MA2001D Mathematics III: Lecture 1

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

References

- S. M. Ross, *Introduction to Probability and statistics for engineers and Scientists*, Academic Press, New Delhi, 2014.
- R. A. Johnson, Miller and Freund's, *Probability and Statistics for Engineers*, Prentice Hall of Indian, New Delhi, 2011.

Let us ask few questions before we start the lecture.

- What is statistics ?
  Ans. Statistics is the art and science of extracting answer from data.
- What is statistic ?
  Ans. Statistic is a property of data (parameter that represent data in some form)(eg. Mean,).
- Why we study statistics ?
  Ans. Statistics help to make decision in an uncertain environment.
- What is the difference between statistics and probability ?

Introduction

What probability do?

Probability allows us to quantify the variability in the outcome of any experiment whose exact outcome cannot be predicted with certainty.

Example: Tossing of a fair (unbiased) coin.

Probability generally has great applications in games, in business to make probability-based predictions, and also probability has extensive applications in this new area of artificial intelligence.

Experiment

Statisticians use the word experiment to describe any process that generates a set of data.

Outcome

The results of the experiment are called outcome.

Types of experiment

1. Deterministic Experiments
2. Random Experiments

Deterministic Experiments

A deterministic experiment is one whose outcome may be predicted with certainty beforehand.

Example: Adding two numbers such as $2 + 3$.

An experiment is said to be random if its results can not be determined beforehand. In other words, a random experiment is one whose outcome is determined by chance.

Random Experiments

A random experiment is an experiment in which:

1. the set of all possible outcomes of the experiment is known in advance;
2. the outcome of a particular performance (trial) of the experiment cannot be predicted in advance;
3. the experiment can be repeated under identical conditions.

Example: How many red lights you encounter on the drive home.

Sample space

Sample space is the set Ω of all possible result or outcomes of an experiment.

Examples

- Tossing a coin, $\Omega = \{Head(H), Tail(T)\}$ -2 outcomes.
- Tossing 2 coins, $\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$ - 4 outcomes.
- Rolling a dice, then the sample space would be $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ - 6 outcomes, if we are interested in the number that shows on the top face. If we are interested only in whether the number is even or odd, then sample space is $\Omega_2 = \{even, odd\}$.

Event

An event is a subset or a portion of the sample space.

- When a Venn diagram is used to describe a random experiment, the sample space is the universal set and events are its subsets.
- If $A, B \subset \Omega \implies A \cup B, A \cap B$, and A^c and all other operation defined for subsets of Ω are events.

Mutually exclusive event

Two events A and B are mutually exclusive or disjoint if $A \cap B = \phi$, that is, if A and B have no elements in Mutually exclusive common.

Example: If we toss two coins, then the sample space is $\Omega = \{HH, HT, TH, TT\}$. Let event A is getting two heads, i.e. $A = \{HH\}$ and event B of getting two tails, i.e. $B = \{TT\}$. Note that $A \cap B = \phi$ implies A and B are Mutually exclusive events.

Mutually exhaustive event

Two events A and B are mutually exhaustive if $A \cup B = \Omega$.

Example

Consider the set of first 10 natural numbers. Check if the following defined events are exhaustive.

A: Selecting a prime number

B: Selecting a multiple of 2

C: Choosing a perfect square number

A, B and C are mutually exhaustive events.

Check mutually exhaustive or mutually exclusive!

Consider the experiment E of tossing a die. Suppose A is the event that an even number occur and let B be the event that a number greater than 3 shows.

– End –

MA2001D MATHEMATICS III: LECTURE 2

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

Axioms of Probability

Let Ω be a sample space of an experiment. The probability is a function $P : \Omega \to [0, 1]$ (or for each event A of the sample space Ω, we assume that a number $P(A)$) is defined and satisfies the following axioms:

1. **Axiom 1: Non-negativity.** $P(A) \geq 0, A \subset \Omega$.
2. **Axiom 2.** $P(\Omega) = 1$.
3. **Axiom 3: Countably Additive.** For any sequence of mutually exclusive events $A_1, A_2, \ldots$ (that is, events for which $A_i \cap A_j = \phi$ when $i \neq j$),
$$P\left(\cup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} P(A_i).$$
We refer to $P(A)$ as the probability of the event A.

Exercise 1

$P(A) + P(A^c) = 1$ for any event $A \subset \Omega$.

Exercise 2

$P(\phi) = 0$.

Exercise 3

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Exercise 4

If $A \subset B$, then $P(A) \leq P(B)$.

Equally likely event

Two event A and B are said to be equally likely if $P(A) = P(B)$.

Classical definition of Probability

Classical method of assigning probabilities is used for random experiments which result in a finite number of equally likely outcomes. Let $\Omega = \{w_1, w_2, \ldots, w_n\}$ be a finite sample space with n possible outcomes. In the classical method of assigning probabilities, the probability of an event A is given by

$$P(A) = \frac{\text{number of outcomes favorable to A}}{\text{total number of outcomes}} = \frac{|A|}{|\Omega|} = \frac{|A|}{n}.$$

Example

If we toss a coin then the sample space for this experiment is $\Omega = \{H, T\}$. All possible subsets of Ω are $\phi, \Omega, \{H\}, \{T\}$. Then $P(\phi) = 0$, $P(\Omega) = 1$, $P(\{H\}) = p$ and $P(\{T\}) = 1 - p$.

- If coin is fair,...
- If coin is biased....

CONDITIONAL PROBABILITY

The conditional probability of A given B, denoted by $P(A|B)$ is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0.$$

EXAMPLE:

A fair die is rolled.

1. Find the probability that the number rolled is a five, given that it is odd.
2. Find the probability that the number rolled is odd, given that it is a five.

