

Ajmi

Ajmi

Sanjana Singh

परीक्षार्थी का पूरा नाम

कक्ष निरीक्षक का नाम

हस्ताक्षर

परीक्षार्थी द्वारा सम्पूर्ण विवरण भर लिए गये है।



R

2018-

भाग-2

M.Sc. Internal

चौधरी चरण सिंह विश्वविद्यालय, मेरठ
Ch. Charan Singh University, Meerut

निम्नलिखित विवरण परीक्षार्थी द्वारा स्वयं भरा जाए (To be filled by the Examinee)

परीक्षा का नाम M.Sc (Name of Exam) वर्ष 20 19 (Year 20.....) भाग/सेमेस्टर II (Part / Semester)
विषय Zoology (Subject) प्रश्न-पत्र/पाठ्यक्रम Biostatistics & Bioinformatics (Paper /Course) पेपर कोड नं. H-2062 (Paper Code No.)
परीक्षा का दिन Wednesday (Day of Examination) दिनांक 1/5/19 (Date)

प्राप्तांक एवं पूर्णांक परीक्षकों द्वारा भरे जायें

पूर्णांक (Max. Marks)

प्रश्नों की क्रम संख्या	a/I	b/II	c/III	d/IV	e/V	f/VI	g/VII	h/VIII	i/IX	j/X	योग
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											

प्राप्तांक

(शब्दों में)	अंकों में
--------------	-----------

जाँचकर्ता के हस्ताक्षर एवं तिथि

परीक्षक के हस्ताक्षर एवं तिथि



चौधरी चरण सिंह विश्वविद्यालय, मेरठ

R

Date Stamp to be affixed here

मार्गदर्शक

(परीक्षार्थी द्वारा भरा जाए)

परीक्षा का नाम M.Sc (Name of Exam) विषय Zoology (Subject) प्रश्न पत्र Biostatistics & Bioinformatics (Paper /Course) दिनांक 1/5/19 (Date)

परीक्षार्थी का अनुक्रमांक (Roll Number)

उत्तर-पुस्तिका क्रमांक

18	P	0	1	7	9	4	6	1				
M	A	0	0	0	0	0	0	0	0	0	0	0
B	1	1	1	1	1	1	1	1	1	1	1	1
C	2	2	2	2	2	2	2	2	2	2	2	2
D	3	3	3	3	3	3	3	3	3	3	3	3
E	4	4	4	4	4	4	4	4	4	4	4	4
F	5	5	5	5	5	5	5	5	5	5	5	5
G	6	6	6	6	6	6	6	6	6	6	6	6
H	7	7	7	7	7	7	7	7	7	7	7	7
I	8	8	8	8	8	8	8	8	8	8	8	8
J	9	9	9	9	9	9	9	9	9	9	9	9
K												
L												
P												
S												
T												
U												
V												
W												

KM-I-01-

कालेज कोड

0	1	8					
0	0	0	0				
1	1	1	1				
2	2	2	2				
3	3	3	3				
4	4	4	4				
5	5	5	5				
6	6	6	6				
7	7	7	7				
8	8	8	8				
9	9	9	9				

(परीक्षार्थी की श्रेणी)

- संस्थागत
- व्यक्तिगत
- बैक पेपर
- अंक सुधार
- भूतपूर्व
- एकल विषय

नामांकन संख्या (Enrollment Number)

पेपर कोड

M	1	5	5	3	9	6	1	4				
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7
8	8	8	8	8	8	8	8	8	8	8	8	8
9	9	9	9	9	9	9	9	9	9	9	9	9

