Module 3 Homework

1) (5 points)

Suppose a random variable X has pdf as $f(x) = 2e^{-2(x-1)}, \quad x > 1$. Which of the following represents P(0 < X < 4)? (Note: you do not need to solve for exact number).

(a) $\int_0^4 2e^{-2(x-1)}dx$;

(b) $\int_1^4 2e^{-2(x-1)}dx$;

(c) $\int_0^4 x2e^{-2(x-1)}dx$;

(d) $\sum_{x=0}^{4} 2e^{-2(x-1)}$;

(e) $\int_1^\infty x2e^{-2(x-1)}dx$.

2) (10 points)

A random variable X has pdf

$$f(x) = \frac{2^x}{x!} e^{-2}, \quad x = 0, 1, 2, \ldots .$$

Find P(X = 1).

Then find $P(-2 < X < 4)$.

Give your answers to at least four decimal places.

3) (5 points)

If two carriers of the gene for albinism marry and have children, then each of their children has a probability of 1/4 of being albino. Let the random variable Y denote the number of their albino children out of all 3 of their children. Then Y follows a binomial(n, p) distribution. Find the values for n and p.

n =_____ p = _____

4) (10 points)

For Y following a binomial (n = 3, p = 0.25) distribution, compute the following:

$P(Y \leq 2) =$

$E(Y) =$

$Var(Y) =$

Give your answers to at least four decimal places.

5) (20 points)

For X following a Chi-square distribution with degree of freedom m = 3, compute the following:

$P(1 < X < 4) =$

$E(X) =$

$Var(X) =$

Give your answers to at least four decimal places.

Also, use a Monte Carlo simulation with sample size n=100,000 to estimate $P(1 < X < 4)$. What is your Monte Carlo estimate? Does it agrees with the answer above?

6) (10 points)

Suppose X follows a Chi-square distribution with degree of freedom m = 5 so that E(X) = 5 and Var(X) = 10. Also, let Y = 4X - 10. Find E(Y) and Var (Y). Does Y follow a Chi-square distribution with degree of freedom m=10?

E(Y) =

Var(Y) =

Does Y follow a Chi-square distribution with degree of freedom m = 10?

7) (20 points)

The Zyxin gene expression values are distributed according to $N(\mu = 1.6, \sigma = 0.4)$.

(a) What is the probability that a randomly chosen patient have the Zyxin gene expression values between 1 and 1.6?

(b) Use a Monte Carlo simulation of sample size n=500,000 to estimate the probability in part (a). Give your R code, and show the value of your estimate.

(c) What is the probability that exactly 2 out of 5 patients have the Zyxin gene expression values between 1 and 1.6?

Please show your work on how to arrive at the answer. Give your answer to at least four decimal places.

8) (20 points)

(a) Hand in a R script that calculates the mean and variance of two random variables X~F(m=2,n=5) and Y~F(m=10,n=5) from their density functions.

(b) Use the formula in Table 3.4.1 to calculate the means and variances directly.

(c) Run your script in (a), and check that your answers agree with those from part (b).