Solution (1)

The sample space for this experiment is the set $\Omega = \{1, 2, 3, 4, 5, 6\}$ consisting of six equally likely outcomes. Let A denote the event ”a five is rolled" and let B denote the event "an odd number is rolled,” so that

$$A = \{5\} \text{ and } B = \{1, 3, 5\}.$$

Since $B = \{1, 3, 5\}$, $A \cap B = \{5\}$, we have $P(A \cap B) = 1/6$ and $P(B) = 1/2$. Thus,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{3}.$$

Solution (2)

$P(B|A) = 1$.

Independent Event

Two events A and B are independent under P (or simply independent) if

$$P(A \cap B) = P(A)P(B).$$

Exercise 1

Probability of Drawing Cards

A deck containing 52 cards is grouped into four suits of clubs, diamonds, hearts, and spades. Each of the clubs, diamonds, hearts, and spades have 13 cards each, which sum up to 52.

The 13 cards in each suit are ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, king. In these, the jack, the queen, and the king are called face cards.

Exercise 1

We can understand the card probability from the following examples.

- The probability of drawing a black card is P(Black card) =$26/52 = 1/2$.
- The probability of drawing a hearts card is P(Hearts)=$13/52 = 1/4$.
- The probability of drawing a face card is P(Face card)=$12/52 = 3/13$.
- The probability of drawing a card numbered 4 is $P(4) = 4/52 = 1/13$.
- The probability of drawing a red card numbered 4 is $P(4Red) = 2/52 = 1/26$.

Exercise 2

Consider the random experiment of tossing 3 coins. Find the probability of getting atleast 2 heads given that the 3^rd one is head.

Solution

Here $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$, event A is getting atleast 2 heads and event B is the 3^rd one is head. Hence

$$A = \{HHH, HHT, HTH, THH\} \text{ and } B = \{HHH, TTH, THH, TTH\}$$

Then $P(A) = 4/8 = 1/2$ and $P(B) = 4/8 = 1/2$. Also, $A \cap B = \{HHH, HTH, THH\}, P(A \cap B) = 3/8$. Thus

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{1/2} = \frac{3}{4}.$$

– End –

MA2001D MATHEMATICS III: LECTURE 3

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

Random Variable

A **random variable** X is function that associates a real number with each element in sample space, that is, $X : \Omega \to \mathbb{R}$.

Types of Random variable

- A **discrete random variable** is a random variable with a finite (or countably infinite) range.
- A **continuous random variable** is a random variable with an interval (either finite or infinite) of real numbers for its range.
  Examples of continuous random variables are age, height, weight etc.

Note

Random variables (r.v.) bare denoted by the capital letters X, Y, Z, etc., to distinguish them from their possible values given in lowercase x, y.

Examples

A random variable representing the number of automobiles sold at a particular dealership on one day would be discrete, while a random variable representing the weight of a person in kilograms (or pounds) would be continuous.

Example

Toss two fair coins, the sample space is

$$\Omega = \{HH, HT, TH, TT\}.$$

If we let X denotes the number of heads that appear, then X is a random variable taking on one of the values $0, 1$, and 2 with respective probabilities

$$P(X = 0) = P(T, T) = \frac{1}{4}$$

$$P(X = 1) = P((T, H), (H, T)) = \frac{2}{4}$$

$$P(X = 2) = P(H, H) = \frac{1}{4}.$$

Probability distribution

The probability distribution is a mathematical function that gives the probabilities of occurrence of different outcomes for an experiment.

DISCRETE RANDOM VARIABLE

PROBABILITY MASS FUNCTION

If X is a discrete random variable and let R_X be the range space of r.v. X. A real valued function $f_X : R_X \to \mathbb{R}$ is said to be probability mass function (p.m.f.) if

$$f_X(x) = P(X = x) = P(\{s \in \Omega : X(s) = x\}), \quad x \in R_X.$$

DISCRETE RANDOM VARIABLE

PROPERTIES:

1. $f_X(x) \geq 0$.
2. $\sum_x f_X(x) = 1$.

Discrete Random Variable

For discrete random- variable, a knowledge of the probability mass function enables us to compute probabilities of arbitrary events. In fact, if A is an event then

$$P(x \in A) = \sum_{x \in A} f_X(x).$$

Exercise

An experiment consists of three independent tosses of a fair coin. Let

$$\begin{aligned} X &= \text{The number of heads} \\ Y &= \text{The number of head runs,} \\ Z &= \text{ The length of head runs,} \end{aligned}$$

a head run being defined as consecutive occurrence of at least two heads, its length then being the number of heads occurring together in three tosses of the coin.

Find the probability function of $(i)X,\ (ii)Y, (iii)Z, (iv)X+Y$ and $(v)XY$.

Cumulative Distribution Function (or distribution function)

Let X be a random variable. Then its cumulative distribution function (cdf) is defined by $F_X(x)$, where

$$F_X(x) = P(X \leq x) = P(\{s \in \Omega : X(s) \leq x\}), \quad x \in \mathbb{R}.$$

If X is discrete r.v. and $f(x)$ is its p.m.f., then

$$F_X(x) = \sum_{t \leq x} f(t).$$

PROPERTIES:

If F is a distribution function of the random variable X and if $a < b$. Then

- $P(a < X \leq b) = F(b) - F(a)$.

$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$.

- $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$.

- $P(a < X < b) = F(b) - F(a) - P(X = b)$.

$0 \leq F_X(x) \leq 1$

If $x \leq y$ then $F_X(x) \leq F_X(y)$.

$\lim_{x \to \infty} F_X(x) = 1$ and $\lim_{x \to -\infty} F_X(x) = 0$.

