Assignment 1

MATH414

Statistics

INSTRUCTIONS

- Write your assignments with a pen on paper and submit it in the physical form.
- The use of ChatGPT or any other generative AI or, in fact, any "calculator" such as "WolframAlpha" is strictly forbidden. If you use any of these, you will get zero in the entire Assignment. The exception to this rule is when you want to check a calculation once you have performed it.
- You are welcome to look up any theory or definitions from any source you like but you must cite it.
- Write your answers cleanly.
- Excuse the large number of questions on this Assignment. Most of them are ungraded and may be used for revision. The ungraded questions are not required to be submitted.

QUESTIONS 1 - GRADED

(1) (Athreya, Sarkar & Tanner, pp79) Let X and Y be random variables with joint distribution given by the chart below.

	X=0	X=1	X=2
Y=0	1/12	0	3/12
Y=1	2/12	1/12	0
Y=2	3/12	1/12	1/12

(a) Compute the marginal distributions of X and Y.
(b) Compute the conditional distribution of X given that Y=2.
(c) Compute the conditional distribution of Y given that X=2.
(d) Carry out a computation to show that X and Y are not independent.

(2) (Athreya, Sarkar & Tanner, pp80) Let X be the result of a fair die roll and let Y be the number of heads in X coin flips.
(a) Both X and (Y |X = n) can be written in terms of common distributions using the notation $\sim$. What is the distribution of X? What is the distribution of (Y |X = n) for n = 1, . . . 6?
(b) Determine the joint distribution for X and Y .
(c) Calculate the marginal distribution of Y .
(d) Compute the conditional distribution of X given that Y = 6.
(e) Compute the conditional distribution of X given that Y = 0.

(f) Perform a computation to prove that X and Y are not independent.

(3) (Athreya, Sarkar & Tanner, pp80) Suppose the number of earthquakes that occur in a year, anywhere in the world, is a Poisson random variable with mean λ. Suppose the probability that any given earthquake has magnitude at least 5 on the Richter scale is p independent of all other quakes. Let $N \sim$ Poisson(λ) be the number of earthquakes in a year and let M be the number of earthquakes in a year with magnitude at least 5, so that (M | N = n) $\sim$ Binomial(n, p).

(a) Calculate the joint distribution of M and N .

(b) Show that the marginal distribution of M is determined by

$$P(M = m) = \frac{1}{m!} e^{-\lambda} (\lambda p)^m \sum_{n=m}^{\infty} \frac{\lambda^{n-m}}{(n-m)!} (1-p)^{n-m}$$

for $m > 0$.

(c) Perform a change of variables (where k = n-m) in the infinite series from part (b) to prove

$$P(M = m) = \frac{1}{m!} e^{-\lambda} (\lambda p)^m \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} (1-p)^k.$$

(d) Use part (c) together with the infinite series equality together with the infinite series for e^x to conclude that $M \sim$ Poisson($\lambda\, p$)

(4) (Athreya, Sarkar & Tanner, pp134) Let $X \sim$ Geometric(p) and let A be the event $(X \leq 3)$. Calculate $E[X|A]$ and $Var[X|A]$.

(5) (Athreya, Sarkar & Tanner, pp134) Let X and Y be described by the joint distribution

	X=-1	X=0	X=1
Y=-1	1/15	2/15	2/15
Y=0	2/15	1/15	2/15
Y=1	2/15	2/15	1/15

(a) Calculate $E[X|Y = -1]$.

(b) Calculate $Var[X|Y = -1]$.

(c) Describe the distribution of $E[X|Y]$.

(d) Describe the distribution of $Var[X|Y]$.

(6) (Athreya, Sarkar & Tanner, pp140) Consider the experiment of flipping two coins. Let X be the number of heads among the coins and let Y be the number of tails among the coins.

(a) Should you expect X and Y to be posivitely correlated, negatively correlated, or uncorrelated? Why?

(b) Calculate Cov [X, Y] to confirm your answer to (a).

