BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)
FIRST SEMESTER 2023-2024

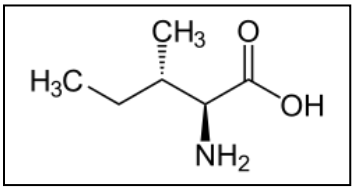
BIOSTATISTICS AND BIOMODELLING (BIO G532)

COMPREHENSIVE EXAMINATION
TOTAL WEIGHTAGE 35% Date: 09.12.2023 DURATION: 3Hrs. (Part A & Part B)
Total Marks (35+35) =70

- **Answer Part A and Part B in separate answer sheets.**
- **Irrelevant answers may attract penalty.**
- **Please answer all parts of the questions in continuation**

PART – A (CLOSED BOOK) (Max. Duration: 1.5 Hrs., Max. Marks 35)

1.a) Side chain conformation during model generation is frequently taken from the rotamer library. How many possible rotamers exist for the following amino acid? Explain with the proper diagram. **[3]**



b) Explain why α-helix have particular significance in DNA binding motifs, including helix-turn-helix motifs, leucine zipper motifs and zinc finger motifs. **[2]**

2.a) Draw and label the standard coordinate frame of DNA double helix. **[2]**

b) Explain slide (step parameter, Y axis translation) and propeller twist (basepair parameter, Y axis rotation) with appropriate diagram. **[3]**

c) If you observe the following characteristic features of a piece of 5 Kb DNA strip, what would be your conclusion about the form of DNA? What will be the end to end length of that DNA strip? **[2]**

[Features: right-handed, ~11 bp per turn, 2.6 nm pitch, most common form for RNA]

3. a) Draw a flowchart depicting the steps of comparative modeling. **[3]**

b) Discuss the possible reasons for errors during comparative modeling. **[3]**

4. State True/ False with Justification: **[4.5]**

a) In an ANOVA, if the sum of squares for error is 400, the sum of squares for treatment is 180, and the total sample size for the four groups compared is 88, then the null hypothesis should not be rejected.

b) A highly peaked frequency distribution curve is known as platykurtic.

c) If the sample size is fixed, if you increase the type I error it will reduce the probability of a Type II error.

5.

a) Suppose the correlation coefficient between height (as measured in feet) versus weight (as measured in pounds) is 0.40. What is the correlation coefficient of height measured in inches versus weight measured in ounces? [12 inches = one foot; 16 ounces = one pound]. In case it can't be determined with the given information please state so with justification. [**2]**

b) Suppose you conduct a significance test for the population proportion and your p- value is 0.184. At 0.10 level of significance, what you will conclude about your hypothesis? **[1.5]**

6. The side effects of a new drug are being tested against a placebo. A simple random sample of 565 patients yields the results below. At a significance level of $\alpha = 0.05$, is there enough evidence to conclude that the treatment is independent of the side effect of nausea? [**3]**

	Drug	Placebo
Nausea	36	13
No nausea	254	262

7. A researcher wishes to determine if Vitamin E supplements could increase cognitive ability among elderly women. In 1999, researchers recruited a sample of elderly women age 75-80. At time of study, the women were randomized to either take vitamin E or placebo for 6 months. At the end of period, women were given cognition test. Higher scores on scores indicate better cognition. The mean and SD of test score of 81 women who took vitamin E supplements was 27 and 6.9 respectively. The mean and SD of test score of 90 women who took placebo was 24 and 6.2 respectively. Apply suitable statistical test to conclude the significance at α=0.05. [**3]**

8. Identify the statistical test to be used in each case: [2x1.5=3]

a) A study was set up to look at whether there was a difference in the mean arterial blood pressure between two groups of volunteers, after 6 weeks of following one of two treatment programs. One group of volunteers were given an exercise regimen to follow for the 6 weeks and the other group were given the same exercise regimen with the addition of an experimental tablet. Identify the test to be used.

b) Students were given different drug treatments before revising for their exams. Some were given a memory drug, some a placebo drug and some no treatment. The exam scores (%) were recorded. Which test you will administer to find the effect of drug?

