4. (a) Discuss various similarities and dissimilarities between correlation and regression. (4)

(b) Calculate standard deviation & standard error from the given dataset : (6)

Marks obtained	18	29	16	12	26	33	25	17	10	11

[This question paper contains 8 printed pages.]

02.01.24(M)

Your Roll No.....

Sr. No. of Question Paper : 1665 **G**

Unique Paper Code : 2163012002

Name of the Paper : Biostatistics and Bioinformatics for Plant Sciences

Name of the Course : **B.Sc. (Hons.) Botany (DSE)**

Semester : III

Duration : 2 Hours Maximum Marks : 60

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **Section A** and **Section B** on separate sheets.
3. Question no. **1** of both the sections is compulsory.
4. Attempt any **three** questions from **Part A** and **three** questions from **Part B** including Q. No. **1** of both the sections.
5. Attempt all parts of a question together.

P.T.O.

Section A

1. (a) Define the following **(any five)** : (5×1=5)

(i) Metabolomics

(ii) Composite database

(iii) Paralogous sequence

(iv) Bioinformatics

(v) Alignment

(vi) Monophyletic clade

(b) Give an example of each of the following **(any five)** : (5×1=5)

(i) Nucleotide sequence database

(vi) _____ is a measure of the tailedness of a distribution.

2. Differentiate between the following **(any two)** : (5×2=10)

(a) Student's t test and chi-square test

(b) Karl Pearson method and Spearman Rank method

(c) Descriptive and inferential statistics

3. Write short note on **(any two)** : (5×2=10)

(a) Limitations and applications of biostatistics

(b) Presentation of data

(c) Quartile deviation–merits and demerits

(ii) In _____ sampling, the population is divided into subgroups, and then samples are randomly selected from each subgroup.

(iii) In the process of data analysis, _____ involves organizing data into tables to make it understandable and informative for the audience.

(iv) The _____ is the simplest measure of dispersion and is calculated as the difference between the highest and lowest values in a dataset.

(v) _____ correlation is a relationship between two variables that move in opposite directions.

(ii) Chemical database

(iii) Literature database

(iv) Alignment tool

(v) Protein database

(vi) File format

2. Differentiate between the following **(any four)** :

(2.5×4=10)

(a) Primary and Secondary database

(b) Genomics and Proteomics

(c) NCBI and PDB

(d) Pairwise sequence and Multiple sequence alignment

(e) Maximum likelihood and Maximum Parsimony method

3. Write short notes on **(any two)** : (5×2=10)

(a) PlantPepDB

(b) Applications of bioinformatics in drug discovery

(c) Concepts of gaps and penalty in alignment

4. (a) Elaborate on various types of standard BLAST. (5)

(b) Draw and label the various parts of a phytogenetic tree. What do each of these parts signify. Comment. (5)

Section B

1. (a) Define the following **(any five)** : (5×1=5)

(i) Inferential statistics

(ii) Skewness

(iii) Alternate hypothesis

(iv) Secondary data

(v) Dispersion

(vi) Mean deviation

(b) Fill in the blanks **(any five)** : (5×1=5)

(i) The formula for coefficient of variation is _____ .