



**CHHATRAPATI SHAHU JI MAHARAJ  
UNIVERSITY, KANPUR**

**प्रश्नBANK**  
Bridge of Academic Novelties in Knowledge



**KANPUR UNIVERSITY'S**  
**QUESTION**  
**BANK**

**B.SC. IV SEM**

**Based On  
NEP  
2020**

**ZOOLOGY**  
**GENE TECHNOLOGY,**  
**IMMUNOLOGY AND**  
**COMPUTATIONAL**  
**BIOLOGY**

- 400+ MCQs
- Brief and Intensive Notes

**DR. NEERAJ VERMA**

**DR. SHAILENDRA PRATAP SINGH**

**B.Sc. Semester – IV****Subject: Zoology****Course Code - B050401T****Course Title - Gene Technology, Immunology  
and Computational Biology****Unit I**

Principles of Gene Manipulation

- Recombinant DNA Technology
- Selection and identification of recombinant cells
- Restriction enzymes, DNA modifying enzymes, Cloning Vectors, Ligation

Recombinant DNA technology, also known as genetic engineering or gene splicing, is a powerful technique that allows scientists to manipulate and combine DNA molecules from different sources. The fundamental principle of recombinant DNA technology involves the creation of new combinations of DNA sequences by cutting and joining DNA molecules.

This includes a number of methodologies. Desired genes or DNA sequences are cut into fragments and then introduced into host cells. 1<sup>st</sup> recombinant DNA was constructed by **Stanley Cohen** and **Herbert Boyer** in 1972. They cut a piece of DNA from a plasmid carrying antibiotic resistance gene from a bacterium *Salmonella typhimurium* and linked it the plasmid of *E. Coli*. So it's a technique of combining genes from two different sources.

**The process of rDNA technology involves following Steps:**

1. Isolation of DNA and find out the gene of interest
2. Cutting of DNA at specific locations and formation of rDNA
3. Amplification of gene of interest using PCR
4. Insertion rDNA into the host cell/organisms
5. Obtaining the desirable gene product
6. Downstream processing

**Tools of rDNA technology**

1. Enzymes
2. Cloning Vectors
3. Competent host

**1. Enzymes**

- a. lysing enzymes
- b. Cleaving enzymes

- c. Synthesizing enzymes
- d. Joining enzymes
- e. Alkaline phosphatase

a. **Lysing Enzymes:**

To open up the cells to obtain DNA

e.g. **Lysozyme** is used to dissolve the bacterial cell wall

**Cellulase** is used for plant cells

**Chitinase** is used to isolate DNA from fungal cells

b. **Cleaving enzymes:**

Used to break DNA molecules- 3 types

- I. Exonuclease – removes nucleotides from terminal ends (5' or 3') of DNA in one strand.
- II. Endonuclease – can cut any point within the DNA except at terminal ends; can cut single or double strands.
- III. Restriction endonucleases (RE) - can cut DNA duplex at certain specific positions e.g. Hind II can cut DNA molecules at a particular point by recognizing a specific sequence of 6 base pair which is called recognition sequence for Hind II. These enzymes are therefore called molecular scissors or chemical scalpels. Actually these enzymes recognize specific palindromes sequences in DNA and cut them. Palindromes are base pair sequences that are same read from left to right or right to left. e.g. 5' – GAATTC – 3'      **a palindrome sequence**

3' – CTTAAG – 5'

**RE – 3 types**

**Type I** - consist of 3 different subunits

**Type II** – simple and used in rDNA technology

**Type III** - intermediate between the two

**Nomenclature**

RE are named after the bacterium from which they are isolated

- I. 1<sup>st</sup> letter of genus is used to denote the 1<sup>st</sup> letter of the enzyme
- II. 1<sup>st</sup> two letter of species make the 2<sup>nd</sup> and 3<sup>rd</sup> letter of the enzyme
- III. 1<sup>st</sup> 3 letters are written in italics form
- IV. 4<sup>th</sup> letter of the enzyme is the 1<sup>st</sup> letter of bacterium strain written in capital
- V. Roman letters indicates the order in which the enzyme was isolated

e.g. EcoRI

c. **Synthetizing enzymes**

Required for DNA synthesis i.e. for DNA replication. They are of two types -

- I. DNA polymerases which require DNA template
- II. Reverse transcriptase require RNA template and newly synthesized DNA is cDNA

d. **Joining enzymes**

DNA ligases –can join two DNA fragments or help in sealing gaps in DNA fragment by forming phospho-di ester bond between them. So they act as **molecular glue**, T4 DNA ligase is most commonly used in rDNA technology.

e. **Alkaline phosphatases**

- These enzymes removes phosphate group from the 5' end of double stranded or single stranded DNA or RNA

2. **Cloning Vectors**

DNA molecules which replicate inside the host cell. It includes-

- a. Plasmids
- b. Cosmids
- c. Bacteriophages
- d. Artificial chromosome vectors e.g. BAC(300-350 kb), YAC (1mb), MAC
- e. Phagemids - constructed by bacteriophage and plasmid
- f. Transposons – e.g. **P- elements in Drosophila**
- g. Viral vectors
- h. Shuttle vectors – used in both *E. coli* and eukaryotic cells

**Characteristics of a cloning vectors:**

It must have-

**Origin of replication (ori)**

DNA sequence which serves as starting point for replication. Can use host cells own's DNA replicative enzymes or some of the plasmids carry genes that code for special enzymes that are specific for plasmid replication. At least one so that plasmids can multiply within the cells independently to the main bacterial chromosome

**Marker gene/selectable marker**

Vector also require a selectable marker to identify and eliminate non-transformants. Generally antibiotic resistance gene is used as selectable marker e.g. tetracycline, ampicillin, kanamycin, chloramphenicol for *E. coli*

**Recognition sites**

A vector must have a one unique RE recognition site. Commonly used vectors contain unique recognition site for many RE. This is referred as polylinker or multiple cloning sites. This provides flexibility in the choice of RE that can be used for cloning.

### Small Size

Cloning vector should be small in size. Less than **10 kb** is desirable for a cloning vector

Plasmid range – 1kb - > 250kb

### Copy number

The number of molecules of an individual vector plasmid normally present in a bacterial cell

Copy no. should be high, range – **1 to 50** copies/bacterium

### Identification of recombinant cells

Not all species of bacteria are equally efficient to uptake DNA. So the bacteria for the purpose of transformation have to undergo some form of physical and/or chemical treatment.

Cells that have undergone this treatment are called competent. This is done by treating them with a specific concentration (50mM) of divalent cation usually  $\text{Ca}^{++}$ . After this treatment the cells are incubated with rDNA on ice, followed by placing them briefly on 42 °C (heat shock) and then putting back them on the ice. This enables the bacteria to take up the r DNA

### Questions

1. Which type of enzyme is commonly used in the process of recombinant DNA technology?
  - A. DNA ligase
  - B. DNA polymerase
  - C. DNA helicase
  - D. Restriction enzyme
2. Which step of recombinant DNA technology involves the introduction of the recombinant DNA into a host organism?
  - A. Cutting DNA
  - B. Isolation of DNA
  - C. Replication and expression
  - D. Selection of vector
3. What is the function of selectable markers in the identification of recombinant cells?
  - A. Isolate DNA from the host organism
  - B. Bind to the recombinant DNA sequence
  - C. Allow growth and survival of recombinant cells



- D. Cut DNA at specific sites
4. The recognition sites of restriction enzymes are usually:
- A. Palindromic
  - B. Complementary to the DNA fragment
  - C. Methylation-resistant
  - D. Non-specific
5. Which type of restriction enzyme is the most commonly used in molecular biology?
- A. Type I
  - B. Type II
  - C. Type III
  - D. Type IV
6. What is the role of DNA ligase in recombinant DNA technology?
- A. Recognize specific DNA sequences
  - B. Add methyl groups to DNA
  - C. Join DNA fragments together
  - D. Unwind the DNA double helix
7. Which enzyme is responsible for synthesizing new DNA strands during DNA replication?
- A. DNA helicase
  - B. DNA polymerase
  - C. DNA ligase
  - D. DNA methyl transferase
8. DNA helicases are involved in:
- A. Joining DNA fragments
  - B. Adding methyl groups to DNA
  - C. Unwinding and separating the DNA double helix
  - D. Sealing nicks and gaps in DNA
9. What is the primary function of topoisomerases?
- A. Synthesizing new DNA strands
  - B. Unwinding DNA helix
  - C. Introducing transient breaks in DNA strands
  - D. Joining DNA fragments together
10. What is the evolutionary significance of restriction enzymes in bacteria?
- A. They protect bacteria from antibiotic resistance

- B. They provide immunity against bacterial infections
- C. They regulate the expression of specific genes
- D. They defend against invading viral DNA
11. DNA methylation is an essential process for:
- A. Repairing DNA damage
- B. Initiating DNA replication
- C. Regulating gene expression
- D. Synthesizing DNA strands
12. Which DNA modifying enzyme creates phosphodiester bonds between adjacent nucleotides?
- A. DNA helicase
- B. DNA polymerase
- C. DNA ligase
- D. DNA methyl transferase
13. The role of DNA polymerase in DNA replication is to:
- A. Unwind the DNA double helix
- B. Add methyl groups to DNA
- C. Synthesize new DNA strands
- D. Introduce transient breaks in DNA
14. What type of cutting pattern does a restriction enzyme with blunt ends produce?
- A. Overhangs
- B. Sticky ends
- C. Single-stranded breaks
- D. No overhangs
15. Recombinant DNA technology involves the insertion, deletion, or modification of:
- A. Proteins
- B. Lipids
- C. Carbohydrates
- D. Genes
16. What is the primary purpose of using selectable markers in the identification of recombinant cells?
- A. To identify the host organism
- B. To cut the DNA into fragments
- C. To introduce foreign DNA into the cells

- D. To distinguish transformed cells from non-transformed cells
17. Which type of restriction enzyme is widely used in molecular biology applications?
- A. Type I
  - B. Type II
  - C. Type III
  - D. Type IV
18. What is the primary purpose of cloning vectors in molecular biology and genetic engineering?
- A. Introduction of therapeutic genes into patient cells
  - B. Propagation and manipulation of DNA fragments in host organisms
  - C. Inhibition of gene expression using RNAi
  - D. Repairing genetic mutations using CRISPR-Cas9
19. Which of the following characteristics is not typically found in cloning vectors?
- A. Origin of replication
  - B. Reporter genes
  - C. Multiple Cloning Site (MCS)
  - D. Transcription factors
20. Which type of gene transfer technique is commonly used to introduce foreign DNA into animal cells?
- A. Transformation
  - B. Transfection
  - C. Gene Gun
  - D. Electroporation
21. What is the primary principle behind gene therapy?
- A. Inhibiting gene expression using RNAi
  - B. Introducing functional genes into patient cells to treat diseases
  - C. Editing genes to correct mutations
  - D. Creating gene libraries for genetic research
22. Which of the following viral vectors is commonly used in gene therapy due to its ability to infect a wide range of cells?
- A. Bacteriophages
  - B. Adeno-associated viruses(AAVs)
  - C. Retroviruses
  - D. Yeast artificial chromosomes(YACs)



23. Which gene transfer technique involves the direct injection of DNA or RNA molecules into the nucleus or cytoplasm of individual cells?
- Transformation
  - Transfection
  - Gene gun
  - Microinjection
24. What is the primary role of restriction enzymes in genetic engineering?
- Cutting DNA at specific recognition sites
  - Adding methyl groups to DNA sequences
  - Repairing DNA mutations
  - Inhibiting gene expression using RNAi
25. Which type of cloning vector is derived from bacterial chromosomes and can carry very large DNA inserts?
- Plasmids
  - Bacterial artificial chromosomes(BACs)
  - Yeast artificial chromosomes(YACs)
  - Bacteriophages
26. What is the primary purpose of selectable markers in cloning vectors?
- Facilitating gene expression studies
  - Identifying and selecting cells containing the vector
  - Inhibiting the replication of cloning vectors
  - Introducing specific mutations in DNA sequences
27. Which type of gene therapy approach uses RNAi or antisense oligonucleotides to inhibit the expression of disease-causing genes?
- Gene silencing therapy
  - Gene editing therapy
  - Gene addition therapy
  - Gene transformation therapy
28. Which viral vector is commonly used in gene therapy and can infect both dividing and non-dividing cells?
- Retroviruses

- B. Lenti viruses  
C. Adenoviruses  
D. Adeno-associated viruses (AAVs)
29. Which gene transfer technique is used to introduce foreign DNA into bacterial, yeast, or plant cells?
- Transformation
  - Transfection
  - Gene gun
  - Microinjection
30. Which type of gene therapy approach involves using gene editing techniques like CRISPR-Cas9 to precisely modify the patient's DNA?
- Gene silencing therapy
  - Gene editing therapy
  - Gene addition therapy
  - Gene transformation therapy
31. Which type of cloning vector is derived from yeast chromosomes and can carry even larger DNA inserts compared to BACs?
- Plasmids
  - Bacterial artificial chromosomes (BACs)
  - Yeast artificial chromosomes (YACs)
  - Bacteriophages
32. Which gene transfer technique involves isolating cells from an animal, genetically modifying them in the laboratory, and then reintroducing them back into the animal?
- In vivo gene therapy
  - Ex vivo gene therapy
  - Gene silencing therapy
  - Gene editing therapy
33. Which type of gene therapy approach involves introducing genes that can enhance the immune response against cancer cells?
- Gene silencing therapy
  - Gene editing therapy
  - Gene addition therapy
  - Cancer immunotherapy using viral vectors

34. Vectors are
- A. Molecule that degrade nucleic acids
  - B. Molecules that help in replication
  - C. Molecules that are able to bond and carry foreign DNA
  - D. Molecules that protect host cells from invasion by foreign DNA
35. The normal role of restriction endonucleases in bacterial cells is
- A. To degrade the bacterial chromosome
  - B. To degrade the invading phage DNA
  - C. To reproduce RNA primers for replication
  - D. None of the above
36. Genetic engineering means
- A. Manipulation of genes
  - B. Manipulation of genetics
  - C. Manipulation of cell content
  - D. All of the above
37. A short molecule containing 2-20 nucleotide is
- A. Mononucleotide
  - B. Oligonucleotide
  - C. Plasmid
  - D. Cosmid
38. Biotechnology is widely used in
- A. Medicine
  - B. Agriculture
  - C. Industry
  - D. All of the Above
39. The most common plasmid vector used in genetic engineering is
- A. pBR 325
  - B. pBR 322
  - C. pBR 328
  - D. pBR 320
40. EcoRI is a
- A. Ligase
  - B. Gyrase
  - C. Polymerase

