Probability and Statistics
Problem Set I

1. Let E, F, G be three events. Find expressions for the events that of E, F, G

   (a) only F occurs
   (b) both E and F but not G occur
   (c) at least one event occurs
   (d) at least two event occur
   (e) all three events occur
   (f) none occurs
   (g) at most one event occurs
   (h) at most two events occur

   Answer: **(a)** $F \cap E^c \cap G^c$, **(b)** $E \cap F \cap G^c$, **(c)** $E \cup F \cup G$, **(d)** $(E \cap F) \cup (E \cap G) \cup (F \cap G)$, **(e)** $E \cap F \cap G$, **(f)** $(E \cup F \cup G)^c$, **(g)** $(E \cap F)^c \cap (E \cap G)^c \cap (F \cap G)^c$, **(h)** $(E \cap F \cap G)^c$

2. Let $\mathcal{S} = \{0, 1, 2, \cdots\}$ and $E \subseteq \mathcal{S}$. Then in each of the following cases, verify P is a probability on $\mathcal{S}$.

   (a) $P(E) = \sum_{x \in E} \frac{e^{-\lambda} \lambda^x}{x!}$, $\lambda > 0$.
   (b) $P(E) = \sum_{x \in E} p(1-p)^x$, $0 < p < 1$.
   (c) $P(E) = 0$, if E has finite number of elements, and $P(E) = 1$, if E has infinite number of elements.

   Answer: (a) Yes (b) Yes (c) No

3. Let $\mathcal{S} = \mathbb{R}$ and $\Sigma = \mathfrak{B}_{\mathbb{R}}$. in each of the following cases, does P define a probability on $\mathcal{S}$?

   (a) For each interval I, let
   $$P(I) = \int_I \frac{1}{\pi(1+x^2)} dx$$
   (b) For each interval I, let $P(I) = 1$ if I is an interval of finite length, and $P(I) = 0$ if I is an infinite interval.
   (c) For each interval I, let $P(I) = 1$ if $I \subseteq (-\infty, 1)$ and $P(I) = \int_I (\frac{1}{2}) dx$ if $I \subseteq [1, \infty)$. (If $I = I_1 \cup I_2$, where $I_1 \subseteq (-\infty, 1)$ and $I_2 \subseteq [1, \infty)$, then $P(I) = P(I_2)$)

4. For events $E_1, E_2, \cdots, E_n$, show that

   (a) $P(\bigcup_{i=1}^{n} E_i) \leq \sum_{i=1}^{n} P(E_i)$.
   This is known as Boole's inequality.
   (b) $P(\bigcap_{i=1}^{n} E_i) \geq \sum_{i=1}^{n} P(E_i) - (n-1)$.
   This is known as Bonferroni's inequality.

(c) $P(\bigcap_{i=1}^{n} E_i) = P(E_1)P(E_2|E_1)P(E_3|E_1 \cap E_2) \cdots P(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1})$, where $P(E_1 \cap E_2 \cap \cdots \cap E_{n-1}) > 0$.

5. Let E and F be two independent events. Then show that

   (a) E^c and F are independent.

   (b) E and F^c are independent.

   (c) E^c and F^c are independent.

6. Let E and F be two events such that $P(E) = p_1 > 0$, $P(F) = p_2 > 0$ and $p_1 + p_2 > 1$. Show that $P(F|E) \geq 1 - \frac{1-p_2}{p_1}$.

7. For any two events E and F, show that $P(E \cap F) - P(E)P(F) = P(E)P(F^c) - P(E \cap F^c) = P(E^c)P(F) - P(E^c \cap F) = P((E \cup F)^c) - P(E^c)P(F^c)$.

8. A cell-phone tower has a circular coverage area of radius 10 km. If a call is initiated from a random point in the coverage area, find the probability that the call comes from within 2 km of the tower.

9. What is the chance that a leap year, selected at random, will contain 53 Sundays?

   Answer: $\frac{2}{7}$.

10. A family has two children. What is the conditional probability that both are boys given that at least one of them is a boy?

    Answer: $\frac{1}{3}$

11. Three numbers are chosen from 1 to 20. Find the probability that they are not consecutive.

    Answer: $\frac{187}{190}$

12. Consider two events A and B with $P(A) = 0.4$ and $P(B) = 0.7$. Determine the maximum and minimum possible values of $P(A \cap B)$ and the conditions under which each of these values is attained.

13. Two digits are chosen at random without replacement from the set of integers $\{1, 2, 3, 4, 5, 6, 7, 8\}$.

    (a) Find the probability that both digits are greater than 5.

    (b) Show that the probability that the sum of the digits will be equal to 5 is the same as the probability that their sum will exceed 13.

    Answer: (a) $\dfrac{\binom{3}{2}}{\binom{8}{2}}$

14. Calculate the probability of drawing from a pack of cards one that is an ace or is a spade or shows an even number $(2, 4, 6, 8, 10)$.

    Answer: 31/52

15. A school contains students in grades $1, 2, 3, 4, 5,$ and 6. Grades $2, 3, 4, 5,$ and 6 all contain the same number of students, but there are twice of this number in grade 1. If a student is selected at random from a list of all the students in the school, what is the probability that she/he will be in grade 3?

16. Suppose that the blood test for some disease is reliable in the following sense: for people who are infected with the disease the test produces a positive result in 99.99% of cases; for people not infected a positive test result is obtained in only 0.02% of cases. Furthermore, assume that in the general population one person in $10,000$ people is infected. A person is selected at random and found to test positive for the disease. What is the probability that the individual is actually infected?

    Answer: 1/3

17. Consider two independent fair coins tosses, in which all four possible outcomes are equally likely. Let $H_1 = \{\text{1st toss is a head}\}$, $H_2 = \{\text{2nd toss is a head}\}$, and $D = \{\text{the two tosses have different results}\}$. Find $P(H_1), P(H_2), P(H_1 \cap H_2), P(H_1|D), P(H_2|D)$, and $P(H_1 \cap H_2|D)$.

    Answer: $P(H_1) = \frac{1}{2} = P(H_2), P(H_1 \cap H_2) = \frac{1}{4}, P(H_1|D) = \frac{1}{2} = P(H_2|D)$, and $P(H_1 \cap H_2|D) = 0$.

18. There are two roads from A to B and two roads from B to C. Each of the four roads has probability p of being blocked by snow, independently of all the others. What is the probability that there is an open road from A to C?

19. Suppose that we have $n \geq 2$ letters and corresponding n addressed envelopes. If these letters are inserted at random in n envelopes, find the probability that no letter is inserted into the correct envelope.

    Answer: $\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n+2}\frac{1}{n!}$.

20. One ball is to be selected from a box containing red, white, blue, yellow, and green balls. If the probability that the selected ball will be red is $1/5$ and the probability that it will be white is $2/5$, what is the probability that it will be blue, yellow, or green?

21. Suppose that $n(\geq 3)$ persons $P_1, \cdots, P_n$ are made to stand in a row at random. Find the probability that there are exactly r persons between P_1 and P_2; here $r \in \{1, 2, \cdots, n-2\}$.