– End –

MA2001D Mathematics III: Lecture 4

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

Cumulative Distribution Function (or distribution function)

Let X be a random variable. Then its cumulative distribution function (cdf) is defined by $F_X(x)$, where

$$F_X(x) = P(X \leq x) = P(\{s \in \Omega : X(s) \leq x\}), \quad x \in \mathbb{R}.$$

If X is discrete r.v. and $f_X(x)$ is its p.m.f., then

$$F_X(x) = \sum_{t \leq x} f_X(t).$$

PROPERTIES:

If F is a distribution function of the random variable X and if $a < b$. Then

- $P(a < X \leq b) = F(b) - F(a)$.

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

- $P(a \leq X \leq b) = F(b) - F(a) + P(X = a)$.

- $P(a < X < b) = F(b) - F(a) - P(X = b)$.

$0 \leq F_X(x) \leq 1$

If $x \leq y$ then $F_X(x) \leq F_X(y)$.

$\lim_{x\to\infty} F_X(x) = 1$ and $\lim_{x\to-\infty} F_X(x) = 0$.

Exercise

Find the value of k for which the the function

$$f(x) = k(2x - 1),$$

is a pmf of a r.v. X where the range set $R_X = \{1, 2, \ldots, 12\}$.

Ans. $k = 1/144$.

Find the CDF of X.

Continuous Random Variable

Probability Density Function

Let X be a continuous random variable. Then the probability density function (p.d.f.) f_X (or simply f) of the r.v. X is defined as

$$f_X(x) = \lim_{\delta x \to 0} \frac{P(x \leq X \leq x + \delta x)}{\delta x}.$$

Note that:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx.$$

Continuous Random Variable

Properties

- $f(x) \geq 0, \ -\infty < x < \infty.$
- $\int_{-\infty}^{\infty} f(x) = 1.$

Cumulative Distribution Function (or distribution function) of continuous r.v.

Let X be a random variable. Then its cumulative distribution function (cdf) is defined by $F_X(x)$, where

$$F_X(x) = P(X \leq x) = P(\{s \in \Omega : X(s) \leq x\}), \quad x \in \mathbb{R}.$$

If X is continuous r.v. and $f_X(x)$ is its p.d.f., then

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt.$$

PROPERTIES:

- $0 \leq F(x) \leq 1$.
- $P(X = x) = 0$.
- $P(a < X < b) = P(a \leq X < b) = P(a < X \leq b) = P(a \leq X \leq b) = F(b) - F(a)$.
- $P(X > x) = 1 - F(x)$.
- $P(X \leq x) = P(X < x) = F(x)$.
- $F(\infty) = 1$ and $F(-\infty) = 0$.

HOW TO FIND P.D.F. OF GIVEN C.D.F.

Let $F(x)$ be a given distribution function (c.d.f.) of continuous r.v. X then p.d.f. of continuous r.v. X is given by

$$\frac{d}{dx}F(x) = f(x).$$

F is a non-decreasing function of x.

Exercise

For what value of the constant c the real valued function f given by

$$f(x) = \begin{cases} c, & \text{if } a \leq x \leq b; \\ 0, & \text{otherwise} \end{cases}$$

Ans.

$c = 1/(b-a)$.

Exercise

The diameter of an electric cable; say X, is assumed to be a continuous random variable with p.d.f. $f(x) = 6x(1-x)$, $0 \leq x \leq 1$. Determine a number b such that $P(X < b) = P(X > b)$.

Hint:

Given that

$$P(X < b) = \int_0^b f(x)\,dx = P(X > b) = \int_b^1 f(x)\,dx.$$

After solving, we get $b = 1/2$.

EXERCISE

PROBLEM

A shipment of 8 similar computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of two computers, find the probability distribution function for the no. of defective.

HINT:

Consider the random variable X: the number of defective computer purchase by school. Thus X can take $\{0, 1, 2\}$ and X is a discrete r.v. Hence compute $P(X = 0), P(X = 1)$, and $P(X = 2)$.

– End –

MA2001D Mathematics III: Lecture 5

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

Mathematical Expectation

Let X be a random variable (r.v.) with p.d.f. (p.m.f.) $f(x)$. Then its mathematical expectation, denoted by $E(X)$ is given by

$$\bullet E(X) = \sum_x x f(x), \qquad \text{for discrete r.v. } X$$

and

$$\bullet E(X) = \int_{-\infty}^{\infty} x f(x) dx \text{ for continuous r.v. } X$$

provided the righthand integral or series is absolutely convergent, i.e., provided

$$\sum_x |x| f(x) < \infty \text{ and } \int_{-\infty}^{\infty} |x| f(x)\, dx < \infty.$$

If we have an infinite sum, it needs to be well-defined. Thus we assume last two conditions.

The expectation describe the average value.

This is also known as the mean, or average or first moment of X and is usually denoted by μ.

Expectation of a function of a Random variable

Consider a r.v. X with p.d.f. (p.m.f.) f. If $g(.)$ is a function such that $g(X)$ is a r.v. and $E[g(X)]$ exists (i.e., is defined), then

$$E(g(X)) = \sum_x g(x)f(x), \qquad \text{for discrete r.v. } X$$

and

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx, \qquad \text{for continuous r.v. } X$$

For constant a, b, and c we have

PROPERTIES:

If X is a random variable, then

- $E(ag(X)) = aE(g(X))$.
- $E(g(X) + b) = E(g(X)) + b$.
- $E(c) = c$.
- $E[X - \mu] = E[X] - \mu = \mu - \mu = 0$.

CONSEQUENCES:

- $E(aX) = aE(X)$.
- $E(X + b) = E(X) + b$.
- $E(aX + b) = aE(X) + b$.

Exercise

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

$$f(x) = \begin{cases} 2x, & \text{if } 0 \leq x \leq 1; \\ 0, & \text{otherwise .} \end{cases}$$

Find the expected value of X.