- D. Restriction enzyme
41. Which of the following is not a product of rDNA technology
- A. Golden Rice
  - B. Bt.Brinjal
  - C. Bt.Cotton
  - D. Dolly
42. Genetically modified DNA fragments are called
- A. Recombinant DNA
  - B. S-DNA
  - C. Mitochondrial DNA
  - D. None of the Above
43. The vectors commonly used for sequencing human genome are
- A. YAC
  - B. Plasmids
  - C. CMV
  - D. M13 Vectors
44. The mechanism of intake of DNA fragments from the surrounding medium by a cell is called
- A. Transduction
  - B. Conjugation
  - C. Transformation
  - D. All of the Above
45. Who discovered restriction enzymes?
- A. Watson and Crick
  - B. Jacob and Monad
  - C. Boyer and Cohen
  - D. Nathen, Arber and Smith
46. A recombinant DNA molecule is produced by
- A. Joining of two DNA fragments
  - B. Joining of two or more DNA fragments
  - C. Both (A) and (B)
  - D. Joining two or more DNA fragments originating from different species
47. The DNA molecule to which the gene of interest is integrated for cloning is called
- A. Carrier

- B. Vector  
C. Transformer  
D. None of the above
48. Restriction enzymes are named for  
A. The Person who discovered  
B. The bacterium they are derived from  
C. The viral DNA that they attack  
D. None of the above
49. Molecular scissor is  
A. Ligase  
B. Helicase  
C. Restriction endonucleases  
D. DNA Polymerase
50. rDNA stands for  
A. Reverse DNA  
B. Replicate DNA  
C. Recombinant DNA  
D. None of the above

### **ANSWERS**

1-D,2-C,3-C,4-A,5-B,6-C,7-B,8-C,9-C,10-D,11-C,12-C,13-C,14-D,15-D,16-D,17- B,18-B,19-D,20-B,21-B,22-B,23-D,24-A,25-B,26-B,27-A,28-D,29-A,30-B,31-C, 32-B,33-D,34-C,35-D,36-A,37-B,38-D,39-B,40-D,41-D,42-A,43-A,44-C,45-D,46-D,47-A,48-B,49-C,50-C

### **Unit II**

#### **Applications of Genetic Engineering**

- Single Cell Proteins
- Biosensors, Biochips
- Crop and Livestock improvement, development of transgenes
- Development of DNA drugs and vaccines

Genetic engineering, a field within biotechnology, involves the deliberate modification of an organism's genetic material to achieve desired traits or outcomes. It has revolutionized various industries and has a wide range of applications in agriculture, medicine, environmental conservation, and industrial production



### Single Cell Proteins

Single cell proteins (SCP) refer to protein-rich food sources derived from microbial cells, typically grown in large-scale fermentation processes e.g. Yeast, Bacteria, Fungi

**Yeast:** *Saccharomyces cerevisiae* is commonly used yeast for SCP production due to its rapid growth, high protein content, and ability to utilize a wide range of substrates.

**Bacteria:** Species like *Escherichia coli* and *Bacillus subtilis* have been utilized for SCP production. They have the advantage of rapid growth and genetic manipulability.

**Fungi:** Filamentous fungi *Aspergillus spp.* And *Fusarium spp*

### Biosensors

Biosensors are analytical devices that combine a biological component (such as enzymes, antibodies, or living cells) with a physicochemical transducer to detect and quantify specific substances or analytes in various samples.

#### Types of Biosensors:

1. Enzymatic Biosensors
2. Immunosensors
3. DNA Biosensors
4. Whole-Cell Biosensors
5. Optical Biosensors

#### Advantages of Biosensors

1. High Sensitivity and Selectivity
2. Rapid Analysis
3. Portability and Miniaturization
4. Cost-Effectiveness

### Biochips

Biochips, also known as microarrays or bioarrays, are miniaturized devices that integrate biological components (such as DNA, proteins, cells, or enzymes) onto a solid substrate, typically a glass slide or a silicon chip. These biochips enable high-throughput analysis, allowing for the simultaneous detection and analysis of thousands to millions of biomolecules in a single experiment.

#### Types of Biochips

DNA Microarrays – Used for gene expression profiling, genotyping, and mutation analysis.

Protein Microarrays – Detection of protein-protein interactions, and antibody screening

Cell-based Microarrays – Used for drug screening, toxicity testing etc.

Tissue Microarrays – Used for gene expression and protein localization across tissues

### **Applications of Biochips**

- Genomics
- Proteomics
- Drug Discovery and Development
- Personalized Medicine
- Biomarker Discovery
- Agriculture and Food Safety

### **Crop improvement**

It involves various techniques and approaches aimed at maximizing agricultural productivity and sustainability.

#### **Modern Breeding Techniques:**

**Genetic Engineering:** Genetic engineering allows the transfer of specific traits across species boundaries, enabling the development of genetically modified (GM) crops with improved traits, such as insect resistance, herbicide tolerance, or enhanced nutritional value.

**Genome Editing:** Genome editing techniques, such as CRISPR-Cas9, enable precise modification of specific genes within an organism's genome, without the need for introducing foreign genes.

#### **Trait Improvement:**

- Enhancing Yield Improvement
- Developing Disease and Pest Resistance varieties
- Develop varieties with Stress Tolerance
- Improve Nutritional Quality

### **Livestock improvement**

These advancements aim to improve various aspects of livestock production, including productivity, disease resistance, animal welfare, and the quality of animal products.

#### **Transgenic Animals:**

Examples of transgenic traits include increased growth rates, improved feed efficiency, enhanced disease resistance, and production of valuable proteins, such as pharmaceuticals or therapeutic antibodies, in the animal's milk or blood.

## DNA Drugs

DNA drugs, also known as gene-based drugs or gene therapies, represent a revolutionary approach in medicine that utilizes DNA molecules as therapeutic agents. These drugs aim to treat diseases by introducing genetic material into a patient's cells, allowing for the correction of genetic defects, regulation of gene expression, or the delivery of therapeutic proteins.

### Types of DNA Drugs:

**Gene Replacement Therapy-** introduce functional copies of a defective gene into cells

**Gene Silencing Therapy -** By introducing small interfering RNA (siRNA)

**Gene Editing Therapy -** CRISPR-Cas9 technology allows for precise editing of the genome

**DNA Vaccines –** DNA Vaccines introduce the genetic material encoding antigenic proteins into the body

DNA vaccines represent a promising avenue for vaccination and immunotherapy. Their unique mechanism of action, ease of design, and potential for broad applications makes them a valuable tool in combating infectious diseases and cancer

### Questions:

1. Which of the following is an application of genetic engineering in agriculture?
  - A. Drug production
  - B. Human gene therapy
  - C. Disease modeling
  - D. Crop improvement
2. Which industry can benefit from genetically engineered microorganisms to produce enzymes with specific properties?
  - A. Textile manufacturing
  - B. Construction
  - C. Oil and gas exploration
  - D. Transportation
3. DNA analysis plays a crucial role in:
  - A. Environmental conservation
  - B. Agricultural development
  - C. Forensic investigations
  - D. Industrial applications
4. Single Cell Proteins (SCP) are:
  - A. Proteins derived from animal sources
  - B. Proteins derived from plant sources

- C. Protein-rich food sources derived from microbial cells  
D. Proteins derived from synthetic sources
5. Which microorganism is commonly used for SCP production due to its rapid growth and high protein content?
- A. *Saccharomyces cerevisiae* (yeast)  
B. *Escherichia coli* (bacteria)  
C. *Aspergillus* spp.(fungi)  
D. All of the above
6. What are the advantages of SCP production?
- A. High Protein Content  
B. Efficient Resource Utilization  
C. Rapid Growth  
D. All of the above
7. Biosensors are analytical devices that:
- A. Combine biological components with physicochemical transducers  
B. Utilize electricity to detect specific substances  
C. Convert light into sound  
D. Are used for space exploration
8. What component of a biosensor provides specificity and selectivity to detect a target analyte?
- a. Signal Processing System  
b. Transducer  
c. Biological Recognition Element  
d. Amplifier
9. Enzymatic biosensors measure the activity of:
- A. Proteins  
B. Enzymes  
C. Nucleic acids  
D. Carbohydrates
10. Which of the following is NOT an application of biosensors?
- A. Medical Diagnostics  
B. Environmental Monitoring  
C. Agricultural Pest Control  
D. Food Safety Testing

11. Which of the following is NOT an application of biosensors?
  - A. Medical Diagnostics
  - B. Environmental Monitoring
  - C. Agricultural Pest Control
  - D. Food Safety Testing
12. What is one advantage of biosensors?
  - A. They require a large sample volume for analysis.
  - B. They are not sensitive enough to detect low-abundance molecules.
  - C. They offer high sensitivity and specificity.
  - D. They are not suitable for real-time analysis.
13. What type of biochip is used to detect complementary DNA or RNA sequences?
  - A. Protein Microarrays
  - B. Tissue Microarrays
  - C. Cell-based Microarrays
  - D. DNA Microarrays
14. What bio chip approach utilizes high-through put genotyping and statistical modeling to predict plant performance based on DNA profiles?
  - A. Genomic Selection
  - B. Marker-Assisted Selection
  - C. Tissue Culture and Micropropagation
  - D. Traditional Breeding
15. Crop improvement aims to enhance:
  - A. Environmental pollution
  - B. Nutritional quality
  - C. Global warming
  - D. Genetic diversity
16. Which breeding technique involves crossing genetically diverse parental lines to create hybrid offspring with desirable traits?
  - A. Selection
  - B. Hybridization
  - C. Mutation breeding
  - D. Somatic embryogenesis
17. Genomics election combines high-through put genotyping with:
  - A. Phenotypic evaluation



- B. In-vitro fertilization  
C. Bioremediation  
D. Statistical modeling
18. Which application of genetic engineering involves the introduction of functional genes into affected cells to correct genetic disorders?
- A. Disease Resistance  
B. Environmental Conservation  
C. Human Gene Therapy  
D. Bioplastics
19. Crop improvement focuses on enhancing yield potential through increased:
- A. Genetic diversity  
B. Photosynthetic efficiency  
C. Heavy metal content  
D. Soil erosion
20. What is the potential benefit of integrating biochips with microfluidics?
- A. Reducing the need for quality control  
B. Enabling the simultaneous detection of multiple analytes  
C. Increasing the cost of analysis  
D. None of the above
21. What is the primary goal of livestock improvement through genetic engineering?
- A. Enhancing animal welfare  
B. Increasing crop yield  
C. Improving environmental conservation  
D. Developing DNA drugs
22. Which technique allows for precise modification of specific genes within an organism's genome?
- A. Gene silencing therapy  
B. Hybridization  
C. Gene editing therapy  
D. Gene replacement therapy
23. DNA vaccines stimulate an immune response by introducing:
- A. Inactivated pathogens  
B. Weak pathogens  
C. Antigenic proteins

- D. Antibodies
24. Which application of DNA vaccines involves stimulating the immune system against tumor-specific antigens?
- A. Cancer immunotherapy
  - B. Infectious diseases
  - C. Veterinary vaccines
  - D. Emerging infectious diseases
25. What is the role of transgenics in crop improvement?
- A. Enhancing disease resistance
  - B. Producing therapeutic proteins
  - C. Correcting genetic defects
  - D. Reducing feed requirements
26. Which technique is used to transfer foreign genes in to livestock embryos for genetic modification?
- A. PCR
  - B. Hybridization
  - C. Electroporation
  - D. Micro projectile bombardment
27. DNA drugs are used for gene replacement therapy to treat:
- A. Infectious diseases
  - B. Neurological disorders
  - C. Genetic disorders
  - D. Cardiovascular diseases
28. What is the primary advantage of DNA vaccines compared to traditional vaccines?
- A. Long-lasting immunity
  - B. Higher efficacy
  - C. Safer administration
  - D. Lower cost
29. Which type of biosensor uses antibodies to specifically bind to target analytes?
- A. Enzymatic biosensor
  - B. DNA biosensor
  - C. Immunosensor
  - D. Whole-cell biosensor
30. Genetic engineering can be applied to enhance crop quality by:

- A. Increasing crop yield  
B. Improving the nutritional content  
C. Enhancing disease resistance  
D. Reducing environmental impact
31. DNA drugs can be used to treat infectious diseases by introducing genes that:  
A. Enhance disease resistance  
B. Inhibit tumor growth  
C. Inhibit viral replication  
D. Improve immune response
32. What is the primary application of single-cell proteins(SCP)in agriculture?  
A. Bio fertilizers and biodegradation  
B. Improved crop quality  
C. Disease resistance  
D. Enzyme production
33. What is the primary challenge in developing transgenics in livestock?  
A. Gene isolation and cloning  
B. Public acceptance  
C. Delivery efficiency  
D. Molecular analysis and characterization
34. DNA vaccines have been studied for various infectious diseases, including all of the following EXCEPT:  
A. HIV  
B. Influenza  
C. COVID-19  
D. Tuberculosis
35. Which type of biochip contains small sections of multiple tissue samples for gene expression analysis?  
A. Protein microarrays  
B. Cell-based microarrays  
C. DNA microarrays  
D. Tissue microarrays
36. Gene editing therapy using CRISPR-Cas9 allows for:  
A. Targeted gene replacement  
B. Gene silencing

- C. Specific gene mutations  
D. All of the above
37. DNA drugs that inhibit the expression of disease-causing genes use specific DNA molecules called:
- A. Enzymes  
B. Antibodies  
C. Small interfering RNA (siRNA)  
D. Messenger RNA (mRNA)
38. What is the primary advantage of using whole-cell biosensors?
- A. High sensitivity  
B. Portability  
C. Rapid analysis  
D. Detection of multiple analytes
39. Gene replacement therapy using DNA drugs is particularly beneficial for diseases caused by:
- A. Multiple genes  
B. Environmental factors  
C. Single genes  
D. Viral infections
40. Which technique is used to introduce foreign genes into plants for crop improvement?
- A. PCR  
B. Agrobacterium-mediated transformation  
C. Micro projectile bombardment  
D. Electroporation
41. DNA drugs are designed to treat genetic disorders by:
- A. Introducing functional copies of a defective gene  
B. Introducing small interfering RNA (siRNA)  
C. Inhibiting viral replication  
D. Enhancing immune responses
42. What is the primary application of biosensors in forensic science?
- A. DNA analysis  
B. Environmental monitoring  
C. Medical diagnostics  
D. Drug discovery

43. DNA vaccines offer potential solutions for rapidly emerging infectious diseases due to their:
- A. Low cost
  - B. Broad applicability
  - C. Ease of design
  - D. High specificity
44. Gene editing therapy using CRISPR-Cas9 allows for precise modification of specific genes within the:
- A. Cytoplasm
  - B. Nucleus
  - C. Mitochondria
  - D. Endoplasmic reticulum
45. Which one of the following is used in industrial preparation of alcohol ?
- A. Aspergillus
  - B. Saccharomyces
  - C. Lactobacillus
  - D. Pseudomonas
46. Biotechnology is used in
- A. Medicine
  - B. Agriculture
  - C. Industry
  - D. All of the Above
47. The full form of GMO is
- A. Genetically mutant organisms
  - B. Genetically modified organisms
  - C. Genetically modern organisms
  - D. Genetically transferred organisms
48. Gene therapy is a technique that helps in
- A. Clonal Propagation
  - B. Producing monoclonal antibodies
  - C. Saving endangered species
  - D. Curing genetic disorders
49. The insulin prepared through genetic engineering is called
- A. Microbial Insulin



- B. Humulin  
 C. Human Insulin  
 D. None of the Above
50. Probiotics are
- A. Antibiotics  
 B. Microbes  
 C. Food Allergen  
 D. Live microbial food supplement

### ANSWERS

1-D,2-A,3-C,4-C,5-D,6-D,7-A,8-C,9-B,10-C,11-C,12-C,13-D,14-A,15-B,16-B,17- D,18-  
 C,19-B,20-B,21-A,22-C,23-C,24-A,25-B,26-C,27-C,28-C,29-C,30-B,31-C, 32-A,33-C,34-  
 D,35-C,36-D,37-C,38-D,39-C,40-B,41-A,42-A,43-B,44-B,45-B, 46- D,47- B, 48-D,49- B,  
 50-D

### UNIT III

#### DNA Diagnostics

- Genetic analysis of human diseases (Hemophilia, Colour blindness), Detection of known and unknown mutations
- Concept of pharmacogenomics and pharmacogenetics

DNA diagnostics, also known as genetic testing or molecular diagnostics, is a branch of medical diagnostics that utilizes DNA analysis to detect and identify genetic variations associated with diseases or conditions.