    Answer: $\frac{2(n-r-1)}{n(n-1)}$

22. Three numbers are chosen at random and without replacement from the set $\{1, 2, ..., 50\}$. Find the probability that the chosen numbers are in (a) arithmetic progression, and (b) geometric progression.

    Answer: (a) $\frac{600}{\binom{50}{3}}$ (b) (a) $\frac{44}{\binom{50}{3}}$

23. The organization that David Jones works for is running a father son dinner for those employees having at least one son. Each of these employees is invited to attend along with his youngest son. If Jones is known to have two children, what is the conditional probability that they are both boys given that he is invited to the dinner?

    Answer: $\frac{1}{3}$

24. An urn contains b black balls and r red balls. One of the balls is drawn at random, but when it is put back in the urn c additional balls of the same color are put in with it. Now suppose that we draw another ball. Find the probability that the first ball drawn was black given that the second ball drawn was red.

    Answer: $\frac{b}{b+r+c}$

25. In answering a question on a multiple-choice test a student either knows the answer or guesses. Let p be the probability that he knows the answer and $1-p$ the probability that he guesses. Assume that a student who guesses at the answer will be correct with probability $1/m$, where m is the number of multiple-choice alternatives. What is the conditional probability that a student knew the answer to a question given that she answered it correctly?

    Answer: $\frac{mp}{1+(m-1)p}$

26. There are three coins in a box. One is a two-headed coin, another is a fair coin, and the third is a biased coin that comes up heads 75 percent of the time. When one of the three coins is selected at random and flipped, it shows heads. What is the probability that it was the two-headed coin?

    Answer: $\frac{4}{9}$

27. A conservative design team, call it C and an innovative design team, call it N, are asked to design a new product within a month. From the past experience we know that

    (a) the probability that team C is successful is $2/3$

    (b) the probability that team N is successful is $1/2$

    (c) the probability that at least one team is successful is $3/4$

    Assuming that exactly one successful design is produced, what is the probability that it was designed by team N?

    Answer: $\frac{1}{4}$

28. A box of fuses contains 20 fuses, of which 5 are defective. If 3 of the fuses are selected at random and removed from the box in succession without replacement, what is the probability that all three fuses are defective?

    Answer: $\frac{1}{114}$

29. There is a 1% probability for a hard drive to crash. Therefore, it has two backups, each having a 2% probability to crash, and all three components are independent of each other. The stored information is lost only in an unfortunate situation when all three devices crash. What is the probability that the information is saved?

    Answer: 0.99

30. A box contains 100 balls, of which r are red. Suppose that the balls are drawn from the box one at a time, at random, without replacement. Determine (a) the probability that the first ball drawn will be red; (b) the probability that the 50th ball drawn will be red; and (c) the probability that the last ball drawn will be red.

31. In screening for a certain disease, the probability that a healthy person wrongly gets a positive result is 0.05. The probability that a diseased person wrongly gets a negative result is 0.002. The overall rate of the disease in the population being screened is 1%. If the person X test gives a positive result, what is the probability that the person X actually have the disease?

    Answer: 0.168

32. Alice is taking a probability class and at the end of each week she can be either up-to-date or she may have fallen behind. If she is up-to-date in a given week, the probability that she will be up-to-date (or behind) in the next week is 0.8 (or 0.2, respectively). If she is behind in a given week, the probability that she will be up-to-date (or behind) in the next week is

0.4 (or 0.6, respectively). Alice is (by default) up-to-date when she starts the class. What is the probability that she is up-to-date after three weeks?

Answer: 0.688

33. Consider an empty box in which four balls are to be placed (one-by-one) according to the following scheme. A fair die is cast each time and the number of spots on the upper face is noted. If the upper face shows up 2 or 5 spots then a white ball is placed in the box. Otherwise a black ball is placed in the box. Given that the first ball placed in the box was white find the probability that the box will contain exactly two black balls.

    Answer: $\frac{4}{9}$

34. On a midterm exam, students X, Y, and Z forgot to sign their papers. Professor knows that they can write a good exam with probabilities $0.8, 0.7$, and 0.5, respectively. After the grading, he notices that two unsigned exams are good and one is bad. Given this information, and assuming that students worked independently of each other, what is the probability that the bad exam belongs to student Z?

    Answer: 0.595

Probability and Statistics
Problem Set II

1. Are the following functions cumulative distribution function?

(a) $F(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 - e^{-x} & \text{if } x \geq 0 \end{cases}$

(b) $F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x, -\infty < x < \infty.$

(c) $F(x) = \begin{cases} 0 & \text{if } x < -5 \\ x & \text{if } -5 \leq x \leq 0.5 \\ 1 & \text{if } x > 0.5 \end{cases}$

Hint: Use the property of cumulative distribution function
Answer: (a) Yes (b) Yes and (c) No

2. Let X be a random variable having the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{2}{3} & \text{if } 0 \leq x < 1 \\ \frac{7-6c}{6} & \text{if } 1 \leq x < 2 \\ \frac{4c^2-9c+6}{4} & \text{if } 2 \leq x \leq 3 \\ 1 & \text{if } x > 3 \end{cases}$$

where c is a real constant.

(a) Find the value of constant c.

(b) Find the p.m.f of X.

(c) Find $P(1 < X < 2), P(2 \leq X < 3), P(0 < X \leq 1), P(1 \leq X \leq 2), P(X \geq 3)$ and $P(X = 2.5)$.

(d) Find the conditional probabilities $P(\{X = 1\}|\{1 \leq X \leq 2\}), P(\{1 \leq X < 2\}|\{X > 1\})$, and $P(\{1 \leq X \leq 2\}|\{X = 1\})$.

Answer: (a) $\frac{1}{4}$, (b) the p.m.f is

$$p(x) = \begin{cases} \frac{2}{3} & \text{if } x = 0 \\ \frac{1}{4} & \text{if } x = 1 \\ \frac{1}{12} & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

(c) $P(1 < X < 2) = 0, P(2 \leq X < 3) = \frac{1}{12}, P(0 < X \leq 1) = \frac{1}{4}, P(1 \leq X \leq 2) = \frac{1}{3}, P(X \geq 3) = 0$ and $P(X = 2.5) = 0$.
(d) $P(\{X = 1\}|\{1 \leq X \leq 2\}) = \frac{3}{4}, P(\{1 \leq X < 2\}|\{X > 1\}) = 0$, and $P(\{1 \leq X \leq 2\}|\{X = 1\}) = 1$.

3. Let us select five cards at random and without replacement from an ordinary deck of playing cards.

(a) Find the p.m.f. of X, the number of hearts in the five cards.

(b) Determine $P(\{X \leq 1\})$.

Answer: (a) the p.m.f is

$$p(x) = \begin{cases} \dfrac{\binom{13}{x}\binom{39}{5-x}}{\binom{52}{5}} & \text{if } x = 0,1,2,3,4,5 \\ 0 & \text{otherwise} \end{cases}$$

(b) $P(\{X \leq 1\}) = p(0) + p(1)$

4. Let X be a random variable having the p.m.f.

$$p(x) = \begin{cases} \frac{c}{(2x-1)(2x+1)} & \text{if } x \in \{1,2,3,\cdots\} \\ 0 & \text{otherwise} \end{cases}$$

where c is a real constant.