Solution

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} xf(x)\,dx \\ &= \int_0^1 x(2x)\,dx \\ &= \int_0^1 2x^2\,dx = \frac{2}{3}. \end{aligned}$$

Variance

If X is a random variable, then variance of r.v. X is defined as

$$\sigma^2 = V(X) = E((X - E(X))^2).$$

Alternate definition of Variance

$$\sigma^2 = V(X) = E(X^2) - (E(X))^2.$$

Standard Deviation

The standard deviation of an r.v. X is

$$\sigma = \sqrt{\sigma} = \sqrt{V(X)}.$$

For constant a, b, and c we have

Properties

- $V(c) = 0$.
- $V(aX) = a^2V(X)$.
- $V(aX + b) = a^2V(X)$.

Exercise

The distribution function of a random variable X is given by

$$F(x) = 1 - (1+x)e^{-x},\ x \geq 0.$$

Find the density function, mean, variance of X.

Note that given random variable X is continuous. Then the probability density function is

$$f(x) = F'(x) = xe^{-x},\ x \geq 0.$$

Also, $E(X) = 2, E(X^2) = 6$, and $V(X) = 2$.

Joint Probability Mass function

Two random variables X and Y are said to be jointly distributed if they are defined on the same probability space.

Joint Probability Mass function

Let X and Y be random variables on a sample space Ω with respective image sets $R_X = \{x_1, \ldots, x_n\}$ and $R_Y = \{y_1, \ldots, y_m\}$. A real valued function $f : R_X \times R_Y \to \mathbb{R}$ defined by

$$f(x, y) = P(X = x, Y = y) = P(X = x \cap Y = y), \; \forall (x, y) \in R_X \times R_Y,$$

is called the joint probability mass function of X and Y.

Properties

- $f(x, y) \geq 0$.
- $\sum_{x \in R_X} \sum_{y \in R_Y} f(x, y) = 1$.

– End –

MA2001D Mathematics III: Lecture 6

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

1. JOINT PROBABILITY FUNCTION
   - Joint Probability Mass function
   - Joint Probability Density function
2. MARGINAL PROBABILITY FUNCTION
   - Marginal Probability Mass function
   - Marginal Probability density function
3. CONDITIONAL PROBABILITY FUNCTION
4. JOINT PROBABILITY DISTRIBUTION FUNCTION
5. MARGINAL DISTRIBUTION FUNCTION
6. INDEPENDENT RANDOM VARIABLE

Joint Probability function

Two random variables X and Y are said to be jointly distributed if they are defined on the same probability space.

Joint Probability Mass function

Let X and Y be discrete random variables on a sample space Ω with respective image sets $R_X = \{x_1, \dots, x_n\}$ and $R_Y = \{y_1, \dots, y_m\}$. A real valued function $f : R_X \times R_Y \to \mathbb{R}$ defined by

$$f_{XY}(x, y) = P(X = x, Y = y) = P(X = x \cap Y = y), \ \forall (x, y) \in R_X \times R_Y,$$

is called the joint probability mass function of X and Y.

Properties

- $f(x, y) \geq 0$.
- $\sum_{x \in R_X} \sum_{y \in R_Y} f(x, y) = 1$.

Joint Probability Density function

Let X and Y be continuous random variables on a sample space Ω. A real valued function $f : R_X \times R_Y \to \mathbb{R}$ defined by

$$f_{XY}(x,y) = \frac{\lim_{\delta x\to 0,\delta y\to 0} P(x \leq X \leq x+\delta x, y \leq Y \leq y+\delta y)}{\delta x\,\delta y}$$

is called the joint p.d.f.

Note that:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)dxdy.$$

PROPERTIES

- $f(x,y) \geq 0.$
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dxdy = 1.$

EXERCISE

For what value of constant k the function given by

$$f(x,y) = \begin{cases} k\,x\,y, & \text{if } x = \{1,2,3\},\ y = \{1,2,3\}; \\ 0, & \text{otherwise} \end{cases}$$

is a joint p.m.f. of r.v. X and Y.

ANS.

$k = 1/36$.

Marginal Probability function

Suppose the joint distribution of two random variables X and Y is given then the probability distribution of X is determined as follows:

Marginal Probability Mass function

Let X and Y be discrete random variables on a sample space Ω with respective image sets $R_X = \{x_1, \ldots, x_n\}$ and $R_Y = \{y_1, \ldots, y_m\}$. Let $f(x, y)$ be a joint probability mass function. The probability distribution of X is determined as follows

$$
\begin{aligned}
f_X(x) = P(X = x) &= \sum_{y \in R_Y} f(x, y) \\
&= P(X = x \cap Y = y_1) + \cdots + P(X = x \cap Y = y_m).
\end{aligned}
$$

Similarly, the probability distribution of Y is determined as follows

$$
\begin{aligned}
f_Y(x) = P(Y = y) &= \sum_{x \in R_X} f(x, y) \\
&= P(X = x_1 \cap Y = y) + \cdots + P(X = x_n \cap Y = y).
\end{aligned}
$$

Marginal Probability density function

Let X and Y be continuous random variables on a sample space Ω. Let $f(x, y)$ be a joint probability density function. The probability distribution of X is determined as follows

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\, dx.$$

Conditional Probability function

Let X and Y be random variables on a sample space Ω. Let $f(x, y)$ be a joint probability function. The conditional probability function of X given Y is define as

$$P(X = x|Y = y) = \frac{f(x, y)}{f_Y(y)}$$

and

$$P(Y = y|X = x) = \frac{f(x, y)}{f_X(x)}$$

where f_Y is a marginal probability function of r.v. Y.