#### **Types of DNA Diagnostics:**

1. **Diagnostic Testing:** Used to confirm or rule out the presence of a suspected genetic condition in individuals showing symptoms.
2. **Carrier Testing:** Determines whether an individual carries a genetic mutation associated with a specific inherited disorder.
3. **Predictive Testing:** Identifies genetic mutations associated with diseases that may

develop in the future, such as certain types of cancer or neurodegenerative disorders.

4. **Pharmacogenetic Testing:** Examines genetic variations that affect an individual's response to medications..

#### **Applications of DNA Diagnostics:**

- a) Diagnose genetic disorders
- b) Testing Cancer
- c) Detecting Infectious disease
- d) Preparation of personalized medicine
- e) Prenatal testing

#### **Genetic analysis of Human Disease**

Genetic analysis of human disease refers to the study and investigation of genetic factors underlying the development and progression of various diseases in humans. It involves the identification and characterization of genetic variations, mutations, and alterations that contribute to the risk, onset, severity, or response to treatment of different diseases

#### **Applications of Genetic Analysis in Human Disease:**

- a) To identify Mendelian disorders
- b) To treat complex diseases
- c) Play important role in cancer therapy
- d) To design precise medicine on individual level

#### **Detection of Known and Unknown Mutation**

Targeted analysis and Genotypic array, Next generation sequencing can be used to detect known mutations. Whereas Whole genome sequencing is used to find out unknown mutations

#### **Concept of Pharmacogenomics**

Pharmacogenomics is a field of study that explores the relationship between an individual's genetic makeup and their response to medications. It focuses on how genetic variations can influence an individual's drug metabolism, efficacy, and adverse reactions.

Pharmacogenomics is a rapidly evolving field that aims to optimize drug therapy by considering an individual's genetic makeup

#### **Concept of Pharmacogenetics**

Pharmacogenetics examines how genetic factors influence drug metabolism, efficacy, and safety, and how this knowledge can be applied to personalize drug treatment.

**Questions**

1. DNA diagnostics is a branch of medical diagnostics that utilizes \_\_\_\_\_ analysis to detect and identify genetic variations associated with diseases or conditions.
  - a. RNA
  - b. Protein
  - c. DNA
  - d. Lipid
2. DNA diagnostics helps in identifying disease-causing factors and providing \_\_\_\_\_.
  - a. Symptomatic relief
  - b. Home remedies
  - c. Accurate diagnoses
  - d. Surgical treatments
3. What type of DNA testing determines whether an individual carries a genetic mutation associated with a specific inherited disorder?
  - a. Diagnostic testing
  - b. Predictive testing
  - c. Carrier testing
  - d. Pharmacogenetic testing
4. Which DNA analysis technique allows the identification of genetic variations or mutations associated with diseases?
  - a. Polymerase Chain Reaction (PCR)
  - b. DNA Sequencing
  - c. Microarray Technology
  - d. Next-Generation Sequencing (NGS)
5. DNA diagnostics is widely used in diagnosing and identifying genetic disorders such as:
  - a. Infectious diseases
  - b. Cardiovascular diseases
  - c. Cancer
  - d. Mental disorders
6. Genetic analysis of human disease helps in understanding the underlying \_\_\_\_\_ of diseases.
  - a. Social causes

- b. Environmental causes
  - c. Genetic factors
  - d. Psychological factors
7. Which genetic analysis approach involves scanning the entire genome of large populations to identify common genetic variations associated with disease susceptibility?
- a. Functional studies
  - b. Family-based studies
  - c. Genome-Wide Association Studies(GWAS)
  - d. Exome sequencing
8. DNA diagnostics aids in predicting individual responses to medications, which is known as:
- a. Diagnostic testing
  - b. Prognostication
  - c. Carrier testing
  - d. Pharmacogenomics
9. DNA sequencing is a technique used in genetic analysis that determines the precise order of \_\_\_\_\_ in a DNA molecule.
- a. Carbohydrates
  - b. Proteins
  - c. Nucleotides
  - d. Amino acids
10. What is the main focus of Exome Sequencing in genetic analysis?
- a. Identifying chromosomal abnormalities
  - b. Studying non-coding regions of the genome
  - c. Sequencing the protein-coding regions of the genome
  - d. Analyzing functional gene pathways
11. Genetic analysis helps assess an individual's genetic predisposition to certain diseases, allowing for targeted \_\_\_\_\_.
- a. Preventive measures
  - b. Treatment plans
  - c. Lifestyle changes

- d. Surgical interventions
12. Which genetic analysis technique allows the simultaneous analysis of thousands to millions of genetic markers associated with diseases?
- Microarray Analysis
  - Polymerase Chain Reaction (PCR)
  - DNA Sequencing
  - Next-Generation Sequencing (NGS)
13. Which application of genetic analysis involves identifying genetic mutations associated with different types of cancer for diagnosis and treatment decisions?
- Mendelian disorders
  - Infectious disease diagnostics
  - Cancer genetics
  - Carrier testing
14. The interpretation of genetic data can be complex and requires expertise in:
- Epidemiology
  - Psychology
  - Bioinformatics
  - Anatomy
15. Genetic analysis of human disease provides crucial insights into the genetic factors contributing to disease development, progression, and treatment response, shaping the future of \_\_\_\_\_.
- Genetic engineering
  - Precision medicine
  - Traditional medicine
  - Public health
16. DNA diagnostics leverage the knowledge of specific \_\_\_\_\_ to identify disease-causing factors.
- RNA sequences
  - Protein structures
  - Genes and genetic mutations
  - Metabolic pathways
17. Which type of DNA testing helps in identifying genetic markers or single nucleotide polymorphisms (SNPs) associated with disease susceptibility?

- a. Predictive testing
  - b. Pharmacogenetic testing
  - c. Genome-Wide Association Studies(GWAS)
  - d. Diagnostic testing
18. What does NGS stand for in the context of genetic analysis techniques?
- a. Non-Genetic Sequencing
  - b. New Genetic Screening
  - c. Next-Generation Sequencing
  - d. Nano-Genomic Studies
19. Genetic analysis raises ethical concerns related to privacy, confidentiality, and potential discrimination based on \_\_\_\_\_.
- a. Socioeconomic status
  - b. Geographic location
  - c. Genetic information
  - d. Medical history
20. Which application of genetic analysis involves identifying causative mutations in monogenic disorders?
- a. Mendelian disorders
  - b. Cancer genetics
  - c. Complex diseases
  - d. Pharmacogenomics
21. DNA diagnostics helps in assessing an individual's genetic predisposition to certain diseases, allowing for targeted \_\_\_\_\_ or early intervention.
- a. Cure
  - b. Palliative care
  - c. Preventive measures
  - d. Diagnostic tests
22. Which genetic analysis approach involves analyzing genetic information from affected individuals and their family members to identify disease-causing mutations or inheritance patterns?
- a. Functional studies
  - b. Family-based studies
  - c. Genome-Wide Association Studies(GWAS)



- d. Exome sequencing
23. What does PCR stand for in the context of genetic analysis techniques?
- Polymerase Chain Reaction
  - Protein Coding Region
  - Pre-Cellular Replication
  - Protein Cellular Reaction
24. Which type of DNA testing aids in predicting disease risk for certain conditions that may develop in the future?
- Diagnostic testing
  - Carrier testing
  - Predictive testing
  - Pharmacogenetic testing
25. Genetic analysis plays a crucial role in oncology, identifying genetic mutations associated with different types of \_\_\_\_\_.
- Diabetes
  - Cardiovascular diseases
  - Cancer
  - Infectious diseases
26. Pharmacogenomics focuses on the relationship between an individual's \_\_\_\_\_ and their response to medications.
- Environment
  - Lifestyle
  - Genetic makeup
  - Diet
27. Genetic variations in drug-metabolizing enzymes can affect the \_\_\_\_\_ of drugs in the body.
- Absorption
  - Distribution
  - Metabolism
  - Excretion
28. Which of the following is NOT a type of genetic variation discussed in the content?
- Single Nucleotide Polymorphisms(SNPs)

- b. Copy Number Variations(CNVs)
- c. Gene Transcription
- d. Gene Expression
29. Pharmacogenetics can help in predicting individuals who may be at a higher risk of \_\_\_\_\_ to certain medications.
- a. Developing allergies
- b. Severe drug interactions
- c. Developing cancer
- d. Adverse drug reactions
30. What is the main goal of pharmacogenetics in drug therapy?
- a. To discover new drugs
- b. To minimize drug resistance
- c. To personalize drug treatment based on genetic profile
- d. To eliminate adverse drug reactions
31. How can genotyping arrays detect known mutations?
- a. By hybridizing with specific DNA sequences
- b. By amplifying DNA regions using PCR
- c. By sequencing the entire genome
- d. By identifying de novo mutations
32. Whole-genome sequencing (WGS) involves sequencing the entire \_\_\_\_\_ to identify both known and unknown mutations.
- a. Exome
- b. Proteome
- c. Genome
- d. Transcriptome
33. What is the purpose of de novo mutation detection in pharmacogenetics?
- a. To identify inherited mutations from parents
- b. To identify novel genetic variations
- c. To identify mutations associated with drug resistance
- d. To assess gene expression levels
34. Pharmacogenomics helps in personalizing drug dosing regimens based on an individual's \_\_\_\_\_.
- a. Weight

- b. Age
  - c. Genetic profile
  - d. Gender
35. Which of the following is NOT a clinical application of pharmacogenomics?
- a. Drug selection based on genetic response
  - b. Personalized drug dosing based on genetic profile
  - c. Identifying infectious diseases
  - d. Predicting adverse drug reactions
36. What challenges may arise in variant classification in pharmacogenomics?
- a. Interpreting gene expression levels
  - b. Analyzing large volumes of sequencing data
  - c. Identifying known mutations
  - d. Differentiating pathogenic mutations from benign variants
37. Which of the following is a technical limitation in mutation detection methods?
- a. High sensitivity
  - b. High specificity
  - c. Missed or misinterpreted mutations
  - d. Quick and accurate results
38. Why is education and awareness essential in implementing pharmacogenomics?
- a. To increase the cost of genetic testing
  - b. To promote the use of traditional medicine
  - c. To help patients understand their genetic profile
  - d. To limit the use of pharmacogenomics in medicine
39. What type of sequencing allows for the identification of genetic variations in both coding and non-coding regions?
- a. Whole-genome sequencing (WGS)
  - b. Whole-exome sequencing (WES)
  - c. Targeted gene panels
  - d. Single-cell sequencing
40. How can advancements in sequencing technologies improve mutation detection?
- a. By reducing the need for bioinformatics tools
  - b. By decreasing the cost of genetic testing

- c. By enhancing the accuracy and efficiency of detection
- d. By focusing on known mutations only
41. Which of the following is NOT a type of genetic variation that can influence drug response?
- a. Gene expression
- b. Copy Number Variations(CNVs)
- c. Micro RNA levels
- d. Single Nucleotide Polymorphisms (SNPs)
42. Pharmacogenetics can help optimize drug therapy in specific populations, such as \_\_\_\_\_.
- a. Pregnant women
- b. Teenagers
- c. Athletes
- d. Children and the elderly
43. How can pharmacogenetics contribute to drug research and development?
- a. By increasing the cost of drug development
- b. By reducing the need for clinical trials
- c. By providing insights in to new disease mechanisms and therapeutic targets
- d. By limiting the use of personalized medicine
44. What type of genetic variations can affect drug absorption, distribution, and elimination?
- a. Drug-metabolizing enzymes
- b. Gene expression levels
- c. Copy Number Variations(CNVs)
- d. Drug transporters
45. What is the main focus of pharmacogenetics in relation to drug targets?
- a. Identifying new drug targets
- b. Assessing the efficiency of drug targets
- c. Understanding how genetic variations affect drug targets
- d. Developing drugs that target specific genetic variations
46. Which of the following is NOT a benefit of pharmacogenetics in drug therapy?
- a. Increased risk of adverse drug reactions
- b. Optimized drug dosing

- c. Improved treatment outcomes
- d. Personalized medicine approaches
47. How can bioinformatics tools be used in pharmacogenetics?
- a. To conduct clinical trials
- b. To analyze sequencing data and classify genetic variations
- c. To determine drug efficacy in patients
- d. To replace genetic testing methods
48. What type of genetic variations are the most common and involve a single nucleotide change in the DNA sequence?
- a. Copy Number Variations (CNVs)
- b. Gene fusions
- c. Single Nucleotide Polymorphisms (SNPs)
- d. Tandem repeats
49. Which of the following is NOT a focus of pharmacogenomics in drug therapy?
- a. Drug metabolism
- b. Drug targets
- c. Drug distribution
- d. Drug response
50. What is the potential impact of identifying unknown mutations in pharmacogenetics?
- a. Discovering new drug targets
- b. Reducing the cost of genetic testing
- c. Decreasing the need for drug development
- d. Enhancing drug safety and efficacy

### ANSWERS

1-c, 2-c, 3-c, 4-b, 5-c, 6-c, 7-c, 8-d, 9-c, 10-c, 11-a, 12-a, 13-c, 14-c, 15-b, 16-c, 17-c, 18-c, 19-c, 20-a, 21-c, 22-b, 23-a, 24-c, 25-c, 26-c, 27-c, 28-c, 29-d, 30-c, 31-a, 32-c, 33-b, 34-c, 35-d, 36-d, 37-c, 38-c, 39-a, 40-c, 41-c, 42-d, 43-c, 44-d, 45-c, 46-a, 47-b, 48-c, 49-c, 50-d

## Unit IV

### Immune System and its components

- Historical perspective of immunology: Innate and acquired immunity, Clonal Selection, complement system
- Structure and functions of different classes of immunoglobulins
- Humoral Immunity and cell mediated immunity
- HLA Complex: organization class I and II HLA molecules

### Historical perspective of immunology

Immunology, the study of the immune system and its functions, has a rich and fascinating historical background. The concept of immunity can be traced back to ancient civilizations, where individuals observed that people who survived certain diseases, such as the plague, were protected from future infections. In the 18th century, Edward Jenner, an English physician, developed the small pox vaccine. He observed that milkmaids who contracted cowpox, a milder disease, seemed to be protected from smallpox. Jenner conducted experiments and successfully demonstrated that inoculating individuals with material from cowpox lesions provided protection against smallpox.