(a) Find the value of constant c.

(b) Find the cumulative distribution function of X.

(c) For the positive integers m and n such that $m < n$, evaluate $P(X < m+1), P(X \geq m), P(m \leq X < n)$ and $P(m < X \leq n)$.

(d) Find the conditional probabilities $P(\{X > 1\}|\{1 \leq X < 4\})$ and $P(\{1 < X < 6\}|\{X \geq 3\})$.

Answer: (a) $c = 2$ (b)the c.d.f is

$$F(x) = \begin{cases} 0 & \text{if } x < 1 \\ 1 - \frac{1}{2[x]+1} & \text{if } x \geq 1 \end{cases}$$

(c) $P(X < m+1) = 1 - \frac{1}{2m+1}, P(X \geq m) = \frac{1}{2m-1}, P(m \leq X < n) = \frac{2(n-m)}{(2m-1)(2n-1)}$ and $P(m < X \leq n) = \frac{2(n-m)}{(2m+1)(2n+1)}$.
(d) $P(\{X > 1\}|\{1 \leq X < 4\}) = \frac{2}{9}$ and $P(\{1 < X < 6\}|\{X \geq 3\}) = \frac{6}{11}$.

5. Let X be a random variable having the cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x^2}{2} & \text{if } 0 \leq x < 1 \\ \frac{x}{2} & \text{if } 1 \leq x < 2 \\ 1 & \text{if } x \geq 2 \end{cases}$$

.

(a) Show that X is of continuous type and find p.d.f of X.

(b) Find $P(1 < X < 2), P(1 \leq X < 2), P(1 < X \leq 2), P(1 \leq X \leq 2), P(X \geq 1)$ and $P(X = 1)$.

Answer: (a) X is of continuous type with p.d.f

$$f(x) = \begin{cases} x & \text{if } 0 \le x < 1 \\ \frac{1}{2} & \text{if } 1 \le x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

(b) $P(1 < X < 2) = \frac{1}{2}, P(1 \le X < 2) = \frac{1}{2}, P(1 < X \le 2) = \frac{1}{2}, P(1 \le X \le 2) = \frac{1}{2}, P(X \ge 1) = \frac{1}{2}$ and $P(X = 1) = 0$.

6. A bag contains ten balls. Among them six are red and four are white. Three balls are drawn at random and not replaced. Find the probability mass function for the number of red balls drawn.
Answer: the p.m.f is

$$p(x) = \begin{cases} \dfrac{\binom{6}{x}\binom{4}{3-x}}{\binom{10}{3}} & \text{if } x = 0, 1, 2, 3 \\ 0 & \text{otherwise} \end{cases}$$

7. Let X be a random variable with p.d.f

$$f(x) = \begin{cases} c(4x - 2x^2) & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

where c is a real constant. Then

(a) Find the value of c.

(b) Find the cumulative distribution function of X.

Answer: (a) $c = 3/8$ (b)the c.d.f

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{4}(3x^2 - x^3) & \text{if } 0 \le x < 2 \\ 1 & \text{if} x \ge 2 \end{cases}$$

8. Let X be a random variable with p.m.f

$$P(X = r) = \binom{n}{r} p^r (1-p)^{n-r}, \ r = 0, 1, 2, \cdots, n, \ 0 \le p \le 1.$$

Find the p.m.f of the random variables (a) $Y = aX + b$ and (b) $Y = X^2$.
Answer: (a) the p.m.f

$$\text{For } a = 0, \ p(y) = \begin{cases} 1 & \text{if } y = b \\ 0 & \text{if } y \ne b \end{cases}$$

$$\text{For } a \ne 0, \ p(y) = \begin{cases} \binom{n}{\frac{y-b}{a}} p^{\frac{y-b}{a}} (1-p)^{n-\frac{y-b}{a}} & \text{if } y \in \{b, a+b, 2a+b, \cdots, na+b\} \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$p(y) = \begin{cases} \binom{n}{\sqrt{y}} p^{\sqrt{y}}(1-p)^{n-\sqrt{y}} & \text{if } y \in \{0, 1, 4, \cdots, n^2\} \\ 0 & \text{otherwise} \end{cases}$$

9. A random variable X has a p.d.f $f(x)$ given by ce^{-x} in the interval $0 < x < \infty$ and zero elsewhere. Find the value of the constant c and hence calculate the probability that X lies in the interval $1 < X \leq 2$.

Answer: $c = 1$ and $P(1 < X \leq 2) = e^{-1} - e^{-2}$.

10. Let X have range $[0, 3]$ and density $f_X(x) = kx^2$. Let $Y = X^3$.
(a) Find k and the cumulative distribution function of X. (b) Compute $E[Y]$.
(c) Compute $Var(Y)$.
(d) Find the probability density function $f_Y(y)$ for Y .
Answer: (a) $k = 1/9$, (b) 13.5, (c) 60.75 (d) $\frac{1}{27}$ on $[0, 27]$.

11. Let X be a random variable with p.m.f.

$$p(x) = \begin{cases} \frac{1}{7} & \text{if } x \in \{-2, -1, 0, 1\} \\ \frac{3}{14} & \text{if } x \in \{2, 3\} \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the m.g.f. of X.

(b) Find the p.m.f. of $Z = X^2$ and its distribution function.

Answer: (a) The m.g.f. of X is $M_X(t) = \frac{1}{7}(e^{-2t} + e^{-t} + e^t + 1) + \frac{3}{14}(e^{2t} + e^{3t})$ (b) The p.m.f. of Z is

$$p_Z(z) = \begin{cases} \frac{1}{7} & \text{if } z = 0 \\ \frac{2}{7} & \text{if } z = 1 \\ \frac{5}{14} & \text{if } z = 4 \\ \frac{3}{14} & \text{if } z = 9 \\ 0 & \text{otherwise} \end{cases}$$

and the distribution function of Z is

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ \frac{1}{7} & \text{if } 0 \leq z < 1 \\ \frac{3}{7} & \text{if } 1 \leq z < 4 \\ \frac{11}{14} & \text{if } 4 \leq z < 9 \\ 1 & z \geq 9 \end{cases}$$

12. Let X be a random variable with p.m.f.

$$p(x) = \begin{cases} \frac{1}{3}(\frac{2}{3})^x & \text{if } x \in \{0, 1, 2, \cdots\} \\ 0 & \text{otherwise} \end{cases}$$

Find the distribution function of $Z = \frac{X}{X+1}$ and hence find p.m.f. of Z.
Answer: The distribution function of Z is

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 - (\frac{2}{3})^{i+1} & \text{if } \frac{i}{i+1} \leq z < \frac{i+1}{i+2}, i \in \{0, 1, 2, \cdots\} \end{cases}$$

and p.m.f. of Z is

$$f_Z(z) = \begin{cases} \frac{1}{3}(\frac{2}{3})^{\frac{z}{1-z}} & \text{if } z \in \{0, \frac{1}{2}, \frac{1}{3}, \cdots\} \\ 0 & \text{otherwise} \end{cases}$$