Exercise

Let (X, Y) be a two dimensional random variable and the joint p.m.f. is given by

$$f(x,y) = \begin{cases} \frac{1}{21}(x+y), & \text{if } x = \{1,2,3\},\ y = \{1,2\}; \\ 0, & \text{otherwise.} \end{cases}$$

Then compute $P(Y|X=2)$, $P(X|Y=2)$, $P(X \leq 2)$, $P(X \leq 2, Y < 3)$ and $P(X \leq 2|Y < 3)$.

ANS.

- $f(Y=2) = 12/21$, $P(X|Y=2) = (x+2)/12$, $f(X=2) = 7/21$.
- $P(X \leq 2) = P(X=1) + P(X=2) = 1$
- $P(X \leq 2, Y < 3) = P(X=1, Y<3) + P(X=2, Y<3) = P(X=1, Y=1) + P(X=1, Y=2) + P(X=2, Y=1) + P(X=2, Y=2) = 12/21$.
- $P(X \leq 2 | Y < 3) = P(X \leq 2, Y < 3)/P(Y < 3) = 1$.

Joint probability distribution function

Let (X, Y) be a two dimensional random variable then their joint distribution function denoted by $F_{XY}(x, y)$ and defined as

$$F_{XY}(x, y) = P(X \leq x, Y \leq y), \forall (x, y) \in R_X \times R_Y.$$

Properties

- $F_{XY}(x, y) \geq 0$.
- $F_{XY}(-\infty, y) = 0 = F(x, -\infty)$.
- $F_{XY}(-\infty, \infty) = 1$.
- $\dfrac{\partial^2 F_{XY}}{\partial x \partial y} = f(x, y)$.

Joint probability distribution function

For Discrete Random Variable

$$F(x, y) = \sum_{r \le x} \sum_{s \le y} f(x, y).$$

For Continuous Random variable

$$F(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dxdy.$$

Marginal Distribution Function

Let (X, Y) be a two dimensional random variable and $f(x, y)$ be a joint probability function. Then it is possible to obtain the individual distribution functions, $F_X(x)$ and $F_Y(y)$ which are termed as marginal distribution function of X and Y respectively w.r.t. the joint probability distribution function $F_{XY}(x, y)$,

$$F_X(x) = P(X \leq x) = P(X \leq x, Y < \infty) = \lim_{y\to\infty} F_{XY}(x, y) = F_{XY}(x, \infty).$$

$$F_Y(y) = P(Y \leq y) = P(X < \infty, Y \leq y) = \lim_{x\to\infty} F_{XY}(x, y) = F_{XY}(\infty, y).$$

Independent Random Variable

Two r.v. X and Y with joint p.d.f. $f_{XY}(x,y)$ and marginal p.d.f.'s f_X and f_Y are said to be independent if and only if

$$f_{XY}(x,y) = f_X(x) f_Y(y)$$

Two r.v. X and Y are said to be independent if and only if

$$F_{XY}(x,y) = F_X(x)\, F_Y(y).$$

For discrete r.v.

Two random variables X and Y are said to be independent if

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

– End –

MA2001D Mathematics III: Lecture 7

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

1. Moments of random variable

2. Moment Generating Function

Recall

Let X be a random variable (r.v.) with p.d.f. (p.m.f.) $f(x)$. Then its mathematical expectation, denoted by $E(X)$ is given by

Expectation for discrete r.v.

$$E(X) = \sum_{x} x f(x).$$

Provided series is absolutely convergent, i.e., $\sum_x |x| f(x) < \infty$.

Expectation for continuous r.v.

$$E(X) = \int_{-\infty}^{\infty} x f(x)\, dx.$$

Provided integral is absolutely convergent, i.e., $\int_{-\infty}^{\infty} |x| f(x)\, dx < \infty$.

r^{th} MOMENT (ABOUT ORIGIN)

RECALL:

$$E(g(X)) = \sum_x g(x)f(x), \qquad \text{for discrete r.v. } X$$

PARTICULAR CASE:

If $g(X) = X^r$ where r is positive real number. We get

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x)dx \text{ or } E(X^r) = \sum_x x^r f(x).$$

which is defined as μ_r', the r^{th} moment (about origin) of the probability distribution.

Thus $\mu_r^{'}$ (about origin) $= E(X^r)$. In particular

- $\mu = \mu_1^{'} = E(X)$
- $\mu_2^{'} = E(X^2)$.

r^{th} CENTRAL MOMENT OR MOMENT (ABOUT MEAN)

PARTICULAR CASE:

If $g(X) = (X - E(X))^r$ where r is a positive real number. We get

$$\mu_r = E((X - E(X))^r) = \int_{-\infty}^{\infty} (x - \mu)^r f(x) dx$$

or

$$\mu_r = E((X - E(X))^r) = \sum_x (x - \mu)^r f(x).$$

In particular, the first central moment is always zero, since $\mu_1 = E(X - E(X)) = 0$ and the second central moment is the variance,

$$\mu_2 = \sigma = E((X - E(X))^2).$$

If X and Y have a joint p.m.f. (or p.d.f.), then

$$E[g(X,Y)] = \sum_x \sum_y g(x,y)f(x,y)$$

or

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\,dxdy.$$

Addition theorem of Expectation

If X and Y are random variables then

$$E(X+Y) = E(X) + E(Y)$$

provided all the expectations exist (or $E(X)$ and $E(Y)$ are finite).

$$\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)\,dydx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)\,dxdy \\
&= \int_{-\infty}^{\infty}xf_X(x)dx + \int_{-\infty}^{\infty}yf_Y(y)dy \\
&= E(X) + E(Y).
\end{aligned}$$

We may show by a simple induction proof that if $E(X_i)$ is finite for all $i = 1, \ldots, n$, then

RESULT:

$$E(X_1 + \cdots + X_n) = E(X_1) + \cdots + E(X_n).$$

The moment generating function $M_X(t)$ of the random variable X is defined for all real values t by

$$M_X(t) = E(e^{tX}).$$

For discrete r.v.

$$M_X(t) = E(e^{tX}) = \sum_x e^{tX} f(x).$$

For continuous r.v.