### Innate immunity

Innate immunity, also known as natural or non-specific immunity, is the first line of defense in the immune system. It provides immediate protection against a wide range of pathogens without prior exposure or specific recognition. Innate immunity is a fundamental defense mechanism that acts rapidly to contain and eliminate invading microorganisms.

### Physical Barriers

- Skin
- Tears
- Saliva
- Mucous membrane of respiratory and urinogenital tract
- HCl

### Cellular Components

- Neutrophils
- Phagocytes
- Natural killer Cells



- Dendritic cells

### **Chemical components**

- Interleukins
- Interferons
- Antimicrobial peptides
- Complement System

### **Adaptive immunity**

Adaptive immunity, also known as acquired or specific immunity, is a sophisticated defense mechanism of the immune system that develops after exposure to specific pathogens or antigens. Unlike innate immunity, which provides immediate but nonspecific protection, adaptive immunity is highly specific, adaptable, and capable of recognizing and targeting specific pathogens or antigens. It is characterized by immunologic memory.

### **Cellular components**

- T-Cells
- B-Cells
- Phagocytes

### **Clonal selection**

Clonal selection is a fundamental concept in immunology that describes how the adaptive immune system selectively expands and activates specific lymphocytes in response to encountering an antigen.

### **Clonal Expansion:**

Upon activation, the selected lymphocyte undergoes clonal expansion. This process involves the rapid proliferation and division of the activated lymphocyte, leading to the formation of a large population of identical cells, known as a clone.

### **The complement system and its components**

The complement system is a complex and crucial component of the immune system that consists of a cascade of proteins that interact with each other and with other immune cells to eliminate pathogens. Complement proteins are produced in the liver and circulate in the bloodstream in their inactive forms. They are designated with the prefix “C” followed by a number. The complement system consists of three main pathways: the classical pathway, the lectin pathway, and the alternative pathway. The classical pathway is primarily activated by

antigen-antibody complexes, the lectin pathway is activated by the recognition of specific carbohydrate structures on pathogens, and the alternative pathway can be spontaneously activated on microbial surfaces.

### **Antibody Classes and Isotypes:**

There are five main antibody classes or isotypes:

- **IgM:** The first antibody produced during an immune response. It is effective at activating the complement system.
- **IgG:** The most abundant antibody in the bloodstream. It provides long-term protection and crosses the placenta to provide passive immunity to the fetus.
- **IgA:** Predominantly found in mucosal secretions, such as saliva, tears, and breast milk, providing localized defense against pathogens.
- **IgE:** Involved in allergic reactions and defense against parasitic infections.
- **IgD:** Functions primarily as a cell surface receptor on B cells, involved in B cell activation.

### **Humoral Immunity**

Humoral immunity plays a critical role in defending against extracellular pathogens, particularly bacteria and viruses circulating in the bloodstream or residing in mucosal surfaces. Antibodies produced by B cells provide diverse mechanisms of protection, including neutralization, opsonization, complement activation, and immune complex clearance. The ability of humoral immunity to generate memory B cells ensures a rapid and effective response upon subsequent encounters with the same pathogen.

### **Cell mediated Immunity**

Cell-mediated immunity is a crucial arm of the adaptive immune response that involves the activation and coordination of various types of immune cells to eliminate intracellular pathogens and target infected or abnormal cells. Unlike humoral immunity, which is mediated by antibodies, cell-mediated immunity relies on the actions of specialized cells, primarily T lymphocytes. Cell-mediated immunity is primarily mediated by T cells, specifically CD4<sup>+</sup> helper T cells (Th cells) and CD8<sup>+</sup> cytotoxic T cells (Tc cells).

### **Major Histocompatibility Complex (MHC):**

**MHC Class I:** MHC class I molecules are present on the surface of almost all nucleated cells. They present intracellular antigens to cytotoxic T cells, allowing them to recognize and eliminate infected or abnormal cells.

**MHC Class II:** MHC class II molecules are primarily expressed on antigen-presenting cells.

They present extracellular antigens to helper T cells, enabling them to provide immune activation signals.

### Questions

1. What cytokines are produced in response to viral infection?
  - A. Interferons
  - B. Monokines
  - C. Interleukins
  - D. Lymphokines
2. Which of the following does not characterize adaptive immunity?
  - A. Immunogenic memory,
  - B. Non-specific antigens,
  - C. Self- and non-self-recognition,
  - D. Diversity
3. Which of the following does not constitute a desirable antigen?
  - A. Large size,
  - B. Foreignness,
  - C. High complexity, and
  - D. Only reproducible by binary fission
4. What particular antibody causes a main immune response?
  - A. IgG,
  - B. IgM,
  - C. IgA,
  - D. IgE.
5. Which of the following substances is NOT present in tears?
  - A. IgA,
  - B. IgE,
  - C. Lactoferrin,
  - D. Lysozyme
6. What type of dendritic cell is not used as an antigen-presenting cell?
  - A. Follicular dendritic cell,
  - B. Langerhans cell,
  - C. Myeloid dendritic cell,
  - D. Lymphoid dendritic cell.

7. Identify the macrophages connected to the brain and spinal cord.
- A. Alveolar macrophages,
  - B. Kupffer cells,
  - C. Mesangial cells,
  - D. Microglial cells.
8. Which chemical does the dendritic cell constitutively express?
- A. The Class I MHC;
  - B. The Class II MHC
  - C. APC
  - D. Antigen.
9. Which of the following cells is not a part of polymorphonuclear leukocytes?
- A. Eosinophils
  - B. Mast cell.
  - C. Basophils,
  - D. Macrophages.
10. Which class of pattern-recognition molecules serves only as signaling receptors?
- A. CRP,
  - B. Toll-like receptor,
  - C. MBL,
  - D. LPS.
11. Which of the following cells produces Anti-bodies?
- A. macrophage.
  - B. B-cells.
  - C. T-cells
  - D. plasma cells
12. Which of the following Antibodies cross the placenta?
- A. IgA.
  - B. IgE
  - C. IgG.
  - D. IgD.
13. The branch of biology, which involves the study of immune systems in all organisms is called\_\_\_\_\_.
- A. Zoology
  - B. Microbiology

- C. Immunology
  - D. Biotechnology
14. Which of the following immunity is obtained during a lifetime?
- A. Acquired immunity
  - B. Active immunity
  - C. Passive immunity
  - D. None of the above.
15. How many types of antibodies are there?
- A. Five.
  - B. Three.
  - C. Two.
  - D. Four.
16. Which of the following cells is involved in cell-mediated immunity?
- A. Leukaemia
  - B. T cells
  - C. Mast cells
  - D. Thrombocytes
17. Which of the following protects our body against disease-causing pathogens?
- A. Respiratory system
  - B. Immune system
  - C. Digestive system
  - D. Respiratory system
18. Hepatitis is an example of \_\_\_\_\_.
- A. Subunit Vaccine
  - B. Killer Vaccine
  - C. Toxoids Vaccine
  - D. Recombinant Vaccine
19. Which of the following statements is true about the IgM of humans?
- A. IgM can cross the placenta
  - B. IgM can protect the mucosal surface
  - C. IgM is produced by high-affinity plasma cells
  - D. IgM is primarily restricted in the circulation
20. Interferons are
- A. Cytokine barriers

- B. Physical barriers
  - C. Cellular barriers
  - D. Physiological barriers
21. Which of the following cells of the immune system do not perform phagocytosis?
- A. Macrophage
  - B. Neutrophil
  - C. Eosinophil
  - D. Basophil
22. Monocytes differentiate into which kind of phagocytic cells?
- A. Neutrophil
  - B. B cell
  - C. Macrophage
  - D. T cell
23. Which immune system component is responsible for engulfing and destroying pathogens through phagocytosis?
- A. Antibodies
  - B. T cells
  - C. Dendritic cells
  - D. Phagocytes
24. The process by which immune cells engulf and destroy pathogens is called:
- A. Antigen presentation
  - B. Opsonization
  - C. Clonal expansion
  - D. Phagocytosis
25. The complement system consists of which of the following pathways?
- A. Adaptive, innate, and alternative pathways
  - B. Classical, adaptive, and innate pathways
  - C. Classical, lectin, and alternative pathways
  - D. Adaptive, lectin, and alternative pathways
26. Which complement proteins involved in opsonization and enhances phagocytosis of pathogens?
- A. C3b
  - B. C5a
  - C. C4a



- D. C3a
27. Fever is an innate immune response triggered by the release of:
- A. Cytokines
  - B. Antibodies
  - C. Pyrogens
  - D. Histamine
28. Which cells produce and release antibodies during an immune response?
- A. T cells
  - B. Phagocytes
  - C. Plasma cells
  - D. Natural killer cells
29. Which immunologist is credited with the discovery of antibodies in the late 19th century?
- A. Louis Pasteur
  - B. Paul Ehrlich
  - C. Karl Landsteiner
  - D. Emil von Behring
30. The “side-chain theory” to explain the specificity of antibody-antigen interactions was proposed by:
- A. Robert Koch
  - B. Paul Ehrlich
  - C. Elie Metchnikoff
  - D. Louis Pasteur
31. The process by which immune cells engulf and destroy pathogens is known as:
- A. Inflammation
  - B. Opsonization
  - C. Membrane attack complex
  - D. Phagocytosis
32. Which component of the innate immune system directly kills infected or abnormal cells?
- A. Dendritic cells
  - B. B cells
  - C. Natural Killer (NK) cells
  - D. Phagocytes

33. Which immune response is characterized by rapid and heightened protection upon re-exposure to the same antigen?
- Primary immune response
  - Secondary immune response
  - Innate immune response
  - Adaptive immune response
34. Which of the following is NOT a function of the complement system?
- Opsonization
  - Phagocytosis
  - Membrane attack complex formation
  - Antibody production
35. The complement system can be activated by:
- Pathogen-associated molecular patterns(PAMPs)
  - Antibodies
  - Cytokines
  - All of the above
36. The discovery of the ABO blood groups had significant implications for:
- Cancer treatment
  - Blood transfusions and transplantation
  - Vaccine development
  - Autoimmune diseases
37. Which immune cells are responsible for antigen presentation to activate the adaptive immune response?
- B cells
  - T cells
  - Plasma cells
  - Dendritic cells
38. What is the name of the process through which lymphocytes proliferate and differentiate upon encountering an antigen?
- Clonal expansion
  - Opsonization
  - Immunization
  - Phagocytosis

39. The regulatory proteins in the complement system help prevent:
- A. Inflammation
  - B. Autoimmune reactions
  - C. Phagocytosis
  - D. Fever
40. Which complement protein forms pores in the membranes of pathogens, leading to cell lysis?
- A. C3b
  - B. C5a
  - C. C4a
  - D. Membrane attack complex (MAC)
41. The process by which immune cells present antigens to activate the adaptive immune response is called:
- A. Opsonization
  - B. Clonal selection
  - C. Phagocytosis
  - D. Antigen presentation
42. Which complement pathway is primarily activated by antigen-antibody complexes?
- A. Classical pathway
  - B. Lectin pathway
  - C. Alternative pathway
  - D. Innate pathway
43. Which component of the complement system triggers the release of inflammatory mediators that promote vasodilation and increase vascular permeability?
- A. C3b
  - B. C5a
  - C. C4a
  - D. Anaphylatoxins
44. Which arm of the adaptive immune response provides defense against extracellular pathogens and toxins?
- A. Cell-mediated immunity
  - B. Humoral immunity
  - C. Innate immunity

- D. Non-specific immunity
45. What are the Y-shaped proteins responsible for antigen recognition in humoral immunity?
- A. Tcells
  - B. Cytokines
  - C. Immunoglobulins
  - D. Antigens
46. Which of the following functions is NOT performed by antibodies in humoral immunity?
- A. Neutralization
  - B. Cell lysis
  - C. Opsonization
  - D. Complement activation
47. B cell activation occurs when a B cell encounters its specific antigen. Where does this typically occur?
- A. Bone marrow
  - B. Thymus
  - C. Spleen
  - D. Liver
48. What is the most abundant antibody in the blood stream that provides long-term protection and crosses the placenta to provide passive immunity to the fetus?
- A. IgM
  - B. IgG
  - C. IgA
  - D. IgE
49. Which antibody is predominantly found in mucosal secretions, providing localized defense against pathogens?
- A. IgM
  - B. IgG
  - C. IgA
  - D. IgE
50. What is the function of Ig D in humoral immunity?
- A. Activates the complement system
  - B. Provides long-term protection

- C. Serves as a cell surface receptor on B cells
- D. Neutralizes pathogens

### ANSWERS

1-B,2-D,3-B,4-B,5-B,6-A,7-D,8-B,9-D,10-B,11-D,12-C,13-D,14-A,15-A,16-B,17-A,18-D,19-D,20-A,21-D,22-D,23-D,24-D,25-C,26-C,27-C,28-C,29-C,30-D,31-B,32-D,33-D,34-D,35-D,36-B,37-B,38-A,39-A,40-B,41-D,42-D,43-A,44-B,45-C,46-C,47-B,48-B,49-C,50-C

### Unit V

#### **Biostatistics I**

- Calculations of mean median, Madian, Mode, Variance, Standard Deviation
- Concept of coefficient of variation, Skewness, Kurtosis
- Elementary idea of probability and application

**Biostatistics** is a hybrid branch which includes Biology and Statistics

**Statistics** – applying statistical methods to biological sciences

**Biostatistics** is a science of measurement, analysis, collection and interpretation of data

#### **Measures of Central tendency**

- If there is a data and anyone want to draw some conclusion/ or to certain analysis/ or their type of data summarization then sometimes you need measures of central tendency
- Generally it is found that in any distribution values of variables tend to congregate around a central value of distribution
- This tendency of distribution is known as central tendency and the measure devised to consider this tendency are known as central tendency