13. Let X be a random variable with p.d.f.

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

Find the distribution function $Z = X^2$ and hence find its p.d.f.
Answer: The distribution function of Z is

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ 1 - e^{-\sqrt{z}} & \text{if } z \geq 0 \end{cases}$$

and p.d.f. of Z is

$$f_Z(z) = \begin{cases} \frac{e^{-\sqrt{z}}}{2\sqrt{z}} & \text{if } z > 0 \\ 0 & \text{otherwise} \end{cases}$$

14. Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} 6x(1-x) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the distribution function $Z = X^2(3 - 2X)$ and hence find its p.d.f.
Answer: The distribution function of Z is

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ z & \text{if } 0 \leq z < 1 \\ 1 & \text{if } z \geq 1 \end{cases}$$

and p.d.f. of Z is

$$f_Z(z) = \begin{cases} 1 & \text{if } 0 < z < 1 \\ 0 & \text{otherwise} \end{cases}$$

15. If m.g.f. of a random variable X is $M_X(t) = \frac{e^t - e^{-2t}}{3t}$, for $t \neq 0$, then find a p.d.f. of X.

16. Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} \frac{x}{2} & \text{if } 0 < x \leq 1 \\ \frac{1}{2} & \text{if } 1 < x \leq 2 \\ \frac{3-x}{2} & \text{if } 2 < x \leq 3 \\ 0 & \text{otherwise} \end{cases}$$

Find the expected value of $Y = X^2 - 5X + 3$.
Answer: $-\frac{11}{6}$

17. Let X be a random variable having the p.m.f.

$$p(x) = \begin{cases} \frac{3}{\pi^2 x^2} & \text{if } x \in \{\pm 1, \pm 2, \pm 3, \cdots\} \\ 0 & \text{otherwise} \end{cases}$$

Then show that expectation of X is not finite.

18. Let X be a random variable having a p.d.f.

$$f(x) = \frac{e^{-|x|}}{2}, \text{ for } x \in \mathbb{R}$$

Then show that expectation of X is finite and find its value.
Answer: 0

19. For any random variable X having the mean μ and finite second moment, show that $E((X-\mu)^2) \leq E((X-c)^2)\ \forall\ c \in \mathbb{R}$.

20. Let X be a continuous random variable with p.d.f. $f_X(x)$ that is symmetric about $\mu \in \mathbb{R}$, i.e., $f_X(\mu + x) = f_X(\mu - x)$, for all $x \in \mathbb{R}$. If $E(X)$ finite, then show that $E(X) = \mu$.

21. Let X be a random variable such that $P(X \leq 0) = 0$ and $\mu = E(X)$ is finite. Show that $P(X \geq 2\mu) \leq 0.5$. (Use Markov inequality)

22. If X is a random variable such that $E(X) = 3$ and $E(X^2) = 13$, then determine a lower bound for $P(-2 < X < 8)$. (Use Chebyshev inequality)
Answer: $\frac{21}{25}$

Note: p.d.f - probability density function, p.m.f - probability mass function, & c.d.f - cumulative distribution function

Probability and Statistics
Problem Set III

Notations:

1. $X \sim \mathrm{Bin}(n,p)$ indicates that the random variable X has Binomial distribution with n trials and success probability p.

2. $X \sim \mathrm{NB}(r,p)$ indicates that the random variable X has negative Binomial distribution with r successes and success probability p.

3. $X \sim \mathrm{Ge}(p)$ indicates that the random variable X has Geometric distribution with success probability p.

4. $X \sim \mathrm{Hyp}(a,n,N)$ indicates that the random variable X has Hyper geometric distribution, where N is the total number of items, a is the number of success and n is the number of selected items .

5. $X \sim \mathrm{P}(\lambda)$ indicates that the random variable X has Poisson distribution with parameter $\lambda > 0$.

6. $X \sim U(\{x_1, x_2, x_3, \cdots, x_N\})$ indicates that the random variable X has Uniform distribution on the set $\{x_1, x_2, x_3, \cdots, x_N\}$.

7. $X \sim U(\alpha, \beta)$ indicates that the random variable X has Uniform distribution on the interval (α, β).

8. $X \sim G(\alpha, \lambda)$ indicates that the random variable X has Gamma distribution with parameters $\alpha > 0$ & $\lambda > 0$.

9. $X \sim \mathrm{Exp}(\lambda)$ indicates that the random variable X has Exponential distribution with parameters $\lambda > 0$.

10. $X \sim N(\mu, \sigma^2)$ indicates that the random variable X has Normal (Gaussian) distribution with parameters $\mu \in \mathbb{R}$ & $\sigma > 0$.

1. Let X be a random variable. Define $X_{(m)} = X(X-1)(X-2)\cdots(X-m+1)$, for $m \in \mathbb{N}$. Then find the expectation $E(X_{(m)})$ of $X_{(m)}$ for the following cases:

   (a) $X \sim \text{Bin}(n, p)$. **Answer:** $= n(n-1)(n-2)\cdots(n-m+1)p^m$

   (b) $X \sim \text{NB}(r, p)$. **Answer:** $= r(r+1)(r+2)\cdots(r+m-1)(\frac{1-p}{p})^m$

   (c) $X \sim \text{Hyp}(a, n, N)$. **Answer:** $= \dfrac{\binom{N-m}{n-m}}{\binom{N}{n}} a(a-1)(a-2)\cdots(a-m+1)$

   (d) $X \sim P(\lambda)$. **Answer:** $= \lambda^m$

2. A person has to open a lock whose key is lost among a set of N keys. Assume that out of these N keys only one can open the lock. To open the lock the person tries keys one by one by choosing, at each attempt, one of the keys at random from the unattempted keys. The unsuccessful keys are not considered for future attempts. Let Y denote the number of attempts the person will have to make to open the lock. Show that $Y \sim U(\{1, 2, 3, \cdots, N\})$ and hence find the mean and the variance of the r.v. Y.
   Answer: $E(Y) = \frac{N+1}{2}$ and $Var(Y) = \frac{N^2-1}{12}$.

3. Each child in a family is equally likely to be a boy or a girl. Find the minimum number of children the family should have so that the probability of it having at least a boy and at least a girl is at least 0.95. (Use Binomial distribution)
   Answer: 6

4. There are 30 applicants for a job, out of which only 20 applicants are qualified for the job. Six applicants are selected at random from these 30 applicants. Find the probability that, among the selected candidates, at least two will be qualified for the job. (Use Hyper geometric distribution)
   Answer: $1 - \dfrac{\binom{20}{0}\binom{10}{6}}{\binom{30}{6}} - \dfrac{\binom{20}{1}\binom{10}{5}}{\binom{30}{6}}$

5. The probability of hitting a target in each shot is 0.002. Find the approximate probability of hitting a target at least twice in 2000 shots. (Use Poisson distribution)
   Answer: $1 - 5e^{-4}$

6. Eighteen balls are placed at random in seven boxes that are labeled $B_1, \cdots, B_7$. Find the probability that boxes with labels B_1, B_2 and B_3 all together contain six balls.
   Answer: $\binom{18}{6}(\frac{3}{7})^6(\frac{4}{7})^{12}$

7. (i) Suppose a group of 100 men aged $60-64$ in Dehradun received a new flu vaccine from a health center in 2014. From the 2014 life table of the health center, it is found that the approximate probability that a man, aged between $60-64$, dies in the next year is 0.02. How likely are, at least 5 out of 100 men who received flu vaccine and aged $60-64$ to die within the next year?
   (ii) What is the probability that amongst the 60 to 64-year old men who got flu

vaccination exactly 25 survive and at least 10 die within the next year? (*you don't need to calculate the exact numerical values of the probabilities*)
Answer: $(i)\, 1 - \sum_{i=0}^{4} \binom{100}{i} (.02)^i (0.98)^{100-i}$, $(ii)\, \binom{100}{25} (.98)^{25} (.02)^{75}$.