(7) (Athreya, Sarkar & Tanner, pp167) Let X be a random variable whose probability function $f : \mathbb{R} \to [0, 1]$ is given by

$$f(x) = \begin{cases} kx^{k-1}e^{-x^k} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the distribution function of X for $k = 2$.

(b) Find the distribution function of X for any k.

(8) (Athreya, Sarkar & Tanner, pp170) Let X be a continuous random variable such that its distribution function is strictly increasing on the set $\{x \in \mathbb{R} : 0 < F(x) < 1\}$. The median of X is the value of x for which $P(X > x) = P(X < x) = 1/2$. Find median of

(a) $X \sim$ Uniform(a, b).

(b) $X \sim$ Exp(λ).

(c) $Z \sim$ Normal(μ, σ^2)

(9) (Athreya, Sarkar & Tanner, pp178) Let $X \sim$ Uniform$(0, 1)$ and let $Y = \sqrt{X}$. Determine the density of Y.

(10) (Athreya, Sarkar & Tanner, pp178) Let $X \sim$ Uniform$(0, 1)$ and let $Y = 1/X$. Determine the density of Y.

(11) A point is chosen randomly in a disk of radius 1. Let R be the distance of the point from the origin. Let Θ be the angle it makes with the positive x-axis.

(a) Find the probability $P(R \leq r)$ that the distance of the point from the centre is less than r for $r \in (0, 1)$. This gives the distribution funtion for R. Differentiate to obtain the density for R.

(b) Find the probability $P(\Theta \leq \theta)$ that the angle that the line joining the point and centre makes with the positive x-axis is less than θ for $\theta \in (0, 2\pi]$. This gives the distribution funtion for Θ. Differentiate to obtain the density for Θ.

(c) Find the probability $P(R \leq r, \Theta \leq \theta)$ that the distance of the point from the centre is less than r for $r \in (0, 1)$ and the angle that the line joining the point and centre makes with the positive x-axis is less than θ for $\theta \in (0, 2\pi]$. This gives the joint distribution funtion for (R, Θ). Differentiate twice (once with respect to r and second with respect to θ) to obtain the density for (R, Θ). Hence, prove that R and Θ are the marginals of (R, Θ) and hence that they are independent.

(d) Define $X = R\cos\Theta$. Use their independence to determine the expected value and variance of X. Similarly, perform these calculations for $Y = R\sin\Theta$. Finally, determine covariance of X and Y.

(12) (Athreya, Sarkar & Tanner, pp238) Let $X \sim \text{Normal}(0, 1)$. Use the moment generating function to calculate $E[X^4]$.

(13) (Athreya, Sarkar & Tanner, pp238) Let $Y \sim \text{Exponential}(\lambda)$.

(a) Calculate the moment generating function $M_Y(t)$.

(b) Hence, find the third and the fourth moments of Y.

Questions 2 - Ungraded

(1) (Athreya, Sarkar & Tanner, pp69) A pair of fair dice are thrown. Let X represent the larger of the two values on the dice and let Y represent the smaller of the two values.

(a) Describe S, the domain of functions X and Y . How many elements are in S?

(b) What are the ranges of X and Y . Do X and Y have the same range? Why or why not?

(c) Describe the distribution of X and describe the distribution of Y by finding the probability mass function of each. Is it true that X and Y have the same distribution?

(2) (Athreya, Sarkar & Tanner, pp69) A pair of fair dice are thrown. Let X represent the number of the first die and let Y represent the number of the second die.

(a) Describe S, the domain of functions X and Y . How many elements are in S?

(b) Describe T , the range of functions X and Y . How many elements are in T ?

(c) Describe the distribution of X and describe the distribution of Y by finding the probability mass function of each. Is it true that X and Y have the same distribution?

(d) Are X and Y the same function? Why or why not?

(3) (Athreya, Sarkar & Tanner, pp78) An urn has four balls labeled 1, 2, 3, and 4. A first ball is drawn and its number is denoted by X. A second ball is then drawn from the three remaining balls in the urn and its number is denoted by Y.

(a) Calculate P (X = 1).

(b) Calculate P (Y = 2 | X = 1).

(c) Calculate P (Y = 2).