## Arithmetic mean

### 1. Series of individual observation (**individual series**)

$$\text{Mean} = \frac{\text{summing up of all the observation}}{\text{Total no. of observation}}$$

$$\bar{X} = \frac{\sum x}{n}$$

$\bar{X}$  = arithmetic mean

$\sum x$  = sum of all values of variables  
i.e.  $x_1 + x_2 + x_3 + \dots + x_n$

$n$  = no. of observations

e.g. Calculate the arithmetic mean of the following set of observations

**7, 6, 8, 10, 13, 14, 18, 20, 22, 26, 30**

$$\begin{aligned}\sum x &= 7+6+8+10+13+14+18+20+22+26+30 \\ &= 174\end{aligned}$$

$$\bar{X} = \frac{\sum x}{n}$$

$$\bar{X} = \frac{174}{11}$$

$$\begin{aligned}\bar{X} &= 15.818 \\ &= 15.82\end{aligned}$$

Q. 2 Calculate the mean of the Series - **10, 11, 17, 19, 23, 27, 34, 41, 58, 64, 75, 96**

$$\begin{aligned}\sum x &= 10+11+17+19+23+27+34+41+58+64+75+96 \\ &= 475\end{aligned}$$

$$\bar{X} = \frac{\sum x}{n}$$

$$\bar{X} = \frac{475}{12}$$

$$\begin{aligned}\bar{X} &= 39.583 \\ &= 39.58\end{aligned}$$



## Median

- Another important and widely used measures of central tendency
- Median is usually defined as that value which divides a distribution so that an equal no. of items occur on either side (divides the group into 2 equal halves, one half all values are greater than median)
- Middle value
- If no. of values is odd median will be middle value  
e.g. 70, 84, 86, 99, 103 (**median** = 86)

For even values.....there will be 2 middle value

e.g. 70, 84, 86, 99, 103, 107, 108, 109

99, 103 → middle value

$$\text{Median} = \frac{99+103}{2}$$

$$= 101$$

### Formula

$$\text{Median} = \left(\frac{n+1}{2}\right)^{\text{th item}} \quad n = \text{Number of items}$$

e.g. 70, 84, 86, 99, 103

First arrange data in either **ascending** or **descending** order then apply the formula

$$\text{Median} = \left(\frac{n+1}{2}\right)^{\text{th item}}$$

$$= \frac{5+1}{2}$$

$$= 3^{\text{rd}} \text{ item}$$

$$= 86$$

Q. In the given series find out the median

10, 18, 17, 19, 10, 15, 11, 17, 12

First arrange data in either **ascending** or **descending** order then apply the formula

10, 10, 11, 12, 15, 17, 17, 18, 19

$$\text{Median} = \left(\frac{n+1}{2}\right)^{\text{th item}}$$

$$\begin{aligned}
 &= \frac{9+1}{2} \\
 &= 5^{\text{th}} \text{ item} \\
 &= \mathbf{15}
 \end{aligned}$$

Q. In the given series find out the median

3.9, 4.0, 5.7, 4.2, 6.6, 7.0, 7.9, 8.0, 9.0, 10.0

First arrange data in either **ascending** or **descending** order then apply the formula

3.9, 4.0, 4.2, 5.7, 6.6, 7.0, 7.9, 8.0, 9.0, 10.0

$$\begin{aligned}
 \text{Median} &= \left(\frac{n+1}{2}\right)^{\text{th}} \text{ item} \\
 &= \mathbf{6.8}
 \end{aligned}$$

### Mode

- Another important measures of central tendency
- Mode means most commonly occurring value
- It is the size of that item which has the maximum frequency (no. of repeats)  
So mode is that value which has maximum no. of repeats.

#### Calculation of mode:

S, No.	1	2	3	4	5	6	7	8	9	10
Wt. (kg)	10	11	10	12	12	11	9	8	11	11

Calculate the mode from a sample of 10 animals

Weights	No. of Animals
8	1
9	1
10	2 <b>Mode = ...</b>
11	4
12	2

S, No.	1	2	3	4	5	6	7	8	9	10
tillers	10	24	27	12	24	24	20	18	15	30

**Mode = 24**

### Measures of dispersion

- Degree of variation of the variables about some central value
- Measures of central tendency
  - ↓
  - Gives the idea of central value
- Measures of dispersion
  - ↓
  - gives an idea of degree of variation of the variables about a central value (measure of spread about a central value)
- Actually it measures the extent to which there are differences between individual observations and the central value

- Dispersion may be defined as the measure of the variation of the variable

e.g.	A	B	C
	16	14	4
	16	15	8
	16	16	16
	16	17	24
	<u>16</u>	<u>18</u>	<u>28</u>
<b>Total</b>	<b>80</b>	<b>80</b>	<b>80</b>
<b>Mean</b>	<b>16</b>	<b>16</b>	<b>16</b>

- In A all the values are same so there is no dispersion
- In B one value is represented by the mean and other values show a slight variation (small dispersion)
- In C the values deviate very widely from one another (large dispersion)
- So dispersion in **C > B**

### Methods of Measuring dispersion

- I. Range
- II. Mean deviation
- III. Variance and standard deviation

#### Variance (s<sup>2</sup>)

- Also called mean square deviation
- Term introduced by R. A. Fisher
- The term variance is also used to describe as “square of the standard deviation”
- It is also defined as the sum of squared differences between each individual value and mean divide by 1 less than the total no. of observations (n)

when sample size is < 30 then apply (n-1)

when sample size is > 30 then apply (n)

$$s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$$

X = variables

$\bar{x}$  = mean

n = No. of observations

- So variance indicate the extent to which variables within the distribution depart from the mean

If  $s^2$  is high  $\rightarrow$  values are spread

If  $s^2$  is small  $\rightarrow$  values are close to each other

### Standard deviation (s) or S.D.

- Defined as the square root of the variance

$$\text{S.D.} = \sqrt{s^2}$$

$$\text{S.D.} = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}}$$

- Most commonly used absolute measure of dispersion
- Introduced by Karl Pearson
- Measure of spread in a set of observations

### **Coefficient of variance (CV)**

$$\text{CV} = \frac{\text{S.D.}}{\bar{X}} \times 100$$

S.D. = Standard deviation  
 $\bar{X}$  = mean

- If CV is smaller data is consistent i.e. less variation in data
- If CV is greater data is said to be more variable
- CV is a relative measure of variation

e.g. Series A  $\bar{X} = 34.60$ , S.D. = 14

Series B  $\bar{X} = 27.10$ , S.D. = 8

$\text{CV} = \frac{\text{Series A}}{\bar{X}} \times 100$ $= \frac{14}{34.60} \times 100$ $= 40.46\%$	$\text{CV} = \frac{\text{Series B}}{\bar{X}} \times 100$ $= \frac{8}{27.10} \times 100$ $= 29.52\%$
--	---

So the result shows that in series A data is more variable

### Skewness

Skewness is a statistical measure that quantifies the asymmetry of a probability distribution or a dataset. It provides insights into the shape and nature of the data distribution, specifically focusing on whether it is symmetric, positively skewed, or negatively skewed.

Skewness is calculated based on the third standardized moment of a dataset. The formula for calculating skewness depends on the method used, but a common method is Pearson's First Coefficient of Skewness (also known as the moment coefficient of skewness). This method is given by the following formula:

$$\text{Skewness} = (3 * (\text{Mean} - \text{Median})) / \text{Standard Deviation}$$

No Skewness - If the Skewness is zero,

Skewness > 0 (Positive Skewness)

Skewness < 0 (Negative Skewness)

In summary, skewness is a statistical measure that assesses the asymmetry of a dataset or probability distribution.

### Kurtosis

Kurtosis is a statistical measure that quantifies the shape of a probability distribution or a dataset. It provides insights into the peakedness or flatness of the distribution and helps assess the presence of outliers or extreme values.

Kurtosis is calculated based on the fourth standardized moment of a dataset. The formula for calculating kurtosis depends on the method used, but a common method is Pearson's Coefficient of Kurtosis (also known as excess kurtosis). This method calculates kurtosis by subtracting 3 from the fourth moment about the mean and dividing by the standard deviation to provide a standardized measure.

**Positive kurtosis:** If the kurtosis value is positive, it indicates that the dataset has a higher peak.

**Negative kurtosis:** If the kurtosis value is negative, it indicates that the data set has a flatter peak.

### Elementary idea of probability

The elementary idea of probability is a fundamental concept in mathematics and statistics that deals with the measurement and analysis of uncertainty and likelihood. It provides a framework for quantifying and understanding the likelihood of different events or outcomes occurring.

Probability is defined as a number between 0 and 1, where 0 represents an event that is impossible, and 1 represents an event that is certain to occur. The concept of probability is

based on the understanding that in a random or uncertain situation, certain outcomes are more likely to occur than others.

### Questions

1. The mean of the square deviation about mean is known as;
  - A. Mean
  - B. Median
  - C. Variance
  - D. Standard deviation
2. Which of the following is not a measure of central tendency?
  - A. Range
  - B. Mean
  - C. Mode
  - D. Median
3. The sum of all data values divided by the number of values is called
  - A. Median
  - B. Mean
  - C. Range
  - D. Mode
4. In a positively skewed distribution, the relationship between mean median and mode is
  - A. Mean = Median = Mode
  - B. Mode < Median < Mean
  - C. Mean < Median < Mode
  - D. Mean > Median > Mode
5. When the dataset has an even number of observations, the median is calculated as
  - A. The largest value in the dataset
  - B. The smallest value in the dataset
  - C. The middle value in the dataset
  - D. The average of the two middle values
6. In a symmetrical distribution the relationship between mean, median, and mode is
  - A. Mean=Median=Mode
  - B. Mean<Median<Mode
  - C. Mode<Median<Mean
  - D. Mean>Median>Mode



7. When the data has a negative skew, the relationship between mean median and mode is
- A. Mean = Median = Mode
  - B. Mode < Median < Mean
  - C. Mean < Median < Mode
  - D. Mean > Median > Mode
8. Which of the following is **not** a measure of central tendency?
- A. Mean & Median
  - B. Median & Mode
  - C. Mean Median & Mode
  - D. Range, Variance & SD
9. Find out the median in the following series –  
**19, 21, 20, 17, 18, 23, 22**
- A. 20
  - B. 19
  - C. 18
  - D. 21
10. Median is
- A. Middle value
  - B. Positional value
  - C. Not affected by the value of extreme ends
  - D. All of the above
11. Most frequently occurring variable in a distribution is called
- A. Mean
  - B. Median
  - C. Mode
  - D. Range
12. Values of median & mode of a series are **14** & **11** respectively. What will be the value of mean?
- A. 9
  - B. 15.5
  - C. 8
  - D. 13.5
13. Which of the following is a measure of Dispersion?

- A. Correlation  
B. Standard Deviation  
C. Variance  
D. Both B & C
14. Standard deviation is  
A. Square of the Variance  
B. Square root of the Variance  
C. co-efficient of Range  
D. All of the above
15. Skewness is  
A. defined as lack of symmetry  
B. absent in symmetrical distribution  
C. is affected by the mean, median & mode  
D. All of the above
16. If a curve is more peaked than the normal curve, it is called  
A. Leptokurtic  
B. Mesokurtic  
C. Platykurtic  
D. None of the above
17. The mode represents the:  
A. Middle value of a dataset  
B. Most frequently occurring value in a dataset  
C. Average of all values in a dataset  
D. Total sum of all values in a dataset
18. Probability theory is an essential tool in various fields because it helps in:  
A. Eliminating outliers in datasets  
B. Making informed decisions in the face of uncertainty  
C. Guaranteeing a certain outcome in every situation  
D. Reducing the sample space to only one possible outcome
19. In statistics, probability forms the basis for:  
A. Weather forecasting  
B. Hypothesis testing  
C. Game development  
D. Population genetics

20. The probability of rolling an odd number on a fair six-sided die is:
- A. 0.5
  - B. 1
  - C. 0
  - D. 0.333
21. The probability of the complement of an event A (not A) is:
- A. 1
  - B. 0
  - C. -1
  - D. 0.5
22. The elementary idea of probability deals with:
- A. Measuring and analyzing uncertainty and likelihood
  - B. Calculating the mean of a dataset
  - C. Identifying outliers in a distribution
  - D. Measuring the Skewness of a dataset
23. A distribution with negative kurtosis indicates:
- A. A tail that extends towards higher values
  - B. A tail that extends towards lower values
  - C. A flatter peak compared to a normal distribution
  - D. A higher peak compared to a normal distribution
24. Kurtosis is a statistical measure that quantifies the:
- A. Symmetry of a distribution
  - B. Variability of datasets
  - C. Peakedness or flatness of a distribution
  - D. Standard deviation of a dataset
25. A distribution with positive skewness indicates:
- A. A perfectly symmetrical distribution
  - B. A tail that extends towards higher values
  - C. A tail that extends towards lower values
  - D. No skewness
26. The coefficient of variation (CV) is useful in which of the following scenarios?
- A. Analyzing outliers

- B. Comparing variability of datasets  
C. Calculating the mean of datasets  
D. Measuring skewness
27. A higher coefficient of variation (CV) indicates:  
A. Greater relative variability  
B. Lower relative variability  
C. No variability  
D. Negative variability
28. The coefficient of variation (CV) is a measure used to compare the relative variability of datasets with different scales or means. It is expressed as a percentage and calculated as:  
A.  $(\text{Mean} / \text{Standard Deviation}) * 100$   
B.  $(\text{Standard Deviation} / \text{Mean}) * 100$   
C.  $(\text{Mean} * \text{Standard Deviation}) * 100$   
D.  $(\text{Standard Deviation} - \text{Mean}) * 100$
29. The primary purpose of biostatistics is to:  
A. Analyze and interpret data related to living organisms  
B. Design computer algorithms for biological research  
C. Develop new medical devices and technologies  
D. Study the chemical composition of biological systems
30. Which measure of central tendency is used to represent the most frequently occurring value in a dataset?  
A. Standard deviation  
B. Median  
C. Mean  
D. Mode
31. The standard deviation is a measure of:  
A. Central tendency  
B. Spread or dispersion of data points around the mean  
C. Frequency of values in a dataset  
D. The total sum of all values in a dataset
32. Biostatistics is used to evaluate the effectiveness of:  
A. Social media campaigns

- B. New drugs and medical devices
  - C. Space exploration missions
  - D. Sports events and tournaments
33. Which statistical measure is sensitive to extreme values or outliers?
- A. Median
  - B. Variance
  - C. Standard deviation
  - D. Mode
34. Biostatistics is applied to understand and solve problems related to:
- A. Human health and biology
  - B. Environmental sustainability
  - C. Technology and engineering
  - D. Economic growth and development
35. Which measure of central tendency is the average value of a set of numbers?
- A. Median
  - B. Variance
  - C. Mean
  - D. Mode
36. Biostatistics plays a critical role in the design and analysis of:
- A. Art exhibitions
  - B. Clinical trials and experiments
  - C. Fashion shows
  - D. Social media campaigns
37. The variance represents the:
- A. Middle value of a dataset
  - B. Spread or dispersion of data points around the mean
  - C. Most frequently occurring value in a dataset
  - D. Total sum of all values in a dataset
38. The mode is the value that:
- A. Represents the average of all values in a dataset
  - B. Appears with the highest frequency in a dataset
  - C. Represents the middle value of a dataset