8. Two teams (say Team A and Team B) play a series of games until one team wins 5 games. If the probability of Team A (Team B) winning any game is 0.7 (0.3), find the probability that the series will end in 8 games.
Answer: 0.188 (approximately).

9. Let $X \sim \text{Ge}(p)$. Find $E(\frac{1}{2^X})$.
Answer: $\frac{p}{p+1}$

10. The *characteristic function* (CF) of a random variable X is defined as

$$C_X(t) = E[e^{itX}] = \begin{cases} \sum_j e^{itx_j} p(x_j) & \text{for discrete case} \\ \int e^{itx} f(x) dx & \text{for continuous case} \end{cases}$$

where $i = \sqrt{-1}$. Show that the characteristic function for Binomial distribution $Bin(n, p)$ and Poisson distribution with mean λ are $(i)\, (pe^{it} + (1-p))^n$ and $(ii)\, e^{\lambda(e^{it}-1)}$.

11. The *cumulunt generating function* is defined as $K_X(t) = ln(\phi_X(t))$, where ϕ_X is the moment generating function of X. Show that $K'(0) = E[X],\ K''(0) = Var(X)$.

12. Let $X \sim U(0, k)$, where k is a positive integer and $Y = X - [X]$, where $[X]$ is the largest integer $\leq X$. Show that $Y \sim U(0, 1)$.

13. Let $X \sim U(0, 1)$, where k is a positive integer and $Y = -\ln(1 - X)$. Find the c.d.f. and p.d.f. of Y.

14. Alvin's driving time to work is between 15 and 20 minutes if the day is sunny, and between 20 and 25 minutes if the day is rainy, with all times being equally likely in each case. Assume that a day is sunny with probability 2/3 and rainy with probability 1/3. What is the p.d.f. of the driving time, viewed as a random variable X?

15. If X is described by a Gaussian distribution of mean μ and variance σ^2, calculate the probabilities that X lies within 1σ, 2σ and 3σ of the mean.
Answer: $0.6826, 0.9544, 0.9974$
(Hint: Use Normal table (z-table). $\Phi(z = 1) = 0.8413$; $\Phi(z = 2) = .9772$; $\Phi(z = 3) = .9987$).

16. Let $X \sim N(\mu, \sigma^2)$, , for some $\mu \in \mathbb{R}$ and $\sigma > 0$ and $Y = aX + b$, where $a \in \mathbb{R} - \{0\}$ and $b \in \mathbb{R}$. Then show that $Y \sim N(a\mu + b, a^2\sigma^2)$.

17. Let $X \sim N(2, 4)$. Find $P(X \leq 0)$, $P(|X| \geq 2)$, $P(1 < X \leq 3)$ and $P(X \leq 3 | X > 1)$.
Answer: $0.1587,\ 0.5228,\ 0.383\ \&\ 0.5539$
(Hint: $F_X(x) = \Phi(\frac{x-\mu}{\sigma})$ and use Normal table (z-table). $\Phi(z = -2) = 0.0228$; $\Phi(z = -1) = 0.1587$; $\Phi(z = 0) = 0.5$; $\Phi(z = 0.5) = .6915$).

18. Suppose the diameter of a certain car component follows the normal distribution with $X \sim N(10, 3^2)$. Find the proportion of these components that have diameter larger than 13.4 mm ($\Phi(z = 1.13) = 0.8708$).
Answer: 0.1292

19. The number of miles that a particular car can run before its battery wears out is exponentially distributed with an average of $10,000$ miles. The owner of the car needs to take a 5000-mile trip. What is the probability that he will be able to complete the trip without having to replace the car battery?
Answer: $e^{-1/2}$

20. Engineers designing the next generation of space shuttles plan to include two fuel pumps one active, the other in reserve. If the primary pump malfunctions, the second is automatically brought on line. Suppose a typical mission is expected to require that fuel be pumped for at most 50 hours. According to the manufacturer's specifications, pumps are expected to fail once every 100 hours. What are the chances that such a fuel pump system would not remain functioning for the full 50 hours?
Answer: $1 - \frac{3}{2}e^{-1/2}$

Probability and Statistics
Problem Set IV

Note: For $n \geq 2$, let $\underline{X} = (X_1, X_2, \cdots, X_n)$ be a random vector with joint cumulative distribution function F. We can also say that $(X_1, X_2, \cdots, X_n)$ is an n-dimensional random variable with joint cumulative distribution function F or $X_1, X_2, \cdots, X_n$ are random variables with joint cumulative distribution function F.

1. Check whether the following functions are joint cdf of some random vector or not.

   (a)

$$F(x,y) = \begin{cases} 1 & \text{if } x + 2y \geq 1 \\ 0 & \text{if } x + 2y < 1 \end{cases}$$

   (b)

$$F(x,y) = \begin{cases} 0 & \text{if } x < 0 \text{ or } x + y < 1 \text{ or } y < 0 \\ 1 & \text{otherwise} \end{cases}$$

   Answer: (a) No, (b) No

2. Suppose a fair coin is tossed three times. Let X be the number of heads and Y be the absolute difference between number of heads and number of tails in three tossing. Then show that $\underline{Z} = (X, Y)$ is a discrete type random vector.

3. Let S be a sample space and P be a probability function defined for all events. Let A and B be two events. Define the following random variables

$$X(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{otherwise} \end{cases} \quad \text{and } Y(w) = \begin{cases} 1 & \text{if } w \in B \\ 0 & \text{otherwise} \end{cases}.$$

   Then show that X and Y are independent if and only if A and B are independent.

4. Let X_1 and X_2 be independent identical distributed random variables with common p.m.f. $P(X = \pm 1) = \frac{1}{2}$. Show that X_1, X_2, X_3 are pairwise independent but not independent, where $X_3 = X_1 X_2$.

5. Let X and Y be two random variables with joint p.d.f.

$$f(x,y) = \begin{cases} \frac{1+xy}{4} & \text{if } |x| < 1, |y| < 1 \\ 0 & \text{otherwise} \end{cases}$$

   Then show that X and Y are not independent but X^2 and Y^2 are independent.

6. Toss a fair coin 3 times. Let X be the number of heads on the first toss, Y be the total number of heads on the last two tosses, and Z be the number of heads on the first two tosses.

   (a) Find the joint p.m.f. for X and Y . Compute $\text{Cov}(X, Y)$.
   (b) Find the joint p.m.f. for X and Z. Compute $\text{Cov}(X, Z)$.

