2.5 < x < 4 \\ 0 & otherwise \end{cases}$$

3. Show that $\mathbb{V}(x) = 0$, if and only if there is a constant *c* such that $P(X = c) = 1$

4. Let X nad Y have joint PDF $f(x, y) = 16xy \;\; for \; 0 < x < 1 \; and \; 0 < y < 1$
    a. What is the $\mathbb{E}(X)$?
    b. What is the (Y) ?

5. Let $X_1, X_2, \dots, X_n \sim Uniform(0,1)$ and let $Y_n = \max\{X_1, X_2, \dots, X_n\}$. Find $\mathbb{E}(Y_n)$.

6. Let $X \sim N(0,1)$ and let $Y = e^X$. Find $\mathbb{E}(Y)$ and $\mathbb{V}(Y)$.

7. Suppose we generate a random variable X in the following way. First we flip a fair coin, if the coin is heads, take X to have a $Unif(0,1)$ distribution. If the coin is tails, take X to have a $Unif(3,4)$ distribution.
    a. Find the mean of X
    b. Find the standard deviation of X

Programming (part B)

1. Suppose a coin has probability of ***p*** falling heads. If we flip coin many times, we would expect the proportion of heads to be near *p*.
    a. Take *p*=0.3 and *n*=1000 and simulate coin flips. Plot the proportion of heads as a function of n.
    b. Repeat the problem for *p*=0.03

2. Suppose we flip a coin *n* times and let *p* denote the probability of heads. Let *X* be the number of heads. We call *X* a binomial random variable. Intuition suggests that *X* will be close to *np*. To see if this is true, we can repeat this experiment many times and average the values. Carry out a simulation and compare the average of the *X*'s to *np*. Try this for *p*=0.3 and *n*=10,100, 1000.

3. Consider tossing a fair die. Let $A = \{2, 4, 6\}$ and $B = \{1,2,3,4\}$. Then, $p(A) = 1/2$, $P(B) = 2/3$ and $\mathrm{P(AB)} = 1/3$. Since $P(AB) \vdots P(A)P(B)$, the events A and B are independent. Simulate

draws from the sample space and verify that $P(AB) = P(A)P(B)$ Where $P(A)$ is the proportion of times A occurred in the simulation and similarly for $P(AB)$ and $P(B)$. Now find two events A and B that are not independent. Compute $P(A)$,$P(B)$ and $P(AB)$. Compare the calculated values to their theoretical values via drawing venn diagram .Report your results and interpret.

4. Use computer simulation to show that changing selected door in "Monty Hall Problem" leads to higher winning probability. Simulate game for 2000 rounds and change the door in first 1000 rounds. Calculate winning probability in two conditions and compare them.

5. Let $X \sim N(3,16)$
   a. Find $P(X < 7)$
   b. Find $P(X > -2)$
   c. Find x such that $P(X > x) = 0.05$
   d. Find $P(-2 \leq X \leq 2)$
   e. Find $P(0 \leq X \leq 4)$
   f. Find x such that $P(|X| > |x|) = 0.05$

6. Consider $X1 \sim Binomial(100, 0.3)$, $X2 \sim Binomial(100,0.5)$ and $X3 \sim Binomial(200,0.5)$. Show X1+X2 and X2+X3 have binomial distributions and find their parameters.

7. Consider $X_i \sim N(0,\frac{1}{i})$. Draw CDF of X_i· for i=1, 100, 1000, 10000. Consider an arbitrary very small number (epsilon). Show X_i converges to 0 in probability and distribution.

8. Let $X, Y \sim Unif(0,1)$ be independent. Simulate the PDF for $X - Y$ and X/Y.

9. Let $X_1, X_2, \dots, X_n$ be $N(0,1)$ random variables and $\bar{X}_n$ is sample mean of first n samples. Plot $\bar{X}_n$ versus n for $1, \dots, 10000$. Repeat for $X_1, X_2, \dots, X_n$ be $Cauchy$. Explain why there is such a difference.

10. Use a computer simulation to generate realizations of a $Poisson(6)$ random variable by approximating it with a $binomial(100, 0.06)$ random variable. What is the average value of X?

11. Write a program to generate a pair of Gaussian random numbers (X_1, X_2) with zero mean and covariance $E(X_1^2) = 1$, $(X_2^2) = \frac{1}{3}$, E(X_1X_2)=$\frac{1}{2}$. Generate 1000 pairs of such numbers, evaluate their sample averages and sample covariance.

12. Try to simulate the following questions:
    a. Generate 1000 samples to estimate $P(|\bar{X} - p| > \varepsilon)$ in Example 5.3 of textbook for n=100, ε=0.2 and p=0.3. Estimate $\bar{X}$ after each n=100 sample generation. Compare the result with boundaries in Example 5.3, Example 5.6 of course textbook.
    b. Repeat part (a) for p=0.5 and report results.