- D. Represents the sum of all values in a dataset
39. The mode represents the:
- A. Middle value of a dataset
  - B. Most frequently occurring value in a dataset
  - C. Average of all values in a dataset
  - D. Total sum of all values in a dataset
40. Biostatistics is closely associated with which field of study?
- A. Epidemiology
  - B. Physics
  - C. Geology
  - D. Astronomy
41. Which statistical measure is used to evaluate the volatility or risk associated with an investment?
- A. Mean
  - B. Variance
  - C. Standard deviation
  - D. Mode
42. Biostatistics helps in making evidence-based decisions in:
- A. Social sciences
  - B. Healthcare and medicine
  - C. Sports and entertainment
  - D. Agriculture and farming
43. What does the standard deviation measure about a dataset?
- A. The average value of the dataset
  - B. The spread or dispersion of data points around the mean
  - C. The most frequently occurring value in the dataset
  - D. The difference between the largest and smallest values in the dataset
44. The median represents the:
- A. Middle value of a dataset
  - B. Most frequently occurring value in a dataset
  - C. Average of all values in a dataset
  - D. Total sum of all values in a dataset



45. Which measure of central tendency is not affected by extreme values or outliers?
- A. Variance
  - B. Mean
  - C. Median
  - D. Standard deviation
46. What is the purpose of biostatistics?
- A. To collect and analyze data related to living organisms
  - B. To study chemical reactions in biological systems
  - C. To analyze data related to climate change
  - D. To design computer algorithms for healthcare
47. What does the standard deviation measure?
- A. Spread or dispersion of data points around the mean
  - B. Most frequently occurring value in a dataset
  - C. Variability of the dataset with respect to the mode
  - D. The total sum of squared differences in a dataset
48. Biostatistics is closely associated with which field of study?
- A. Epidemiology
  - B. Physics
  - C. Geology
  - D. Astronomy
49. In genetics and biology, probability is used to analyze:
- A. Economic indicators
  - B. Price movements
  - C. Genetic inheritance and predict the likelihood of genetic traits or diseases
  - D. Weather patterns
50. The probability measure assigns a probability value to each:
- A. Random experiment
  - B. Event in the sample space
  - C. Outlier in a dataset
  - D. Skewness value

**ANSWERS**

1-C,2-A,3-B,4-D,5-D,6-A,7-C,8-D,9-A,10-D,11-C,12-B,13-D,14-B,15-D,16-A,17-B,18-B,19-B,20-D,21-B,22-A,23-C,24-C,25-B,26-B,27-A,28-B,29-A,30-D,31-B,32-B,33-C,34-A,35-C,36-B,37-B,38-B,39-B,40-A,41-C,42-B,43-B,44-C,45-C,46-A,47-A,48-A,49-C,50-B

**Unit VI****Biostatistics II**

- Data Summarizing: frequency distribution, graphical representation, pie diagram, histogram
- Tests of significance: one and Two sample tests, t-test and Chi-square test

Data summarization is a crucial step in the data analysis process that aims to provide a concise and meaningful representation of the data, allowing analysts and decision-makers to gain insights quickly and make informed conclusions

**Frequency distribution**

Frequency distribution is a statistical technique used to organize and summarize data by displaying the frequency or count of observations falling into various categories or intervals. It provides a systematic and concise representation of the data.

**Graphical representation**

Graphical representation, also known as data visualization, is a powerful technique that uses visual elements such as charts, graphs, and maps to represent data and convey information effectively. It provides a visual depiction of data patterns, trends, and relationships, allowing users to understand complex information quickly and intuitively.

**Pie diagram**

A pie diagram, also known as a pie chart, is a circular graphical representation used to display categorical data or proportions. It is an effective way to visualize the distribution of a single categorical variable or the relative proportions of different categories within a whole.

**Histogram**

A histogram is a graphical representation that organizes and displays the distribution of a continuous variable. It provides a visual representation of the frequency or count of observations falling into different intervals or bins along a continuous scale. Histograms are widely used in statistics and data analysis to understand the shape, central tendency, variability, and outliers within a dataset.

**Test of significance**

A test of significance, also known as a statistical hypothesis test, is a method used to determine the statistical significance of an observed effect or relationship in a sample. The

null hypothesis represents the assumption of no effect, no difference, or no relationship between variables in the population. The alternative hypothesis represents the opposite of the null hypothesis, suggesting the presence of an effect, difference, or relationship. The significance level, denoted as  $\alpha$ , is the predetermined threshold that defines the level of evidence required to reject the null hypothesis.

**One-Sample t-test:** The one-sample t-test is used to compare the mean of a single sample to a known or hypothesized population mean. The null hypothesis states that there is no significant difference between the sample statistic (mean or proportion) and the hypothesized population value. The alternative hypothesis represents the opposite, suggesting a significant difference

### Two sample test

- **Independent Samples t-test:** The independent samples t-test compares the means of two independent samples to assess whether there is a significant difference between them. It is suitable for continuous data and assumes that the two groups are independent and normally distributed.
- **Paired Samples t-test:** The paired samples t-test compares the means of two related or paired samples. It is used when the data is collected from the same subjects or units under different conditions or at different time points. It assesses whether there is a significant difference between the paired measurements.

### Chi-square test

The chi-square test is a statistical test used to determine whether there is a significant association or relationship between two categorical variables. It is based on the chi-square statistic, which measures the difference between the observed frequencies and the expected frequencies under the assumption of independence.

- Chi-Square Test of Independence
- Chi-Square Goodness of Fit Test

### Questions

1. Data summarization aims to:
  - A. Increase data complexity
  - B. Remove outliers from the data
  - C. Provide a concise representation of data
  - D. Manipulate data for visualization
2. Which step comes first in the data summarization process?

- A. Data Visualization
  - B. Data Cleaning
  - C. Dimensionality Reduction
  - D. Aggregation and Grouping
3. Descriptive statistics include measures like:
- A. Mode and Range
  - B. Standard Deviation and Variance
  - C. Mean and Median
  - D. Skewness and Kurtosis
4. Data visualization techniques are used to:
- A. Hide patterns in the data
  - B. Make data complex for analysis
  - C. Present data visually
  - D. Increase data dimensions
5. Aggregation in data summarization involves:
- A. Reducing data complexity
  - B. Dividing data into categories
  - C. Combining multiple data points in to a single value
  - D. Handling missing values in data
6. Dimensionality reduction techniques are applied to:
- A. Add more variables to the data set
  - B. Summarize data with the most important features
  - C. Remove outliers from the data
  - D. Visualize data with 3 Deffects
7. Frequency distribution is used to summarize data by:
- A. Representing data with pie charts
  - B. Displaying the distribution of a continuous variable
  - C. Organizing data in to interval sorbins
  - D. Identifying central tendency in the data
8. Cumulative frequency represents:
- A. The frequency of the largest category
  - B. The total count of observations up to a specific interval
  - C. The average frequency of all intervals
  - D. The number of intervals in a frequency table

9. A bar chart is useful for visualizing:
  - A. The distribution of continuous data
  - B. The frequency distribution of categorical data
  - C. Trends overtime
  - D. The relationship between two continuous variables
10. A scatter plot is used to display:
  - A. The distribution of a continuous variable
  - B. The frequency distribution of categorical data
  - C. The relationship between two continuous variables
  - D. The proportion of each category in a whole
11. Which chart is best suited to show the composition of a whole?
  - A. Scatter plot
  - B. Histogram
  - C. Pie chart
  - D. Bar chart
12. The central angle of each slice in a pie chart represents:
  - A. The proportion of data points in that category
  - B. The width of the interval in the frequency distribution
  - C. The significance level in hypothesis testing
  - D. The mean of the data points in that category
13. Histograms are used to represent the distribution of:
  - A. Categorical data
  - B. Discrete variables
  - C. Continuous variables
  - D. Cumulative frequencies
14. The height of each bar in a histogram represents:
  - A. The frequency or count of observations in each interval
  - B. The proportion of data points in that category
  - C. The significance level in hypothesis testing
  - D. The width of the interval in the frequency distribution
15. A significance level( $\alpha$ ) in hypothesis testing is used to:
  - A. Determine the frequency distribution of data
  - B. Set the width of intervals in a histogram
  - C. Define the level of evidence required to reject the null hypothesis



- D. Calculate the central angle in a pie chart
16. A test of significance is also known as:
- A. Data summarization
  - B. Data visualization
  - C. Statistical hypothesis test
  - D. Frequency distribution
17. The null hypothesis ( $H_0$ ) in a significance test represents:
- A. The assumption of no effector relationship
  - B. The observed data in the sample
  - C. The data cleaning process
  - D. The alternative hypothesis( $H_1$ )
18. The P-value in a significance test quantifies:
- A. The probability of a Type I error
  - B. The strength of evidence against the null hypothesis
  - C. The central tendency of the data
  - D. The proportion of outliers in the data
19. Which test statistic is commonly used to compare means between two groups?
- A. Chi-square test
  - B. T-test
  - C. ANOVA
  - D. Regression analysis
20. Histograms are used to visualize the distribution of which type of data?
- A. Categorical data
  - B. Discrete data
  - C. Continuous data
  - D. Ordinal data
21. What does the width of each interval represent in a histogram?
- A. Frequency of observations in each interval
  - B. Proportion of data points in each interval
  - C. Size of the sample
  - D. Size of the entire population
22. The purpose of data cleaning in the data summarization process is to:
- A. Create 3D effects in visualizations
  - B. Handle missing values and outliers



- C. Reduce the number of intervals in the frequency distribution  
D. Convert continuous data to categorical data
23. Which graphical representation is used to display the distribution of categorical data?  
A. Scatter plot  
B. Pie chart  
C. Histogram  
D. Line chart
24. In a pie chart, the central angle of each slice is proportional to:  
A. The number of categories  
B. The percentage of data points in that category  
C. The width of each interval in the histogram  
D. The number of observations in each group
25. The primary purpose of a frequency distribution is to:  
A. Provide a concise representation of data  
B. Organize data into intervals or bins  
C. Create a visual representation of data  
D. Compare different data sets or subgroups
26. The one-sample t-test is used to compare the mean of a single sample to:  
A. Another sample's mean  
B. A known or hypothesized population mean  
C. The median of the same sample  
D. The variance of the same sample
27. The one-sample Z-test is preferred over the t-test when:  
A. The population standard deviation is unknown  
B. The sample size is large  
C. The sample mean is known  
D. The sample is not normally distributed
28. The one-sample proportion test is suitable for analyzing data with:  
A. Continuous variables  
B. More than two categories  
C. Categorical data with two possible outcomes  
D. Ordinal data
29. The null hypothesis in a one-sample test states that:  
A. The sample mean is equal to the population mean

- B. The sample mean is greater than the population mean  
C. The sample mean is less than the population mean  
D. There is no significant difference between the sample mean and the population mean
30. The significance level( $\alpha$ ) is used to:
- A. Determine the sample size  
B. Set the range of values in the data  
C. Define the threshold for rejecting the null hypothesis  
D. Calculate the standard deviation of the sample
31. The independent samples t-test compares:
- A. The means of two dependent samples  
B. The means of two related samples  
C. The means of two independent samples  
D. The means of two non-normal samples
32. The paired samples t-test is appropriate when:
- A. The sample size is large  
B. The samples are independent  
C. The data is continuous and normally distributed  
D. The samples are related or paired
33. The chi-square test of independence examines the relationship between:
- A. Two continuous variables  
B. Two categorical variables  
C. A categorical and a continuous variable  
D. A categorical and an ordinal variable
34. The degrees of freedom for a chi-square test of independence are calculated based on:
- A. The sample size of the two groups  
B. The number of rows and columns in the contingency table  
C. The variance of the sample data  
D. The number of categories in each group
35. The chi-square goodness of fit test is used to:
- A. Compare the means of two independent samples  
B. Assess the independence of two categorical variables  
C. Determine if the observed frequencies match an expected distribution

- D. Test the significance of a single sample's mean
36. The t-test is used to compare:
- A. The means of two independent samples
  - B. The means of two related samples
  - C. The means of two continuous variables
  - D. The means of two categorical variables
37. The paired samples t-test is appropriate when:
- A. The data is continuous and normally distributed
  - B. The samples are independent
  - C. The samples are related or paired
  - D. The population standard deviation is known
38. The chi-square test is used to analyze:
- A. Continuous data
  - B. Categorical data
  - C. Ordinal data
  - D. Discrete data
39. The chi-square test of independence is used to examine the relationship between:
- A. Two continuous variables
  - B. Two categorical variables
  - C. A categorical and a continuous variable
  - D. A continuous and an ordinal variable
40. The chi-square goodness of fit test is used to assess:
- A. The difference between two sample means
  - B. The relationship between two categorical variables
  - C. Whether the observed frequencies match an expected distribution
  - D. The significance of a single sample's mean
41. The null hypothesis( $H_0$ ) in a one-sample test states that:
- A. There is a significant difference between the sample mean and the population mean
  - B. The sample mean is equal to the population mean
  - C. The sample mean is greater than the population mean
  - D. The sample mean is less than the population mean
42. The significance level( $\alpha$ ) in a one-sample test determines:
- A. The size of the sample

- B. The threshold for rejecting the null hypothesis  
C. The variance of the population  
D. The range of values in the data
43. The independent samples t-test is used when the two groups are:  
A. Related or paired  
B. Independent and normally distributed  
C. Dependent and non-normally distributed  
D. Categorical
44. The paired samples t-test is used when the data is collected from:  
A. Two independent groups  
B. Two related or paired groups  
C. Two non-normally distributed groups  
D. Two large groups
45. The chi-square test is appropriate for analyzing data with:  
A. Continuous variables  
B. More than two categories  
C. Categorical data with two possible outcomes  
D. Ordinal data
46. The chi-square test of independence examines whether there is a relationship between:  
A. Two continuous variables  
B. Two categorical variables  
C. A categorical and a continuous variable  
D. A categorical and an ordinal variable
47. Which test compares the means of two independent samples?  
A. One-sample t-test  
B. Independent samples t-test  
C. One-sample Z-test  
D. Chi-square test
48. The degrees of freedom for a chi-square test of independence are calculated as:  
A.  $(r - 1) \times (c - 1)$   
B.  $(r + c) \times (r - 1) \times (c - 1)$   
C.  $(r + c) \times (r + c - 1)$   
D.  $(r - 1) + (c - 1)$

49. The chi-square test is suitable for comparing:
- Two sample means
  - Two independent samples
  - The variability of two samples
  - The relationship between two categorical variables
50. The t-test is used when the data is:
- Continuous and normally distributed
  - Categorical and non-normally distributed
  - Ordinal and independent
  - Related or paired

### **ANSWERS**

1-C,2-B,3-C,4-C,5-C,6-B,7-C,8-B,9-B,10-C,11-C,12-A,13-C,14-A,15-C,16-C,17-A,18-B,19-B,20-C,21-A,22-B,23-B,24-B,25-A,26-B,27-B,28-C,29-D,30-C,31-C,32-D,33-B,34-B,35-C,36-A,37-C,38-B,39-B,40-C,41-B,42-B,43-B,44-B,45-C,46-B,47-B,48-A,49-D,50-A

### **Unit VII**

#### Basics of computers

- Basics (CPU, I/O Units) and operating systems
- Concept of homepages and websites, World Wide Web, URLs, Using Search engines

A computer is an electronic device capable of performing various tasks by executing instructions and processing data. It is composed of hardware components, such as the central processing unit (CPU), memory, storage devices, input/output devices, and software programs that enable it to carry out specific functions

#### **Central Processing Unit (CPU)**

The CPU is often referred to as the “brain” of the computer. It is responsible for executing instructions and performing calculations.