   Answer: Calculate the joint p.m.f.s yourselves. Show that X and Y are independent.
   (b) $\text{Cov}(X, Z) = 1/4$

7. Roll a dice ($n = 1, 2, ...6$). Two events s_1 and s_2 are defined as follows:

$$s_1 = \begin{cases} 1 & \text{if } n \text{ is even number} \\ 0 & \text{otherwise} \end{cases}$$

$$s_2 = \begin{cases} 1 & \text{if n is prime number} \\ 0 & \text{otherwise} \end{cases}$$

Find the joint p.m.f. $P(s_1, s_2)$. Also find the covariance and correlation coefficient between s_1 and s_2.
Answer: $P(1,1) = P(0,0) = 1/6, P(1,0) = P(0,1) = 2/6$, $\text{Cov}(s_1, s_2) = -\frac{1}{12}$ $\rho(s_1, s_2) = -\frac{1}{3}$

8. The joint probability mass function of two discrete random variables X and Y is given by

$$p(x,y) = \begin{cases} c(2x+y) & \text{if } (x,y) \in \{0,1,2\} \times \{0,1,2,3\} \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of c.

(b) Find $P(X = 2, Y = 1)$, $P(X \geq 1, Y \leq 2)$

(c) Find the marginal p.m.f. of X and Y.

(d) Are X and Y independent.

Answer:(a) $\frac{1}{42}$, (b) $\frac{5}{42}$ & $\frac{4}{7}$(c)The p.m.f. of X is

$$p_X(x) = \begin{cases} \frac{1}{7} & \text{if } x = 0 \\ \frac{1}{3} & \text{if } x = 1 \\ \frac{11}{21} & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

and the p.m.f. of Y is

$$p_Y(y) = \begin{cases} \frac{1}{7} & \text{if } y = 0 \\ \frac{3}{14} & \text{if } y = 1 \\ \frac{2}{7} & \text{if } y = 2 \\ \frac{5}{14} & \text{if } y = 3 \\ 0 & \text{otherwise} \end{cases}$$

(d) No

9. **(a)** Prove that the correlation $\rho(X, Y)$ between two random variables X and Y remains unchanged under the transformation $X = aX + b$, where a, b are constants.

(b) Recall the relation between degrees Fahrenheit and degrees Celsius

$$\text{degrees Celsius} = \frac{5}{9} \text{ degrees Fahrenheit } -\frac{160}{9}$$

Let X and Y be the daily high temperature in degrees Fahrenheit for the summer in Delhi and Allahabad. Let T and S be the same temperatures in degrees Celsius.

Suppose that $\text{Cov}(X, Y) = 4$ and $\rho(X, Y) = 0.8$. Compute $\text{Cov}(T, S)$ and $\rho(T, S)$.
Answer: $\text{Cov}(T, S) = 100/81$, $\rho(T, S) = 0.8$

10. A card is drawn at random from a normal 52-card pack and its identity noted. The card is replaced, the pack shuffled and the process repeated. Random variables W, X, Y , Z are defined as follows:

    W = 2 if the drawn card is a heart; W = 0 otherwise.

    X = 4 if the drawn card is an ace, king, or queen; X = 2 if the card is a jack or ten; X = 0 otherwise.

    Y = 1 if the drawn card is red; Y = 0 otherwise.

    Z = 2 if the drawn card is black and an ace, king or queen; Z = 0 otherwise.

    Establish the correlations between each pair of random variables W, X, Y , Z.
    Answer: $\rho(X, Z) = 0.598, \rho(Y, Z) = -0.361, \rho(W, Z) = -0.209, \rho(Y, W) = 0.577$.

11. The probability distribution for the number of eggs in a clutch is $P(\lambda)$, and the probability that each egg will hatch is p (independently of the size of the clutch). Show by direct calculation that the probability distribution for the number of chicks that hatch is $P(\lambda p)$. ($P(\lambda)$ denotes Poisson distribution with mean λ).

12. A miner is trapped in a mine containing three doors. The first door leads to a tunnel that takes him to safety after two hours of travel. The second door leads to a tunnel that returns him to the mine after three hours of travel. The third door leads to a tunnel that returns him to his mine after five hours. Assuming that the miner is at all times equally likely to choose any one of the doors, what is the expected length of time until the miner reaches safety?
    Answer: 10.

13. A and B roll a pair of dice in turn, with A rolling first. A's objective is to obtain a sum of 6, and B's is to obtain a sum of 7. The game ends when either player reaches his or her objective, and that player is declared the winner. (a) Find the probability that A is the winner. (b) Find the expected number of rolls of the dice. (c) Find the variance of the number of rolls of the dice.
    Answer: (a) 30/61 (b) 402/61, (c) Var(X)$\approx$ 36.24

14. At a party N men throw their hats into the center of a room. The hats are mixed up and each man randomly selects one. Let X be the number of men those get their own hat. Find $E(X)$ and $Var(X)$.
    Answer: 1 and 1

15. Let (X, Y) have the joint p.m.f as follows

$$p(x,y) = \begin{cases} \frac{2}{15} & \text{if } (x,y) = (1,1) \\ \frac{4}{15} & \text{if } (x,y) = (1,2) \\ \frac{3}{15} & \text{if } (x,y) = (1,3) \\ \frac{1}{15} & \text{if } (x,y) = (2,1) \\ \frac{1}{15} & \text{if } (x,y) = (2,2) \\ \frac{4}{15} & \text{if } (x,y) = (2,3) \\ 0 & \text{otherwise} \end{cases}$$

Find the correlation coefficient of X and Y.
Answer: $\frac{21}{\sqrt{7236}}$

16. Let $X_1, X_2, \cdots, X_n$ be independent random variables. Suppose that $X_i's$ are continuous type with same cumulative distribution function for every $1 \le i \le n$. Let $Y_1 = max(X_1, X_2, \cdots, X_n)$ and $Y_2 = min(X_1, X_2, \cdots, X_n)$. Find the cumulative distribution functions of Y_1 and Y_2.
Answer: Suppose F is the c.d.f. of $X_i's$. Then the c.d.f. of Y_1 is $(F)^n$ and the c.d.f. of Y_2 is $1 - (1 - F)^n$.

17. If the number of typographical errors per page type by a certain typist follows a Poisson distribution with a mean of λ, find the probability that the total number of errors in 10 randomly selected pages is 10.
Answer: $\frac{e^{-10\lambda}(10\lambda)^{10}}{10!}$

18. We start with a stick of length l. We break it at a point which is chosen randomly and uniformly over its length and keep the piece that contains the left end of the stick. We then repeat the same process on the piece that we were left with. What is the expected length of the piece that we are left with after breaking twice?
Answer: $\frac{l}{4}$

19. Let $X_1, X_2, \cdots, X_n$ be independent random variable such that $X_i \sim N(\mu_i, \sigma_i^2)$, $\mu_i \in \mathbb{R}, \sigma_i > 0$, $i = 1, \cdots, n$. If $a_1, a_2, \cdots, a_n$ are real constant, such that not all of them are zero, then show that $\displaystyle\sum_{i=1}^{n} a_i X_i \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$

20. Let $X_1, X_2, \cdots, X_n$ be independent random variable such that $X_i \sim \text{NB}(r_i, p)$, $r_i \in \mathbb{N}$, $i = 1, \cdots, n$ and $0 < p < 1$. Then show that $\displaystyle\sum_{i=1}^{n} X_i \sim \text{NB}\left(\sum_{i=1}^{n} r_i, p\right)$

21. Let (X, Y) be random vector with joint p.d.f.

$$f(x,y) = \begin{cases} cx + 1 & \text{if } x, y \ge 0 \text{ and } x + y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of c.