H	2	0	6	2
---	---	---	---	---

परीक्षार्थी का पूरा नाम

Sanjana Singh

कक्ष निरीक्षक का नाम

Ajmi

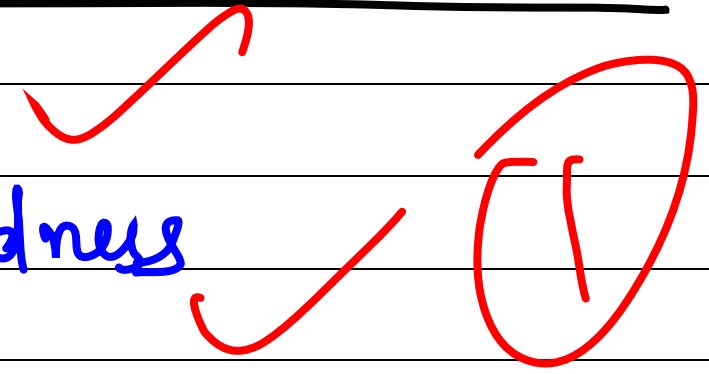
Section - A

Ans 1 → Hypothesis -

- Hypothesis is used in the chi-square test.
- It is used to ^{infer} accept or reject the null hypothesis.
- When tabulated value is larger than we accept the null hypothesis.
- When T.V is smaller than chi-square value then null hypothesis is rejected.

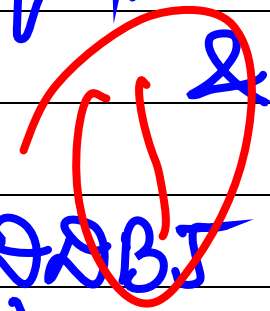
Ans 2 → Two Genetic disorders —

- Haemophilia
- Colour Blindness



Ans 3 → Gene Bank —

- Gene Bank are used to preserve the protein sequence, nucleotide sequence, DNA & RNA.
- In this the DDBJ (DNA data Bank of Japan), EMBL at EBI (European Bioinformatics Institute) are



work with collaboration.

- We use the sequence database at the daily basis.

Ans 4 → full form of http -

http → Hyper Text Transfer Protocol

Ans 5 → DDBJ -

- DDJI stands for DNA Data Bank of Japan.

- DDBJ used for the collection of DNA sequences.
- It is located at the National Institute of Genetics (NIG) at the Shizuoka Prefecture of Japan.
- It is also a member of International Nucleotide Sequence Database Collaboration (INSDC).

Section - B

Ans 6 → Characters

Red - Tall

Red - Dwarf

White - Tall

White - Dwarf

No. of Plants

850

250

350

~~150~~

1600

$$= \frac{1600}{16}$$

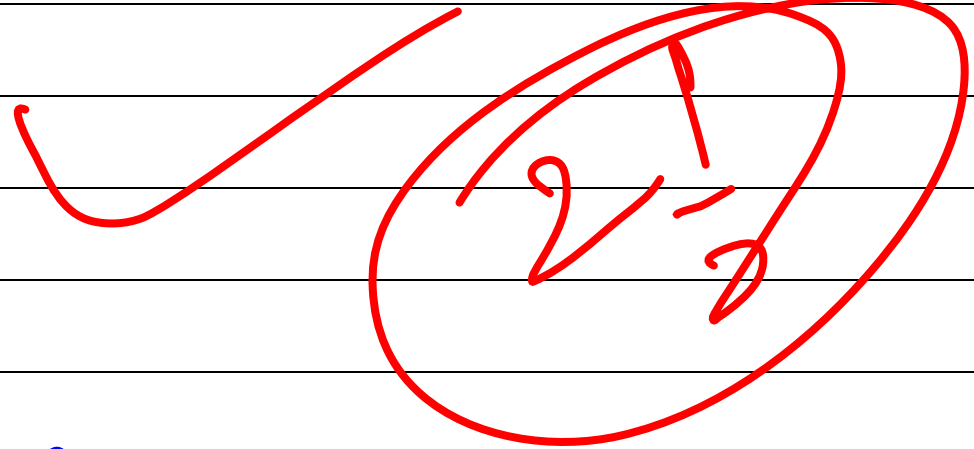
$$= 100$$

Characters	No. of plants O	Expected v. E	O-E	(O-E) ²	$\frac{(O-E)^2}{E}$
Red-Tall	850	100 × 9 = 900	-50	2500	2.77
Red-Dwarf	250	100 × 3 = 300	-50	2500	8.33
White-Tall	350	100 × 3 = 300	50	2500	8.33
White-dwarf	150	100 × 1 = 100	50	2500	25

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

$$\chi^2 = 44.43$$

$$\begin{aligned} df &= n-1 \\ &= 4-1 \\ &= 3 \end{aligned}$$



T.V. > C.V. (Null hypothesis accepted)