(d) Calculate P (X = 1, Y = 2).

(e) Are X and Y independent? Why or why not?

(4) (Athreya, Sarkar & Tanner, pp96) Let X $\sim$ Bernoulli(p) and Y$\sim$ Bernoulli(q) be independent.

(a) Prove that XY is a Bernoulli random variable. What is its parameter?

(b) Prove that (1 - X) is a Bernoulli random variable. What is its parameter?

(c) Prove that X + Y - XY is a Bernoulli random variable. What is its parameter?

(5) (Athreya, Sarkar & Tanner, pp97) Let $X_1, X_2, ..., X_n$ be an i.i.d. sequence of discrete random variables and let Z be the maximum of these n variables. Let r be a real number and let $R = P(X_1 \leq r)$. Prove that $P(Z \leq r) = R^n$.

(6) (Athreya, Sarkar & Tanner, pp97) Let $X_1, X_2, ..., X_n$ be an i.i.d. sequence of discrete random variables and let Z be the minimum of these n variables. Let r be a real number and let $R = P(X_1 \leq r)$. Prove that $P(Z \leq r) = 1 - (1 - R)^n$.

(7) (Athreya, Sarkar & Tanner, pp109) Suppose X and Y are random variables. Suppose E[X] =∞ and E[Y] = $-\infty$

(a) Provide an example to show that E[X+Y]=∞ is possible.

(b) Provide an example to show that E[X+Y]=$-\infty$ is possible.

(c) Provide an example to show that E[X+Y] may have finite value.

(8) (Athreya, Sarkar & Tanner, pp140) In previous sections it was shown that if X and Y are independent, then Var[X + Y] = Var[X] + Var [Y]. If X and Y are dependent, the result is typically not true, but the covariance provides a way to relate the variances of X and Y to the variance of their sum.

(a) Show that for any discrete random variables X and Y, Var[X+Y]=Var[X]+Var[Y]+ 2 Cov[X,Y]

(b) Use (a) to conclude that when X and Y are positively correlated, then Var [X + Y] $>$ Var[X] + Var[Y], while when X and Y are negatively correlated, Var[X + Y] $<$ Var[X] + Var[Y].

(c) Use induction to prove a similar result for n random variables.

(9) (Athreya, Sarkar & Tanner, pp119) Suppose we begin rolling a die, and let X be the number of rolls needed before we see the first 3.

(a) Show that E[X] = 6.

(b) Calculate SD[X].

(c) Viewing SD[X] as a typical distance of X from its expected value, would it seem unusual to roll the die more than nine times before seeing a 3?

(d) Calculate the actual probability P (X$>$9).

(e) Calculate the probability X produces a result within one standard deviation of its expected value.

(10) (Athreya, Sarkar & Tanner, pp134) A standard light bulb has an average lifetime of four years with a standard deviation of one year. A Super D-Lux lightbulb has an average lifetime of eight years with a standard devaition of three years. A box contains many bulbs – 90% of which are standard bulbs and 10% of which are Super D-Lux bulbs. A bulb is selected at random from the box. What are the average and standard deviation of the lifetime of the selected bulb?

(11) (Athreya, Sarkar & Tanner, pp219) Suppose X has p.d.f. given by

$$f(x) = \begin{cases} \frac{\cos(x)}{2} & \text{if } -\pi/2 \leq x \leq \pi/2, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Compute the distribution function of X.

(b) Compute its expected value and variance.

(12) (Athreya, Sarkar & Tanner, pp220)

(a) Let $\lambda > 0$, X and Y be two independent Exponential(λ) random variables. Let Z = X + Y. Compute the density of Z.

(b) Compute the sum of n independent Exponential(λ) r.v.s. This is the same as Gamma(n,λ).

(c) Compute the expected value and variance of Exponential(λ) and therefore that of Gamma(n,λ).

(13) (Athreya, Sarkar & Tanner, pp227) Let (X, Y) be uniformly distributed on the triangle $0 < x < y < 1$.

(a) Compute $E[X|Y = 1/6]$.