(b) Find the marginal p.d.f. of X and Y.

(c) Find $P(Y < 2X^2)$.

Answer:(a) 3, (b)The p.d.f. of X is

$$f_X(x) = \begin{cases} (3x+1)(1-x) & \text{if } 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$

and the p.d.f. of Y is

$$f_Y(y) = \begin{cases} \frac{1}{2}(5-3y)(1-y) & \text{if } 0 \le y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(c) $\frac{53}{96}$

22. Let (X, Y) be random vector with joint p.d.f.

$$f(x, y) = \begin{cases} 6e^{-(2x+3y)} & \text{if } x, y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

(a) Are X and Y independent.

(b) Find $E(Y|X > 2)$.

(c) Find $P(X > Y)$.

Answer:(a)Yes, (b)$\frac{1}{3}$, (c)$\frac{3}{5}$.

23. Let X be random variable with p.d.f.

$$f_X(x) = \begin{cases} 2x & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

We know that given $X = x$, the random variable Y uniformly distributed on $[-x, x]$.

(a) Find joint p.d.f. of X and Y.

(b) Find p.d.f. of Y.

(c) Find $P(|Y| < X^3)$.

Answer:(a)The joint p.d.f. is

$$f(x, y) = \begin{cases} 1 & \text{if } |y| \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(b) the p.d.f. of Y is

$$f_Y(y) = \begin{cases} 1-|y| & \text{if } |y| \le 1 \\ 0 & \text{otherwise} \end{cases}$$

(c) $\frac{1}{2}$

24. Let X be random variable with p.d.f.

$$f_X(x) = \begin{cases} 4x(1-x^2) & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

For a fixed $x \in (0, 1)$, the conditional p.d.f. of Y given $X = x$ is

$$f_{Y|X}(y|x) = \begin{cases} \frac{2y}{1-x^2} & \text{if } x < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the conditional p.d.f. of X given $Y = y$ for appropriate values of y

(b) Find $E(X|Y = 0.5)$ and $\text{Var}(X|Y = 0.5)$.

(c) Find $P(0 < Y < \frac{1}{3})$ and $P(\frac{1}{3} < Y < \frac{2}{3}|X = 0.5)$.

Answer: (a) For $0 < y < 1$

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

(b)$E(X|Y = 0.5) = \frac{1}{3}$ and $\text{Var}(X|Y = 0.5) = \frac{1}{72}$, (c)$P(0 < Y < \frac{1}{3}) = \frac{1}{81}$ and $P(\frac{1}{3} < Y < \frac{2}{3}|X = 0.5) = \frac{7}{27}$.

25. Let (X, Y) be random vector with joint p.d.f.

$$f(x, y) = \begin{cases} \frac{1}{2}ye^{-xy} & \text{if } 0 < x < \infty, 0 < y < 2 \\ 0 & \text{otherwise} \end{cases}$$

Find $E(e^{\frac{X}{2}}|Y = 1)$.
Answer: 2

26. Let $\underline{X} = (X_1, X_2)$ be a random vector with joint p.d.f.

$$f(x_1, x_2) = \begin{cases} 2 & \text{if } 0 \le x_1 \le x_2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the cumulative distribution function of $Y = X_1 + X_2$ and hence find the p.d.f. of Y.
Answer:(a)The c.d.f. is

$$F_Y(y) = \begin{cases} 0 & \text{if } y < 0 \\ \frac{y^2}{2} & \text{if } 0 \le y < 1 \\ 1 - \frac{(y-2)^2}{2} & \text{if } 1 \le y < 2 \\ 1 & \text{if } y \ge 2 \end{cases}$$

(b) the p.d.f. of Y is

$$f_Y(y) = \begin{cases} y & \text{if } 0 \le y \le 1 \\ 2 - y & \text{if } 1 \le y \le 2 \\ 0 & \text{otherwise} \end{cases}$$

27. Let $\underline{X} = (X_1, X_2)$ be a random vector with joint p.d.f.

$$f(x_1, x_2) = \begin{cases} 8x_1x_2 & \text{if } 0 < x_1 < x_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Find the pdf of $Y_1 = \frac{X_1}{X_2}$.
Answer: The p.d.f. is

$$f_{Y_1}(y) = \begin{cases} 2y & \text{if } 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}$$

28. Let $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function defined by

$$f(x,y) = \begin{cases} \frac{e^{-(y+\frac{x}{y})}}{y} & \text{if } 0 < x, y < \infty \\ 0 & \text{otherwise} \end{cases}$$

(a) Verify that f is a joint probability density function of some random vector (X, Y).

(b) Find $\mathrm{Cov}(X, Y)$.

Answer: 1

29. Let X and Y be jointly continuous random variables with joint p.d.f.

$$f(x,y) = \begin{cases} x + cy^2 & \text{if } 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the constant c.

(b) Find joint c.d.f.

(c) $P(0 \leq X \leq \frac{1}{2}, 0 \leq Y \leq \frac{1}{2})$.

(d) Find the marginal p.d.f. of X and Y.

Answer: (a) $\frac{3}{2}$, (b)The joint c.d.f. is

$$F_{X,Y}(x,y) = \begin{cases} 0 & \text{if } x < 0 \text{ or } y < 0 \\ \frac{1}{2}x^2y + \frac{1}{2}xy^3 & \text{if } 0 \leq x, y \leq 1 \\ \frac{1}{2}x^2 + \frac{1}{2}x & \text{if } 0 \leq x \leq 1 \text{ and } y \geq 1 \\ \frac{1}{2}y + \frac{1}{2}y^3 & \text{if } 0 \leq y \leq 1 \text{ and } x \geq 1 \\ 1 & \text{if } x \geq 1 \text{ and } y \geq 1 \end{cases}$$

, (c) $\frac{3}{32}$, (d)The p.d.f. of X is

$$f_X(x) = \begin{cases} x + \frac{1}{2} & \text{if } 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and the p.d.f. of Y is

$$f_Y(y) = \begin{cases} \frac{3}{2}y^2 + \frac{1}{2} & \text{if } 0 \leq y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

30. Let $\underline{X} = (X_1, X_2)$ be a random vector with joint probability density function

$$f_X(x_1, x_2) = \begin{cases} \frac{1}{2}e^{-x_1} & \text{if } 0 < x_2 < x_1 < \infty \\ \frac{1}{2}e^{-x_2} & \text{if } 0 < x_1 < x_2 < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Find the joint probability density function of $Y_1 = X_1 + X_2$ and $Y_2 = \dfrac{X_2}{X_1 + X_2}$.