#### **Parts of a Computer**

##### **Input device**

- Keyboard
- Mouse
- Scanner
- Microphone
- Touchscreen

**Output Device**

- Monitor
- Printer
- Speaker
- Headphone
- Projector

**Operating System**

An operating system (OS) is a software program that acts as an interface between the user and the computer hardware. It manages the computer's resources, provides a platform for running applications, and enables users to interact with the computer system.

**Website**

A website is a collection of interconnected webpages that are accessible via the internet. It is a virtual space where individuals, businesses, organizations, and other entities can share information, provide services, and interact with users.

**Homepage**

The home page of a website is the main or introductory page that serves as the starting point for visitors when they access the site. It is typically the first page users encounter when they enter the website's URL or click on the website's link from search engine results or other sources.

**World Wide Web**

The World Wide Web (WWW), commonly known as the web, is a vast network of interconnected hypertext documents and resources accessible through the internet. It was developed by Sir Tim Berners-Lee in the late 1980s and has since become a fundamental component of the modern digital world.

**Uniform Resource Locators (URLs)**

Each web page on the web has a unique address called a Uniform Resource Locator (URL). URLs provide a standardized way to locate and access web pages. Users enter a URL in the address bar of a web browser to navigate to a specific web page.

**Search Engine**

A search engine is a software program or service that allows users to search and retrieve information from the internet. e.g. Google, Bing, Yahoo



**Questions**

1. Computers can execute instructions and process data at extremely high speeds, measured in terms of:
  - A. Bytes per second(B/s)
  - B. Clock cycles per second(Hz)
  - C. Characters per second(C/s)
  - D. Instructions per minute(IPM)
2. Which type of storage device allows for long-term data retention and quick access to information?
  - A. CPU cache memory
  - B. Random Access Memory (RAM)
  - C. Hard Disk Drives (HDDs)
  - D. Read-Only Memory (ROM)
3. Computers are known for their high accuracy in executing instructions, minimizing errors due to:
  - A. CPU over heating
  - B. Human fallibility
  - C. Insufficient storage capacity
  - D. Insufficient processing power
4. The versatility of computers allows them to be programmed to perform a wide range of tasks, including:
  - A. Baking cakes
  - B. Word processing
  - C. Driving cars
  - D. Painting pictures
5. Which characteristic of computers allows them to automate repetitive tasks and increase efficiency?
  - A. Speed
  - B. Versatility
  - C. Scalability
  - D. Automation
6. Computers can connect to networks, enabling communication and data exchange between different devices through:

- A. Sensors
  - B. Processors
  - C. Connectors
  - D. System buses
7. What does scalability in computers refer to?
- A. The ability to withstand physical impacts
  - B. The ability to adapt to changing requirements by adding or removing components
  - C. The ability to process data with high accuracy
  - D. The ability to display content on multiple screens simultaneously
8. Which characteristic of computers ensures their reliable operation over extended periods with low failure rates?
- A. Versatility
  - B. Reliability
  - C. Speed
  - D. Automation
9. The primary function of an operating system is to manage the computer's resources efficiently, including:
- A. Sending emails
  - B. Allocating CPU time to processes
  - C. Operating peripherals
  - D. Generating reports
10. The operating system schedules and prioritizes tasks, allowing multiple processes to run concurrently. This feature is known as:
- A. Multitasking
  - B. Multithreading
  - C. Multiprocessing
  - D. Multidimensional
11. Which type of memory is volatile and holds data and instructions temporarily during program execution?
- A. Cache memory
  - B. Read-Only Memory (ROM)
  - C. Random Access Memory(RAM)
  - D. Solid-State Drive (SSD)

12. The operating system manages the organization and access of files and directories on storage devices through its:
- A. Cache management system
  - B. File system management
  - C. Task scheduling algorithm
  - D. Memory allocation system
13. What is the function of device drivers in an operating system?
- A. Allocate memory to running processes
  - B. Manage file access permissions
  - C. Facilitate communication and coordination with hardware devices
  - D. Provide a user interface for interacting with the computer system
14. The operating system incorporates security measures to protect the computer system and its data, such as:
- A. Virtual memory management
  - B. Multitasking support
  - C. User authentication and access control mechanisms
  - D. File system integrity and security
15. What does an operating system use to handle errors, exceptions, and faults that may occur during system operation?
- A. Virtual memory
  - B. File system
  - C. Fault tolerance measures
  - D. Data compression techniques
16. A website is a collection of interconnected web pages that are accessible via the:
- A. Telephone network
  - B. Internet
  - C. Postal service
  - D. Satellite communication system
17. Hyperlinks are used in websites to:
- A. Encrypt data for security purposes
  - B. Connect web pages and allow users to navigate between them
  - C. Perform arithmetic calculations on the web server
  - D. Create interactive animations

18. Which element of a website enhances user experience by adapting to different screen sizes and resolutions?
- A. Hyperlinks
  - B. Multimedia content
  - C. Responsive design
  - D. Navigation
19. Websites can include various types of multimedia content, such as images, videos, audio files, and interactive elements. This enhances the:
- A. SEO ranking of the website
  - B. Aesthetic appeal and engagement of the website
  - C. Security of the website
  - D. Accessibility of the website
20. The content management system (CMS) of a website facilitates:
- A. Allocation of CPU time to different processes
  - B. Creation and editing of content without technical knowledge
  - C. File system management
  - D. System monitoring and diagnostics
21. Which feature of a website allows users to engage with the content or perform actions, such as data submission and search?
- A. Multimedia content
  - B. Navigation
  - C. Responsive design
  - D. Interactivity
22. In the context of web sites, accessibility refers to:
- A. The ability of a website to handle high traffic
  - B. The ability of a website to load quickly
  - C. The ability of a website to be easily found on search engines
  - D. The ability of a website to be usable by individuals with disabilities
23. What is the purpose of the user interface in a website?
- A. To manage the computer's resources efficiently
  - B. To connect to networks and enable data exchange
  - C. To facilitate communication between different devices
  - D. To allow users to interact with the website and perform tasks

24. Websites use system buses to facilitate communication and coordination between different components of the computer. Which of the following is NOT a component of the system bus?
- A. Address bus
  - B. Data bus
  - C. Control bus
  - D. Arithmetic Logic Unit (ALU)
25. Which type of memory is used to store permanent instructions and data that are essential for booting up the computer?
- A. Cache memory
  - B. Random Access Memory (RAM)
  - C. Hard Disk Drives (HDDs)
  - D. Read-Only Memory (ROM)
26. What is the main purpose of the homepage of a website?
- A. To show case the website's back end infrastructure
  - B. To create a strong first impression and guide users to explore further
  - C. To provide a detailed overview of the website's backend code
  - D. To host multimedia content for promotional purposes
27. Which of the following is NOT a characteristic of a homepage?
- A. Introduction and branding
  - B. Navigation and menu
  - C. Search engine algorithms
  - D. Visual design
28. The navigation menu on a homepage helps users:
- A. Access webpages with broken links
  - B. Locate resources on the internet
  - C. Explore different sections and pages of the website
  - D. Access the website's backend code
29. Which component of a homepage provides a brief description of the website's purpose and value proposition?
- A. Search bar
  - B. Introduction and branding
  - C. Social media integration
  - D. Multimedia content

30. What is the purpose of call-to-action elements on a homepage?
- A. To encourage users to navigate to other websites
  - B. To display multimedia content
  - C. To guide users to take specific actions, such as signing up or making a purchase
  - D. To provide a visual representation of the website's purpose
31. The World Wide Web is built on the concept of:
- A. Binary code
  - B. Hypertext
  - C. Assembly language
  - D. Machine learning
32. What are individual webpages written in to create content for the World Wide Web?
- A. Java Script
  - B. HTML
  - C. CSS
  - D. XML
33. How do users access specific webpages or resources on the World Wide Web?
- A. By dialing phone numbers
  - B. By entering a Uniform Resource Locator (URL) in a web browser
  - C. By sending physical letters
  - D. By using satellite communication systems
34. Which component of a URL specifies the protocol used to access the resource?
- A. Host
  - B. Scheme
  - C. Path
  - D. Query
35. The World Wide Web relies on the underlying infrastructure of:
- A. Operating systems
  - B. Web browsers
  - C. The internet
  - D. Search engines
36. What does URL stand for?
- A. Universal Routing Locator
  - B. Uniform Resource Locator
  - C. Unique Resource Locator



- D. United Resource Locator
37. Which component of a URL represents the specific location or file on the server's file system?
- A. Scheme
  - B. Host
  - C. Path
  - D. Query
38. URLs are used to identify and locate resources on the:
- A. Intranet
  - B. Internet
  - C. Extranet
  - D. Mainframe computers
39. URLs are designed to be human-readable, allowing users to recognize and understand the:
- A. Computer's IP address
  - B. Data encryption algorithm
  - C. Resource they point to
  - D. Browser's version
40. What is the purpose of URL encoding?
- A. To make URLs case-insensitive
  - B. To allow URLs to be interpreted correctly by web browsers and servers
  - C. To convert URLs in to binary code for faster processing
  - D. To make URLs visually appealing
41. What is the primary function of a search engine?
- A. To create web pages
  - B. To analyze and index web pages
  - C. To provide social media integration
  - D. To facilitate e-commerce transactions
42. Which search engine is the most widely used and known for its comprehensive index and accurate search results?
- A. Yahoo
  - B. Bing
  - C. Google
  - D. Ask.com

43. Search engines use complex algorithms to analyze and index web pages, enabling users to find relevant content based on their:
- A. GPS coordinates
  - B. Social media profiles
  - C. Search queries
  - D. Weather forecasts
44. How do search engines deliver search results to users?
- A. By sending emails
  - B. By displaying advertisements only
  - C. By utilizing advanced search operators
  - D. By using algorithms to rank and display relevant webpages
45. Which search engine provides a question-and-answer-based search experience, allowing users to enter questions in natural language?
- A. Google
  - B. Bing
  - C. Yahoo
  - D. Ask.com
46. Which component of a computer is responsible for executing instructions and performing calculations?
- A. Memory
  - B. Storage Devices
  - C. Input Devices
  - D. Central Processing Unit (CPU)
47. What type of memory is volatile and is used for temporarily holding data and instructions during program execution?
- A. Read-Only Memory (ROM)
  - B. Cache Memory
  - C. Random Access Memory (RAM)
  - D. Solid-State Drives (SSDs)
48. Which of the following is the primary function of an operating system (OS)?
- A. Creating web pages
  - B. Managing computer hardware resources
  - C. Designing graphical elements for a website
  - D. Providing internet connectivity

49. What is the function of a searching engine in the context to the World Wide Web?
- Creating web pages and websites
  - Hosting and managing websites
  - Allowing users to search and retrieve information from the internet
  - Providing internet connectivity to users
50. Which of the following is a characteristic of URLs?
- They are case-insensitive.
  - They cannot contain special characters.
  - They are used for creating multimedia content.
  - They do not contain domain names.

### ANSWERS

1-B,2-C,3-B,4-B,5-D,6-C,7-B,8-B,9-B,10-A,11-C,12-B,13-C,14-C,15-C,16-B,17-B,18-C,19-B,20-B,21-D,22-D,23-D,24-D,25-D,26-B,27-C,28-C,29-B,30-C,31-B,32-B,33-B,34-B,35-C,36-B,37-C,38-B,39-C,40-B,41-B,42-C,43-C,44-D,45-D,46-D,47-C,48-B,49-A,50-A

### Unit – VIII

#### Bioinformatics

- Databases: nucleic acids genome, protein sequences and structures, Bibliography
- Sequence analysis (homology): Sequence alignments – BLAST, CLUSTALW
- Phylogenetic analysis

#### Database:

A database is used to store and organize the data logically. It is a computerized archive that contains information which can be retrieved easily

A **biological database** is a collection of data that is organized so that its contents can easily be accessed, managed, and updated.

Biological databases can be broadly classified into **sequence and structure** databases. Sequence databases are applicable to both **nucleic acid sequences and protein sequences**, whereas **structure database** is applicable to only **Proteins**.

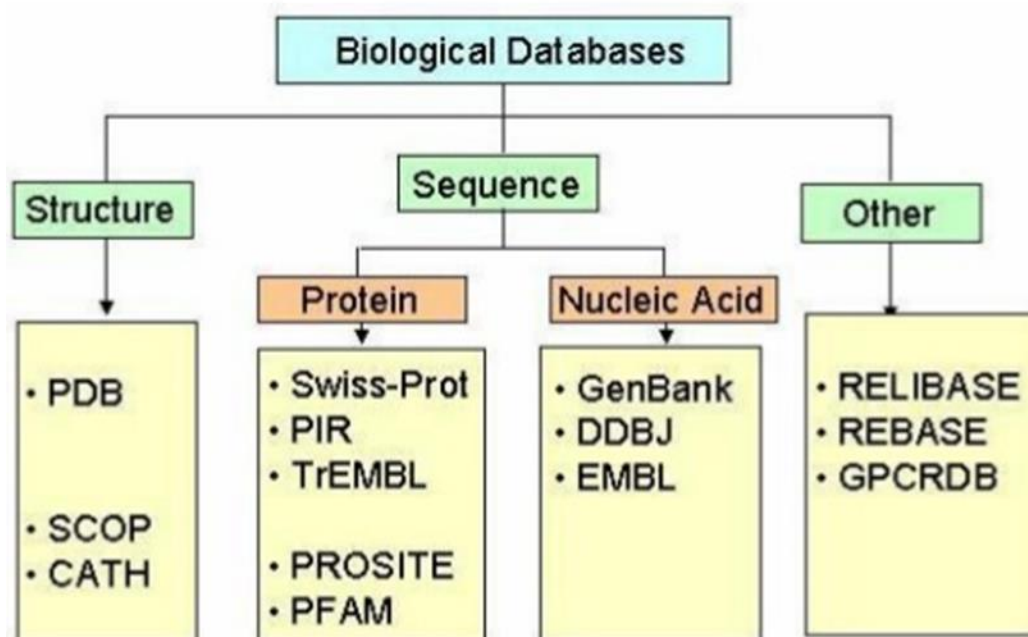
The first database was created within a short period after the Insulin protein sequence was made available in 1956. During this period, three dimensional structures of proteins were studied and the well-known **Protein Data Bank** was developed as the first protein structure database with only 10 entries in 1972. This has now grown in to a large database with over 10,000 entries.