(b) Compute $E[(X - Y)^2]$.

(14) (Athreya, Sarkar & Tanner, pp227) X is a random variable with mean 3 and variance 2. Y is a random variable with mean -1 and variance 6. The covariance of X and Y is -2. Let $U = X + Y$ and $V = X - Y$. Find the correlation coefficient of U and V.

(15) (Athreya, Sarkar & Tanner, pp228) Suppose Y is uniformly distributed on $(0, 1)$ and suppose for $0 < y < 1$ the conditional density of $X|Y = y$ is given by

$$f_{X|Y=y}(x) = \begin{cases} \frac{2x}{y^2} & \text{if } 0 < x < y, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that, as a function of x, $f_{X|Y=y}$ is a density.

(b) Compute the joint pdf of (X, Y) and the marginal density of X.

(c) Compute the expected value and variance of X given that $Y = y$ with $0 < y < 1$.

(16) (Athreya, Sarkar & Tanner, pp228) Let (X, Y) have joint ptobability density function $f : \mathbb{R}^2 \to \mathbb{R}$. Show that $\text{Var}[X|Y = y] = E[X^2|Y = y] - E[X|Y = y]^2$.

(17) (Athreya, Sarkar & Tanner, pp167) Let $R > 0$ and $X \sim$ Uniform$[0, R]$. Let $Y \sim \min\{X, R/10\}$. Find the distribution function of Y.

(18) (Athreya, Sarkar & Tanner, pp166) Let X be a continuous random variable with distribution function F : R $\to$ [0, 1]. Then G : R $\to$ [0, 1] given by G(x) = 1 - F (x) is called the reliability function of X or the right tail distribution function of X. Suppose T $\sim$ Exponential(λ) for some $\lambda > 0$, then find the reliability function of T .

(19) (Athreya, Sarkar & Tanner, pp193-194) Let $c > 0$. Suppose that X and Y are random variables with joint probability density

$$f(x) = \begin{cases} c(xy + 1) & \text{if } 0 \leq x \leq 1 \text{ and } 0 \leq y \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find c.

(b) Compute the marginal densities f_X and f_Y and the conditional density $f_{X|Y=b}$.

(20) (Athreya, Sarkar & Tanner, pp194) Let $A = \{(x, y) \in \mathbb{R}^2 : x > 0, y > 0, x + y < 1\}$. Let X and Y be random variables defined by the joint density $f(x, y) = 24xy$ if $x \in A$ andd 0 elsewhere.

(a) Verify the claim that $f(x, y)$ is a density.

(b) Show that X and Y are dependent random variables.

(c) Explain why (b) is true despite f being in a product form.

(21) (Athreya, Sarkar & Tanner, pp208) Let X and Y be two random variables with the joint p.d.f. given by

$$f(x) = \begin{cases} \lambda^2 e^{-\lambda y} & \text{if } 0 \leq x \leq y, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Compute the marginal distribution of X and Y.

(b) Compute the conditional distribution of $(Y|X = x)$ for some $x > 0$.

(22) (Athreya, Sarkar & Tanner, pp208) Let $a, b > 0$. Let $X \sim$ Gamma(a, b) and $Y \sim$ Exponential(X).

(a) Find the joint density of X and Y.

(b) Find the marginal density of Y.

(c) Find the conditional density of $(X|Y = y)$.

(23) (Athreya, Sarkar & Tanner, pp209) Let $a, b, c > 0$. Let $X \sim$ Gamma(a, c) and $Y \sim$ Gamma(b, c) be two independent random variables. Prove that $X + Y$ is distributed as Gamma$(a + b, c)$.

(24) (Athreya, Sarkar & Tanner, pp209) Let $a, b, c > 0$. Let $X \sim$ Gamma(a, c) and $Y \sim$ Gamma(b, c) be two independent random variables. Prove that $X + Y$ is distributed as Gamma$(a + b, c)$.

(a) Let $W = \frac{Y}{X}$. Find the probability density function of W.

(b) Let $Z = \frac{X}{X+Y}$. Find the probability density function of Z.

(c) Are X and Z independent?