Answer:

$$f_{(Y_1,Y_2)}(y_1, y_2) = \begin{cases} \dfrac{1}{2}|y_1|e^{-y_1(1-y_2)} & \text{if } 0 < y_1y_2 < y_1(1-y_2) < \infty \\ \dfrac{1}{2}|y_1|e^{-y_1y_2} & \text{if } 0 < y_1(1-y_2) < y_1y_2 < \infty \\ 0 & \text{otherwise} \end{cases}$$

31. Let X_1, X_2, X_3 be independent identically distributed $Exp(1)$ random variables. Find the joint p.d.f. of $Y_1 = X_1 + X_2 + X_3$, $Y_2 = \frac{X_1+X_2}{X_1+X_2+X_3}$, and $Y_3 = \frac{X_1}{X_1+X_2}$. Also find the marginal p.d.f. of Y_1, Y_2, and Y_3 and hence show that Y_1, Y_2, and Y_3 are independent.
Answer: The joint p.d.f. is

$$f_{(Y_1,Y_2,Y_3)}(y_1, y_2, y_3) = \begin{cases} y_1^2 y_2 e^{-y_1} & \text{if } 0 < y_1 < \infty, 0 < y_2, y_3 < 1 \\ 0 & \text{otherwise} \end{cases}$$

The marginal p.d.f. of Y_1 is

$$f_{Y_1}(y_1) = \begin{cases} \frac{1}{2} y_1^2 e^{-y_1} & \text{if } 0 < y_1 < \infty \\ 0 & \text{otherwise} \end{cases}$$

The marginal p.d.f. of Y_2 is

$$f_{Y_2}(y_2) = \begin{cases} 2y_2 & \text{if } 0 < y_2 < 1 \\ 0 & \text{otherwise} \end{cases}$$

The marginal p.d.f. of Y_3 is

$$f_{Y_3}(y_3) = \begin{cases} 1 & \text{if } 0 < y_3 < 1 \\ 0 & \text{otherwise} \end{cases}$$

Probability and Statistics
Problem Set V

Note: Normal Approximation to Binomial Bin(n, p) is good if $np(1-p) \geq 10$. Similarly Normal Approximation to Poisson $P(\lambda)$ is good if $\lambda \geq 10$.
Table of continuity correction:

Discrete	Continuous
$P(X = a)$	$P(a - 0.5 < X < a + 0.5)$
$P(X > a)$	$P(X > a + 0.5)$
$P(X \leq a)$	$P(X < a + 0.5)$
$P(X < a)$	$P(X < a - 0.5)$
$P(X \geq a)$	$P(X > a - 0.5)$

1. Let $X \sim \mathrm{N}(\mu, \sigma^2)$ be a normal random variable. Then show that the random variable $Y = (X-\mu)^2/\sigma^2$ is distributed as gamma distribution $G(1/2, 1/2)$. Let us now consider n independent Normal random variables $X_i \sim \mathrm{N}(\mu_i, \sigma_i^2)$, $i = 1, 2, ..., n$, and define the new variable
$$\chi_n^2 = \sum_{i=1}^{n} \frac{(X - \mu_i)^2}{\sigma_i^2}.$$
Show that χ_n^2, will be distributed as $G(1/2, n/2)$, which have a p.d.f.
$$f(\chi_n^2) = \frac{1}{2^{n/2}\Gamma(n/2)}(\chi_n^2)^{n/2-1}e^{(-\frac{1}{2}\chi_n^2)}.$$
This is known as the *chi-squared distribution* of order n and has numerous applications in statistics.

2. Let $X_1, X_2, \cdots, X_{10}$ be independent Poisson random variables with mean 1. Use the central limit theorem to approximate $P((X_1 + X_2 + \cdots + X_{10}) \geq 15)$.

3. Let X_i, $i = 1, 2, \cdots, 10$ be independent random variables, each being uniformly distributed over $(0, 1)$. Estimate $P((X_1 + X_2 + \cdots + X_{10}) > 7)$.
Answer: 0.0139

4. The lifetime of a special type of battery is a random variable with mean 40 hours and standard deviation 20 hours. A battery is used until it fails, at which point it is replaced by a new one. Assuming a stockpile of 25 such batteries, the lifetimes of which are independent, approximate the probability that over 1100 hours of use can be obtained.
Answer: 0.1587

5. Let X be the number of times that a fair coin, flipped 40 times, lands heads. Find the probability that $X = 20$. Use the normal approximation and then compare it to the exact solution.
Answer: 0.1272

6. The ideal size of a first-year class at a particular college is 150 students. The college, knowing from past experience that, on the average, only 30 percent of those accepted for admission will actually attend, uses a policy of approving the applications of 450 students. Compute the probability that more than 150 first-year students attend this college.
   Answer: 0.06

7. A College would like to have 1050 freshmen. This college cannot accommodate more than 1060. Assume that each applicant accepts with probability 0.6 and that the acceptances can be modeled by Bernoulli trials. If the college accepts 1700, what is the probability that it will have too many acceptances?
   Answer: 0.0216

8. A die is rolled 420 times. What is the probability that the sum of the rolls lies between 1400 and 1550?
   Answer: 0.967

9. We load on a plane 100 packages whose weights are independent random variables that are uniformly distributed between 5 and 50 pounds. What is the probability that the total weight will exceed 3000 pounds?
   Answer: 0.0274

10. The National Health and Nutrition Examination Survey of 19881994 (NHANES III, A-1) estimated the mean serum cholesterol level for U.S. females aged 2074 years to be 204 mg/dl. The estimate of the standard deviation was approximately 44. Using these estimates as the mean m and standard deviation s for the U.S. population, consider the sampling distribution of the sample mean based on samples of size 50 drawn from women in this age group. What is the mean of the sampling distribution? The standard error?
    Answer: 204 and 6.2225

11. Suppose the 45 percent of the population favors a certain candidate in an upcoming election. If a random sample size 200 is chosen, find
    (a) the expected value and standard deviation of the number of members of the sample that favor the candidate
    (b) the probability that more than half the members of the sample favor the candidate.

    Answer: 0.0678

12. Suppose cars arrive at a parking lot at a rate of 50 per hour. Lets assume that the process is a Poisson random variable with $\lambda = 50$. Compute the probability that in the next hour the number of cars that arrive at this parking lot will be between 54 and 62.
    Answer: 0.2718

13. An insurance company has $25,000$ automobile policy holders. If the yearly claim of a policy holder is a random variable with mean 320 and standard deviation 540, approximate the probability that the total yearly claim exceeds 8.3 million.
    Answer: 0.00023.

14. In a Municipal election to select the post of the Chairman, suppose 50% of the population supports Arjun, 20% supports Sarika, and the rest are split between Nisar, John and Rohini. A poll asks 400 random people who they support. Use the central limit theorem to estimate the probability that less than 25% of those polled prefer Nisar, John and Rohini?
Answer: 0.0145.

15. If the mean and standard deviation of serum iron values for healthy men are 120 and 15 micrograms per 100 ml, respectively, what is the probability that a random sample of 50 normal men will yield a mean between 115 and 125 micrograms per 100 ml?
Answer: 0.9818.