Databases in general can be classified into **primary**, **secondary** and **composite** databases.

A **primary** database contains information of the sequence or structure alone. Examples of these include **Swiss-Prot & PIR** for protein sequences, **GenBank & DDBJ** for Genome sequences and the **PDB** (Protein Data bank) for protein structures.

A **secondary** database contains derived information from the primary database. A secondary sequence database contains information like the conserved sequence, signature sequence and active site residues of the protein families arrived by multiple sequence alignment of a set of related proteins. A secondary structure database contains entries of the PDB in an organized way. Some of the secondary database created and hosted by various researchers at their individual laboratories includes **SCOP**, developed at Cambridge University; **CATH** developed at University College of London, **PROSITE** of Swiss Institute of Bioinformatics, eMOTIF at Stanford.

**Composite** database amalgamates a variety of different primary database sources, which obviates the need to search multiple resources.



### BLAST:

Basic Local Alignment Search Tool (BLAST) is a sequence similarity search program for comparing biological sequences such as amino acids sequences of different proteins or the nucleotides of DNA sequences with sequence database or library sequences. It is an In silico Hybridisation experiment used to identify significant similarities between query sequences with the library sequences. BLAST was designed by Eugene Myers, Samuel Karlin, Stephen

Altschul, Warren Gish, David J. Lipman and Webb Miller (1990, 1994, 1997). It was originally developed and controlled by NCBI

**Input** – FASTA FORMAT

**Output**- HTML, XML, PLAIN TEXT FORMAT

### **CLUSTAL:**

Clustal is a series of widely used computer programs used in Bioinformatics for multiple sequence alignment. There have been many versions of Clustal over the development of the algorithm. Clustal Omega has the widest variety of operating systems out of all the Clustal tools.

There have been many variations of the Clustal software. The original software for global multiple sequence alignments. It is created by Des Higgins in 1988. It was based on deriving phylogenetic trees from pairwise sequences of amino acids or nucleotides.

### **ClustalW:**

The third generation, released in 1994. It greatly improved upon the previous versions. It improved upon the progressive alignment algorithm in various ways, including allowing individual sequences to be weighted down or up according to similarity or divergence respectively in a partial alignment. It also included the ability to run the program in batch mode from the command line.

### **Phylogenetic tree**

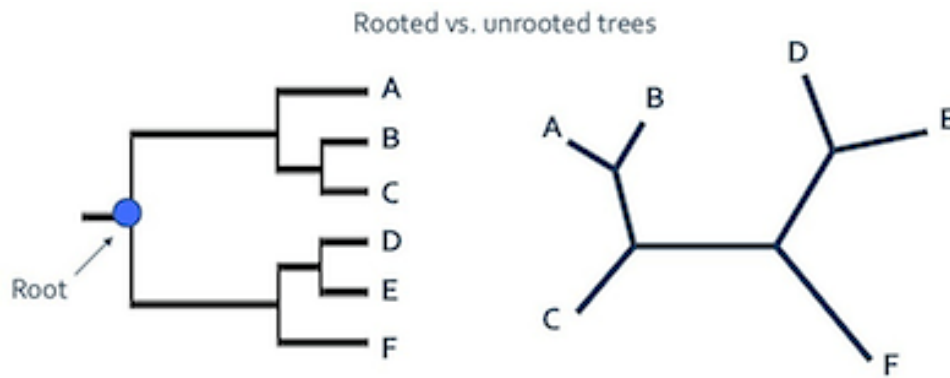
A process for showing relatedness or genetic relatedness among species. It shows the evolutionary relationship among the species. We construct a diagram which gives the ideas of evolutionary relationship. We can predict which organisms are more or less closely related to each other. Two types of phylogenetic trees

1. Rooted- emerging species is known
2. Unrooted – emerging species is not known

To construct a rooted phylogenetic tree we need to know physiological, morphological and genetic information about the species

For unrooted trees we don't know the ancestor so we have less idea about the species. We can not construct an evolutionary diagram properly. It's a radial like structure without any kind of rooting





### Questions

1. Bioinformatics is an interdisciplinary field that combines:
  - A. Biology and chemistry
  - B. Biology and computer science
  - C. Physics and mathematics
  - D. Computer science and physics
2. What is the primary role of bioinformatics in the field of genomics?
  - A. Analyzing protein structures
  - B. Predicting drug interactions
  - C. Studying genetic variations
  - D. Investigating metabolic pathways
3. What is the purpose of genome annotation in bioinformatics?
  - A. Studying evolutionary relationships
  - B. Identifying genes and regulatory elements
  - C. Predicting protein structures
  - D. Analyzing gene expression data
4. Comparative genomics help sin:
  - A. Designing drugs
  - B. Identifying microbial communities
  - C. Studying evolutionary relationships
  - D. Predicting protein functions
5. The three-dimensional structures of proteins can be predicted using:
  - A. Genome annotation
  - B. Comparative genomics
  - C. Structural biology



- D. Bioinformatics methods
6. Functional genomics helps in understanding the roles of genes in:
- A. Protein-protein interactions
  - B. Drug responses
  - C. Biological processes
  - D. Structural biology
7. Meta genomics involves the study of genetic material obtained from:
- A. Environmental samples
  - B. Human tissues
  - C. Bacterial cultures
  - D. Animal models
8. Pharmacogenomics integrates genomic data with information about:
- A. Protein structures
  - B. Drug responses
  - C. Evolutionary relationships
  - D. Metabolic pathways
9. A data base is a structured collection of data organized into:
- A. Rows and columns
  - B. Books and chapters
  - C. Sentences and paragraphs
  - D. Genes and proteins
10. Which of the following is a key feature of data bases that ensures data accuracy?
- A. Data integration
  - B. Data retrieval
  - C. Data security
  - D. Data scalability
11. Which language is commonly used to interact with databases?
- A. Python
  - B. Java
  - C. C++
  - D. SQL
12. Data bases enforce data integrity through:
- A. Data visualization
  - B. Data relationships

- C. Data modification  
D. Data constraints
13. What type of relationship allows multiple records in one table to be associated with multiple records in another table?  
A. One-to-one  
B. One-to-many  
C. Many-to-one  
D. Many-to-many
14. Biological data bases serve as comprehensive repositories of:  
A. Nucleotide sequences  
B. Protein structures  
C. Gene expression data  
D. All of the above
15. Which database is one of the largest and most comprehensive nucleic acid databases?  
A. GenBank  
B. ENA  
C. DDBJ  
D. RefSeq
16. What is the primary purpose of reference genomes in nucleic acid databases?  
A. Data retrieval and analysis  
B. Standardizing genetic research  
C. Storing protein sequences  
D. Supporting metagenomics studies
17. The UCSC Genome Browser is known for:  
A. Data integration  
B. Protein structure prediction  
C. Comparative genomics analysis  
D. Drug discovery
18. Nucleic acid data bases are essential for researchers in the field of:  
A. Bioinformatics  
B. Structural biology  
C. Pharmacology  
D. Ecology

19. The European Nucleotide Archive (ENA) collaborates with which other data bases as part of the International Nucleotide Sequence Database Collaboration (INSDC)?
- GenBank and DDBJ
  - GenBank and RefSeq
  - RefSeq and DDBJ
  - GenBank, ENA, and DDBJ
20. Which area of biology does bioinformatics NOT play a crucial role in?
- Genomics
  - Proteomics
  - Ecology
  - Transcriptomics
21. What is the primary role of genome data bases in genomics research?
- Storing and providing access to protein sequences
  - Storing and providing access to complete or partial genomic sequences of organisms
  - Integrating data from diverse sources for functional analysis
  - Classifying proteins based on their structural features
22. Which of the following data bases provides annotated and high-quality reference sequences for proteins and nucleotides?
- UniProt
  - GenBank
  - Ensembl
  - RefSeq
23. Which database focuses on clustering similar protein sequence stored to reduce redundancy?
- PDBsum
  - InterPro
  - UniRef
  - Pfam
24. What is the purpose of a bibliography in academic and research work?
- To provide a list of related websites for further exploration
  - To give credit to the authors and sources of information used in a document
  - To classify proteins based on their structural features
  - To arrange the sequences in a specific order for readability

25. Which component of the FASTA format begins with a greater-than symbol(">") followed by a unique identifier for the sequence?
- Sequence identifier line
  - Sequence data
  - Citation
  - Order
26. The FASTA format is commonly used in which of the following applications?
- Storing nucleotide sequences in genome databases
  - Structural classification of proteins
  - Creating phylogenetic trees
  - Formatting bibliographies
27. Which database is a hierarchical protein structure classification database that groups proteins into families, superfamilies, folds, and classes?
- CATH
  - UniProt
  - GenBank
  - SCOP
28. What type of bibliography includes a brief summary or evaluation (annotation) of each cited source?
- Works Cited
  - References
  - Annotated Bibliography
  - Annotated Webography
29. In which database can researchers find comprehensive genome annotations, gene annotations, functional annotations, genetic variations, and other genomic information for a wide range of organisms?
- UniProt
  - FlyBase
  - UCSC Genome Browser
  - Ensembl
30. What does the FASTA format primarily represent?
- Three-dimensional protein structures
  - Nucleotide or protein sequences
  - Protein families and domains

- D. Genome annotations and gene predictions
31. Which database serves as a central hub for protein-related information, integrating data from various sources like Swiss-Prot and TrEMBL?
- A. SCOP
  - B. CATH
  - C. UniProt
  - D. Pfam
32. The FASTA format is commonly used for which of the following applications?
- A. Storing and providing access to complete or partial genomic sequences
  - B. Protein structure prediction
  - C. Genome annotation and gene prediction
  - D. Creating phylogenetic trees
33. How sources in a bibliography are typically arranged?
- A. Based on the length of the content
  - B. In the order of citation in the paper
  - C. In alphabetical order based on authors' last names
  - D. In chronological order of publication
34. Which protein structure data base provides interactive visualization and analysis tools for exploring protein structures and interactions?
- A. CATH
  - B. SCOP
  - C. UniProt
  - D. Protein Data Bank (PDB)
35. Which type of bibliography focuses specifically on web-based sources, such as websites and online articles?
- A. Annotated Bibliography
  - B. Annotated Webography
  - C. Works Cited
  - D. References
36. Which data base is widely used for comparing query sequences against a database of known sequences to identify sequence similarities and functional annotations?
- A. UniProt
  - B. GenBank
  - C. Ensembl

- D. BLAST
37. The Genome Data Viewer is a part of which data base and provides interactive visualization and analysis tools for exploring genome data?
- A. GenBank
  - B. RefSeq
  - C. UniProt
  - D. PDB
38. The CATH database categorizes proteins into hierarchical levels based on which of the following?
- A. Sequence similarities
  - B. Structural features and evolutionary relationships
  - C. Functional annotations
  - D. Gene expression levels
39. Which database is primarily focused on storing and providing access to complete or draft genome sequences of various organisms?
- A. Protein Data Bank (PDB)
  - B. UniProt
  - C. GenBank
  - D. CATH
40. What is the purpose of including annotations in an annotated bibliography?
- A. To acknowledge the authors of the sources
  - B. To provide a brief summary or evaluation of each cited source
  - C. To list the sources alphabetically
  - D. To classify proteins based on their structural features
41. Sequence analysis is a fundamental approach in bioinformatics that involves the examination, comparison, and interpretation of which biological sequences?
- A. Amino acid sequences
  - B. DNA, RNA, and protein sequences
  - C. DNA sequences only
  - D. RNA sequences only
42. What is the primary purpose of functional annotation in sequence analysis?
- A. Identifying homologous sequences
  - B. Predicting protein structures
  - C. Inferring evolutionary relationships



- D. Annotating the functions of genes, proteins, and non-coding RNAs
43. Which method of sequence analysis allows the comparison of sequences from different species to gain insights into evolutionary relationships?
- A. Motif discovery
  - B. Hidden Markov Models (HMMs)
  - C. Comparative genomics
  - D. Global alignment
44. Sequence analysis is essential in understanding the genetic basis of diseases. How does it achieve this?
- A. By predicting protein structures
  - B. By identifying orthologs and paralogs
  - C. By analyzing DNA or protein sequences to identify genetic variations and mutations
  - D. By constructing phylogenetic trees
45. Which technique in sequence analysis predicts the three-dimensional structure of proteins based on their amino acid sequences?
- A. Comparative genomics
  - B. Phylogenetic analysis
  - C. Pairwise sequence alignment
  - D. Homology modeling
46. Multiple sequence alignment (MSA) allows the comparison of how many biological sequences simultaneously?
- A. One sequence
  - B. Two sequences
  - C. Three or more sequences
  - D. Four sequences
47. Which alignment algorithm starts with pairwise alignments and it adds more sequences to construct the final multiple alignment?
- A. Progressive alignment
  - B. Hidden Markov Models (HMMs)
  - C. Iterative refinement
  - D. Dynamic programming
48. BLAST is an algorithm used for which of the following tasks?
- A. Multiple sequence alignment

- B. Protein structure prediction
  - C. Functional annotation
  - D. Sequence similarity searches
49. What is the statistical measure used in BLAST to estimate the significance of sequence alignments?
- A. Sum-of-pairs score
  - B. Hidden Markov Models (HMMs)
  - C. Maximum likelihood
  - D. E-value
50. What is the primary purpose of pairwise sequence alignment?
- A. Identifying homologous sequences
  - B. Inferring evolutionary relationships
  - C. Aligning three or more sequences
  - D. Comparing sequences from different species
51. Which tool is used for motif discovery in a set of sequences to identify functional elements?
- A. BLAST
  - B. Needleman-Wunsch algorithm
  - C. MEME
  - D. Smith-Waterman algorithm

**ANSWERS**

1-B,2-C,3-B,4-C,5-D,6-C,7-A,8-B,9-A,10-C,11-D,12-D,13-D,14-D,15-A,16-B,17-C,18-A,19-D,20-C,21-B,22-D,23-C,24-B,25-A,26-A,27-A,28-C,29-D,30-B,31-C,32-A,33-C,34-D,35-B,36-D,37-A,38-B,39-C,40-B,41-B,42-D,43-C,44-C,45-D,46-C,47-A,48-D,49-D, 50-A, 51-C