T.V. < C.V. (Null hypothesis rejected)

Ans 8 → SWISS - PORT

Summary

- Introduction
- It consist of two sections
 - SWISS - PORT
 - Tr - EMBL
- Features of SWISS - PORT
 - Annotation Data
 - Core Data
 - Annotation
 - Minimal of redundancy
 - Integration of database
 - Documentation

→ Tr-EMBL : A computer annotated supplement

→ Why TrEMBL

→ TrEMBL

Sp-TrEMBL

REM-TrEMBL

→ Conclusion

• Introduction -

It is the Universal protein Knowledgebase (UniProt KB) is the central hub for the collection of

information on protein.

• It consist of two sections

SWISS PORT

TrEMBL

(Translated European
Molecular Biological
Laboratory)

→ Reviewed
→ Manually
annotated

→ Unreviewed
→ Computationally
annotated

→ Records with information

→ Records with the

collected from curator
& lecturer to
computationally annotated.

autolift of manually
annotated.

- Introduction -

- SWISS-PROT strives to provide high level of annotation, minimal level of redundancy & integration with other databases.

- Tr-EMBL a computational annotated supplement.

- It is helpful for protein sequences,

nucleotide sequences & other databases.

- Features of SWISS - PROT -

→ Annotation Data

```
graph TD; A[Annotation Data] --> B[Core Data]; A --> C[Annotation];
```

- Core Data -

→ The citation information.

→ Taxonomic characters

- Annotation -

→ Domains & sites
e.g. Calcium binding regions, Zinc fingers.

→ Quaternary Structure
e.g. Homodimer, Heterodimer

→ Secondary Structure
e.g. Alpha helix, Beta Sheet etc.

→ Minimal level of redundancy —

- The flow of the change is low
- The merged data is coherent & cohesive.
- Data is collected from creator & literature.

→ Integration of database —

- Database is integrated with other

nucleotide sequence database , protein
sequence database etc.

→ Documentation —

- Documentation are kept up-to-date.
- Documents are checked on daily basis.
- TeX-EMBL : A computer annotated supplement

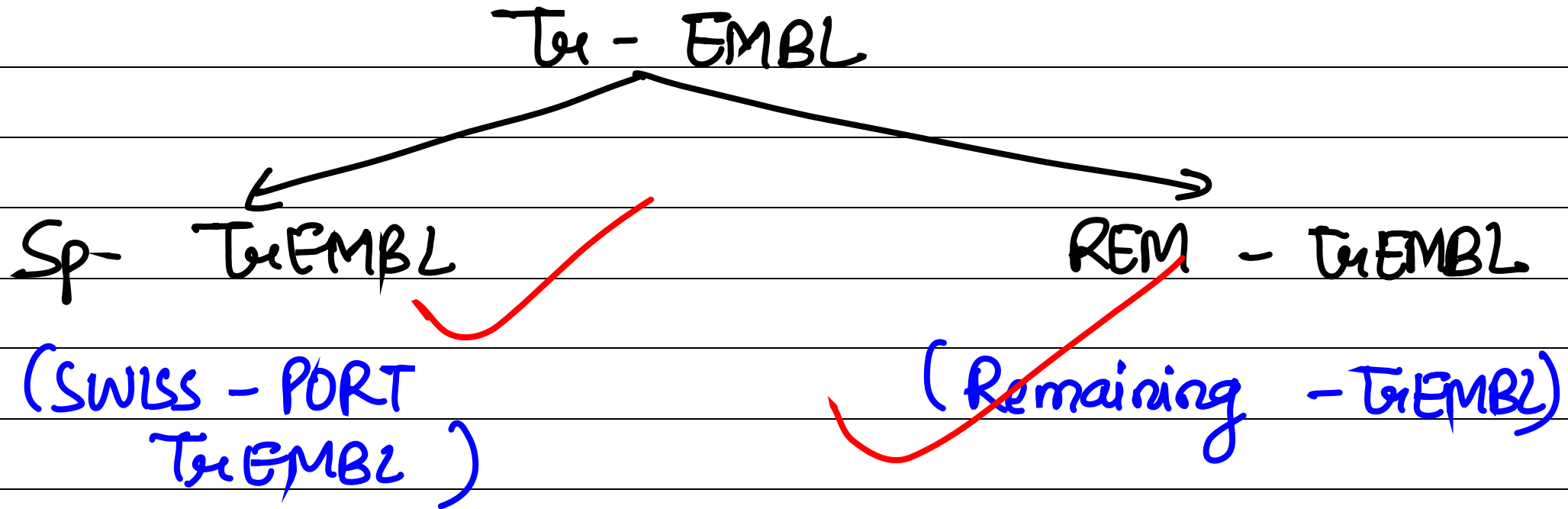
Tr-EMBL (Translated European Molecular Biology Laboratory) since 1996.