****************************Good Luck**

BIRLA INSTITUTE OF TECHNOLOGY AND SCIENCE, PILANI (RAJ.)
FIRST SEMESTER 2023-2024
BIOSTATISTICS AND BIOMODELLING (BIO G532)
COMPREHENSIVE EXAMINATION
TOTAL WEIGHTAGE 35% Date: 08.12.2023 DURATION: 3Hrs. (Part A & Part B)
Total Marks (35+35) =70

PART – B (OPEN BOOK) (Max. duration: 1.5 Hr., Max. Marks 35)

1. a) A synthetic polypeptide of 20 amino acids was designed to adopt a β-sheet structure in solution. NMR demonstrated that it did indeed adopt this conformation, forming a 3-strand antiparallel structure. The sequence (from the N-terminal end) of this polypeptide is: "RGWSVQNGKYTNNGKTTEGR". Draw a careful sketch showing how this polypeptide might fold, and identify the amino acid that form the turns. **[3]**

b) Peptide nucleic acid (PNA) is an artificially synthesized oligonucleotide mimics combing parts of DNA and protein. What would be probable structure of PNA. Justify your answer. [Hints: It combines either DNA backbone with protein side chain or protein main chain with DNA bases] [**4**]

c) Explain the procedure for modeling structurally variable regions (SVR) in comparative modeling. **[3]**

2. a) The following structural parameters are from three different forms of DNA double helix. Identify them with proper justification. In each form, values are given for CG/GC pairs and CpG/GpC steps. All symbols are representing standard DNA notation. **[4]**

	β	δ	P	χ	Inclin	Rise	X-disp	Prop	Slide
Form-1	**-146°**	**156°**	**192°**	**-98°**	**-6°**	**3.38Å**	**-0.7Å**	**-12°**	**0.23 Å**
Form-2	**-176°**	**140°**	**156°**	**-161°**	**7°**	**3.7Å**	**-1.2Å**	**2.6°**	**0.5 Å**
Form-3	**-152°**	**82°**	**38°**	**-154°**	**19°**	**2.56Å**	**-5.4Å**	**-11°**	**-1.5 Å**

b) Among all standard base pair and step parameters of DNA double helix, identify the four most closely linked parameters. (pairwise) **[3]**

3. Identify the sampling method: **[3]**

a) Engelbert chooses the elements for his sample by giving particular attention for each sub-population. He sees to it that every computed stratum sample is the same with the other strata and that the respondents are chosen randomly. What sampling design is used?

b) A college principal conducted an ethnographic probe into the problems faced by tribal students. Which method of sampling will be most appropriate?

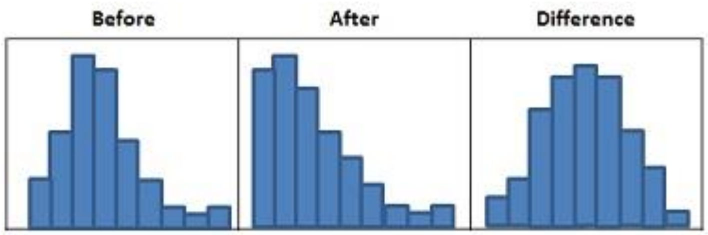
4. A certain experimental drug claims a cure rate of at least 75% for males with prostate cancer. Describe both the Type I and Type II errors in context. Which error is the more serious? **[2]**

5. State whether the data from the given studies can be analyzed using a t-test. If yes, which t-test will be used and if no which other statistical tool should be used? [**6]**

a) A study was conducted to investigate the relationship between sheep live weight (kg) and its chest girth (cm). A random sample of 66 sheep was weighed and simultaneously had their chest girth measured.

b) A histogram of these CD4+ cell counts has shown that the distribution is negatively skewed. We want to test for differences between the average values in Healthy volunteers compared to immunocompromised patients.

c) The following Histograms display the distribution of reported alcohol consumption (units) in patients diagnosed with alcoholic liver disease before an intervention and after the intervention has been completed. A histogram of the difference (before minus after) is also presented. We were interested in testing to see if there had been a significant change in reported alcohol consumption?



5. A research team wants to investigate the usefulness of relaxation training for reducing levels of anxiety in individuals experiencing stress. They identify 30 people at random from a group of 100 who have "high stress" jobs. The 30 people are divided into two groups. One group acts as the control group - they receive no training. The second group (of 15) receives the relaxation training. Now the samples were paired on the dimensions of sex and job type. The subjects in each pair of the group are then given an anxiety inventory. The summarized results appear below where higher scores indicate greater anxiety. Evaluate her experiment using the criteria of $p < .05$. **[3]**

Pairs	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Control:	38	40	35	36	35	32	31	30	28	26	24	21	18	34	22
Relax:	35	32	30	34	30	32	28	27	22	22	18	17	17	25	21

6. A research study was conducted to examine the clinical efficacy of a new antidepressant. Depressed patients were randomly assigned to one of three groups: a placebo group, a group that received a low dose of the drug, and a group that received a moderate dose of the drug. After four weeks of treatment, the patients completed the Beck Depression Inventory. The higher the score, the more depressed the patient. The data are presented below. Compute the appropriate test. **[4]**

Placebo	Low Dose	Moderate Dose
38	22	14
47	19	26
39	8	11
25	23	18
42	31	5

**************************Good Luck***********************************