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tX} f(x)\, dx.$$

Observation

$$M_X(t) = E(e^{tX}) = E\left(1 + tX + \frac{(tX)^2}{2!} + \dots\right)$$
$$= 1 + tE[X] + \frac{t^2}{2!}E(X^2) + \cdots = \sum_{n=0}^{\infty} \frac{(t)^n}{n!} E(X^n).$$

Note that

$$\mu_r^{'} = \frac{d^r}{dt^r}\{M_X(t)\}|_{t=0}.$$

THE ABOVE RESULT ALSO IMPLIES THAT THE MOMENT GENERATING FUNCTION UNIQUELY DETERMINES THE DISTRIBUTION. THAT IS, THERE EXISTS A ONE-TO-ONE CORRESPONDENCE BETWEEN THE MOMENT GENERATING FUNCTION AND THE DISTRIBUTION FUNCTION OF A RANDOM VARIABLE.

PROPERTIES

- $M_{cX}(t) = M_X(ct) = E(e^{cXt})$.
- $M_{aX+b}(t) = e^{tb} M_X(at)$.

Exercise

Let X be a random variable and the p.m.f. of X is given by

$$p_X(x) = \frac{3}{4}\left(\frac{1}{4}\right)^x, \, x = 0, 1, 2, \ldots.$$

Find the MGF of X.

Hint:

$$M_X(t) = E(e^{tX}) = \sum_{x=0}^{\infty} e^{tx} \frac{3}{4}\left(\frac{1}{4}\right)^x = \frac{3}{4}\sum_{x=0}^{\infty}\left(\frac{e^t}{4}\right)^x.$$

If $e^t/4 < 1$, then

$$M_X(t) = \frac{3}{4}\left[\frac{1}{1 - e^t/4}\right].$$

$$E(X) = M_X^{'}(0) = \frac{3/4}{4(1 - 1/4)^2} = \frac{1}{3}.$$

Exercise

The value of a piece of factory equipment after three years of use is $100(0.5)^x$ where X is a random variable having moment generating function

$$M_X(t) = \frac{1}{1-2t} \text{ for } t < \frac{1}{2}.$$

Calculate the expected value of this piece of equipment after three years of use.

Ans.

$$f_X(x) = 100(0.5)^x,\ E(X) = \frac{100}{1-2(\ln(0.5))} = 41.906.$$

– End –

MA2001D MATHEMATICS III: LECTURE 7

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

1. BERNOULLI DISTRIBUTION

2. BINOMIAL DISTRIBUTION

BERNOULLI DISTRIBUTION

BERNOULLI TRIALS

A random experiment is called a Bernoulli Trials if

- There are only two possible outcomes.
- The probability of success is the same for each trial.
- The outcomes from different trials are independent.

Bernoulli distribution

Bernoulli distribution

A r.v. X is said to have Bernoulli random variable with parameter p, shown as $X \sim Bernoulli(p)$, if X has only two possible values 0 and 1 (or Success and Failure), and

$$P(X=1)=p, P(X=0)=1-p,$$

where p is the probability that the trial is success. The p.m.f. of X is given by

$$f(x)=p^x(1-p)^{1-x};\ x=0,1.$$

BERNOULLI DISTRIBUTION

PROPERTIES:

- $\mu = E(X) = 1 \cdot p + 0 \cdot (1-p) = p$.
- $\sigma^2 = V(X) = p(1-p)$.
- $M_X(t) = E(e^{tX}) = (1-p) + pe^t$.
- $E(X^r) = \mu_r' = p,\ r = 1, 2, \ldots$

Binomial distribution

Binomial distribution

For some $n \in \mathbb{N}$ and $0 < p < 1$, the probability mass function of a Binomial random variable X with parameters n and p is given by

$$f(x;n;p) = {}^nC_x p^x q^{n-x};\ x = 0, 1, \ldots, n;\ q = 1 - p$$

• The number of successes in n Bernoulli trials with probability of success equal to p, follows the Binomial probability distribution.

• For $n = 1$, i.e. a single experiment, the binomial distribution is a Bernoulli distribution.

Binomial distribution

Remarks

$$\sum_{x=0}^{n} f(x) = \sum_{x=0}^{n} {}^{n}C_{x} p^{x} q^{n-x} = (q+p)^{n} = 1.$$

Properties

- $E(X) = np$.
- $Var(X) = np(1-p)$.
- $M_X(t) = ((1-p) + pe^{t})^{n}$.

EXERCISE

EXAMPLE

Ten coins are thrown simultaneously. Find the probability of getting at least seven heads.

SOLUTION

Let p= probability of getting a head=$1/2$ and q= probability of getting a tail=$1/2$.
The probability of getting x heads in a random throw of 10 coins is

$$f(x) = {}^{10}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x}.$$

– End –

MA2001D MATHEMATICS III: LECTURE 9

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

EXERCISE

EXAMPLE

Ten fair coins are thrown simultaneously. Find the probability of getting at least seven heads.

SOLUTION

Let p= probability of getting a head=$1/2$ and q= probability of getting a tail=$1/2$.
The probability of getting x heads in a random throw of 10 coins is

$$P(X = x) = f(x) = {}^{10}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x}.$$

The probability of getting at least seven heads =
$P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$.