• Why Tr-EMBL?

→ The flow of mechanism of genome projects are present in this.

→ It is computer annotated supplement.

→ Much more of the literature gave more information.



SWISS - PROT $T_{\alpha} \text{EMBL}$ -

- It contain ~~sequence~~ ~~to~~ which is incorporated to Swiss port.

REM - TREMBL -

- It contains those sequences which are not incorporated to SWISS-Prot.

Conclusion -

- It is present in bioinformatics & in this time the bioinformatics are present in colleges & universities to express the protein & nucleotide sequences.

Section - C

<u>Ans</u>	Marks	No. of students
------------	-------	-----------------

	0-10	10
--	------	----

	10-20	5
--	-------	---

	20-30	5
--	-------	---

	30-40	8
--	-------	---

	40-50	12
--	-------	----

	50-60	10
--	-------	----

	60-70	21
--	-------	----

	70-80	11
--	-------	----

Total
Students

10	
5	
5	
8	
12	
10	
21	
11	
100	

20

?

.

$$\text{Mean} = \frac{\sum X}{N}$$

$$= \frac{100}{8}$$

$$\text{Mean} = 12.5$$

$$\text{Median} = \frac{L + \left(\frac{n}{2} - C\right) i}{f}$$

$$= \frac{40 + \left(\frac{100}{2} - 8\right) 10}{12}$$

$$= \frac{40 + (50 - 8) 10}{12}$$

$$= \frac{40 + (42) 10}{12}$$

$$= 40 + \frac{420}{12}$$

$$= 40 + 35$$

Median = 75

$$\text{Mode} = L + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times i$$

$$= 60 + \frac{21 - 18}{2 \times 21 - 18 - 11} \times 10$$

$$= 60 + \frac{3}{42 - 29} \times 10$$

$$= 60 + \frac{3}{13} \times 10$$

$$= 60 + \frac{30}{13}$$

0.23
 $13 \overline{) 30}$
 $\underline{26}$
 40
 $\underline{39}$
 1

~~$60 + 0.23 \times 10$~~

$$= 60 + 0.23$$

$$\boxed{\text{Mode} = 60.23}$$

$$\text{Mode} = 3 \text{ Median} - 2 \text{ Mean}$$

$$3 \text{ Median} = \text{Mode} + 2 \text{ Mean}$$

$$= 60.23 + 2 \times 12.5$$

$$= 60.23 + 25$$

$$= 85.23$$

$$\text{Median} = 85.23 \times 3$$

$$= 255.69$$

3

Ans 10) Sequence Database

- Sequence database is a biological database that collect the protein sequence database, nucleotide sequence database.
- The uniprot is an example of protein sequence database.

Sequence Database is of three types —

Gene Bank —

- Gene Bank is used to preserved the sequence database of protein & nucleotide.
- It work in collaboration with DDBJ & EMBL at EBI.
- These three ~~are~~ work together & maintain the stability of bioinformatics.

AABI —

- AABI stands for DNA Data Bank of Japan.
- AABI is a biological sequences that collects the DNA sequences.
- AABI is located at NIG (National Institute of Genetics) at Shizuoka Prefecture of Japan.
- It works with INSDC (International Nucleotide Sequence Database Collaboration).

EMBL at EBI

- EMBL (European Molecular Biological Laboratory) at EBI (European Bioinformatics Institute).
- EMBL is a computational sequence which preserved the data or sequence database in computer, floppy disk, harddrive, etc.
- They are work in collaboration.

Conclusion —

- In bioinformatics the sequence database is helpful to identify the nature & process of the protein & nucleotide sequences.