If X and Y are independent random variable with MGF $M_X(t)$ and $M_Y(t)$, then

$$M_{X+Y}(t) = M_X(t)M_Y(t).$$

Additive Property of Binomial Distribution:

Let $X \sim B(n_1, p_1)$ and $Y \sim B(n_2, p_2)$ are independent random variable. Then

$$M_X(t) = (q_1 + p_1 e^t)^{n_1} \text{ and } M_Y(t) = (q_2 + p_2 e^t)^{n_2}.$$

What is the distribution of $X + Y$?

If $p_1 = p_2 = p$, then we get $X + Y \sim B(n_1 + n_2, p)$.

GEOMETRIC DISTRIBUTION

OBSERVATION

Conduct a sequence of Bernoulli trials until first success. The sample space is

$$\Omega = \{S, FS, FFS, FFFS, \ldots\}$$

and random variable X=The number of trials needed for the first success=$\{1, 2, 3, \ldots\}$.

Geometric Distribution

Definition:

A discrete random variable X is said to follow the Geometric distribution with parameter p, written as $X \sim Geometric(p)$, if its p.m.f. is given by

$$f_X(x) = p(1-p)^{x-1}, \text{ for } x = 1, 2, 3, \ldots.$$

EXAMPLE 1

Suppose a dice is repeatedly rolled until "3" is obtained. Then the probability of getting "3" is $p = 1/6$ and the random variable, X, can take on a value of $1, 2, 3, \ldots,$ until the first success is obtained.

Geometric Distribution

Properties:

- $E(X) = 1/p$.
- $Var(X) = \dfrac{1-p}{p^2}$.
- $M_X(t) = \dfrac{pe^t}{1-(1-p)e^t}$.
- $F_X(x) = 1-(1-p)^x$ for $x = 1, 2, 3, \ldots$.

– End –

MA2001D Mathematics III: Lecture 10

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

Geometric Distribution

Recall

Conduct a sequence of Bernoulli trials until first success. The sample space is

$$\Omega = \{S, FS, FFS, FFFS, \dots\}$$

and random variable X=The number of trials needed for the first success=$\{1, 2, 3, \dots\}$.

GEOMETRIC DISTRIBUTION

RECALL:

A discrete random variable X is said to follow the Geometric distribution with parameter p, written as $X \sim Geometric(p)$ (or $X \sim G(p)$), if its p.m.f. is given by

$$f_X(x) = p(1-p)^{x-1}, \text{ for } x = 1, 2, 3, \ldots.$$

VERIFY

$$\sum_{x=1}^{\infty} f_X(x) = 1.$$

Geometric Distribution

Example 1

Suppose a dice is repeatedly rolled until "3" is obtained. Then the probability of getting "3" is $p = 1/6$ and the random variable, X, can take on a value of $1, 2, 3, \ldots$, until the first success is obtained.

Geometric Distribution

Properties:

- $E(X) = 1/p$.
- $Var(X) = \dfrac{1-p}{p^2}$.
- $M_X(t) = \dfrac{pe^t}{1-(1-p)e^t}$.
- $F_X(x) = 1-(1-p)^x$ for $x = 1, 2, 3, \ldots$.

Geometric Distribution

Example 2

If a patient is waiting for a suitable blood donor and the probability that the selected donor will be a match is 0.2, then find the expected number of donors who will be tested till a match is found including the matched donor.

Sol.

As we are looking for only one success this is a geometric distribution. Here $p = 0.2$ and $E(X) = 1/p = 5$.

Geometric Distribution

Example

A light bulb manufacturing factory finds 3 in every 60 light bulbs defective. What is the probability that the first defective light bulb will be found when the 6^{th} one is tested?

Solution

As the probability of the first defective light bulb needs to be determined hence, this is a geometric distribution. Here $p = 3/60$ and

$$P(X = 6) = (1 - p)^{x-1} p = 0.0386.$$

Hypergeometric Distribution

A hypergeometric experiment is a statistical experiment that has the following properties:

- A sample of size n is randomly selected **without replacement** from a population of N items.
- In the population, K items can be classified as successes, and $N - K$ items can be classified as failures.

Hypergeometric Distribution

Definition

A discrete random variable X is said to follow the Hypergeometric distribution with parameters N, n, and K, written as $X \sim H(N, n, K)$

$$f(x; N, n, K) = P(X = x) = \frac{{}^{K}C_{x} \, {}^{N-K}C_{n-x}}{{}^{N}C_{n}},$$

where K is the number of success in the population, N is the population size, n is the sample size.

- Note that it would not be a binomial experiment.
- A binomial experiment requires that the probability of success be constant on every trial.
- With the Hypergeometric experiment, the probability of a success changes on every trial.
- If the number of elements in the sample (n) are much smaller than in the population, the

$$\text{Hypergeometric distribution} \approx \text{ Binomial Distribution.}$$

HYPERGEOMETRIC DISTRIBUTION

PROPERTIES:

- $E(X) = \frac{nk}{N}$.
- $Var(X) = n \cdot \frac{K}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1}$.

Hypergeometric Distribution

Exercise

Suppose we randomly select 5 cards without replacement from an ordinary deck of playing cards. What is the probability of getting exactly 2 red cards (i.e., hearts or diamonds)?

Hint:

N is 52, $K = 26$; since there are 26 red cards in the deck, $n = 5$; since we randomly select 5 cards from the deck, and $x = 2$; since 2 of the cards we select are red. Thus,

$$P(X = 2) = \frac{{}^{26}C_2 \; {}^{26}C_3}{{}^{52}C_5} = 0.325.$$

Hypergeometric Distribution

Exercise

A box contains 50 light bulbs of which 5 are defective and 45 are not. a quality control Inspector randomly samples 4 bulbs without replacement. Let X the number of defective bulbs selected. Find the probability mass function, $f(x)$, of the discrete random variable X.

Hypergeometric Distribution

Applications:

- One of the most common applications of the hypergeometric distribution is in industrial quality control, such as calculating probabilities for defective parts produced in a factory.

– End –

MA2001D Mathematics III: Lecture 11

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

1. POISSON DISTRIBUTION

Exercise

If X and Y are two independent random variable, then

$$E(XY) = E(X)E(Y).$$

Relationship between Binomial and Poisson distribution

Poisson distribution is a limiting case of the binomial distribution under the following conditions:

- n, the number of trials is indefinitely large, i.e. $n \to \infty$.
- the constant probability of success for each trial is indefinitely small, i.e. $p \to 0$.
- $np = \lambda$

Poisson Distribution

Definition:

A discrete random variable X that takes on one of the values $0, 1, 2, \ldots$ is said to be a Poisson random variable with parameter λ if, for some $\lambda > 0$, p.m.f. of X is given by

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

We shall use the notation $X \sim P(\lambda)$ to denote that X is a Poisson variate with parameter λ.

Verify

$$\sum_{x=0}^{\infty} f_X(x) = 1.$$

Distribution Function (c.d.f.)

$$F(X) = P(X \leq x) = \sum_{r=0}^{x} f(r) = e^{-\lambda} \sum_{r=0}^{x} \frac{\lambda^r}{r!}, \quad x = 0, 1, 2 \ldots .$$

Poisson Distribution

Properties:

- $E(X) = \lambda$.
- $E(X^2) = \lambda(\lambda + 1)$.
- $Var(X) = \lambda$.
- $M_X(t) = e^{(e^t - 1)\lambda}$.

Poisson Distribution

Additive Property of Poisson distribution

Let $X \sim P(\lambda_1)$ and $Y \sim P(\lambda_2)$ are independent random variable. Then

$$M_X(t) = e^{\lambda_1(e^t-1)} \text{ and } M_Y(t) = e^{\lambda_2(e^t-1)}$$

Thus,

$$M_{X+Y}(t) = M_X(t)M_Y(t) = e^{\lambda_1(e^t-1)}e^{\lambda_2(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)}.$$

Hence $X + Y \sim P(\lambda_1 + \lambda_2)$.

Poisson Distribution

Exercise

Let X and Y are two random variables with moment generating function $M_(t) = e^{2(e^t-1)}$ and $M_Y(t) = e^{3(e^t-1)}$. Which of the following are correct?

1. $X + Y \sim P(5)$.
2. $E(X+Y) = 5$.

1. False
2. True

POISSON DISTRIBUTION

EXERCISE

Six coins are tossed $6,400$ times. Using the Poisson distribution, find the approximate probability of getting six heads r times.

HINT:

The probability of obtaining six beads in one throw of six coins (a single trial), is $p = 1/2^6$. Thus

$$P(X = r) = \frac{e^{-100}100^r}{r!}, \quad r = 0, 1, 2, \ldots.$$

– End –

MA2001D Mathematics III: Lecture 12

Instructor: Vibhuti Arora

Department of Mathematics
National Institute of Technology Calicut

1. Uniform Distribution

2. Gamma Distribution

Continuous Probability Distribution

Uniform Distribution

Definition:

A continuous random variable X is said to follow Uniform distribution over the interval (a, b) if the probability density function of X is given by

$$f_X(x) = \frac{1}{b-a}, \quad a \leq x \leq b.$$

and it is written as $X \sim Uniform(a, b)$.

Uniform Distribution

Verify

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

Distribution Function (c.d.f.)

Since $F(x) = \int_{-\infty}^{x} f(x)dx$. Thus,

$$F(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & \text{if } a < x < b; \\ 1, & \text{if } x \geq b. \end{cases}$$

Uniform Distribution

Properties:

1. $E(X) = \dfrac{a+b}{2}$.
2. $Var(X) = \dfrac{(a-b)^2}{12}$.
3. $M_X(t) = \dfrac{e^{bt} - e^{at}}{(b-a)t}$.

Exercise

Suppose that a large conference room for a certain company can be reserved for no more than 4 hours. However, the use of the conference room is such that both long and short conferences occur quite often. In fact, it can be assumed that length X of a conference has a uniform distribution on the interval $[0,4]$.

- What is the probability density function?
- What is the probability that any given conference lasts at least 3 hours?

$$f(x) = \begin{cases} \frac{1}{4}, & \text{if } 0 \leq x \leq 4; \\ 0, & \text{otherwise .} \end{cases}$$

$$P(X \geq 3) = \int_3^4 \frac{1}{4} dx = \frac{1}{4}$$

Gamma Distribution

The factorial function

Suppose $f(x)$ denotes the factorial function

$$f(x) = x!, \quad x = 0, 1, 2, \ldots.$$

The gamma function

A continuous extension of the factorial function is the gamma function.

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy \quad \alpha > 0.$$

Gamma Distribution

Properties

- $\Gamma(1) = 1$.
- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 0$.
  Proof: Exercise! (Use by-part).
- $\Gamma(n) = (n-1)!$.
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Gamma Distribution

Standard Gamma Distribution

Let X be a continuous random variable is said to follow standard gamma distribution, $X \sim Gamma(\alpha, 1)$, it the probability density function is given by

$$f(x) = \frac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)}, \quad x \geq 0.$$

Gamma Distribution

A continuous random variable X is said to follow gamma distribution with parameter $\alpha, \beta > 0$, $X \sim Gamma(\alpha, \beta)$, if the probability density function is given by

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, \quad x \geq 0.$$

Gamma Distribution

Verify:

$\int_{-\infty}^{\infty} f(x)\,dx = 1.$

Properties:

- $M_X(t) = \left(\dfrac{1}{1-\beta t}\right)^{\alpha}$, for $t < 1/\beta$.
- $E(X) = \mu = \alpha\beta$.
- $Var(X) = \alpha\beta^2$.

– End –