# MURANG'A UNIVERSITY OF TECHNOLOGY

## SCHOOL OF PURE, APPLIED AND HEALTH SCIENCES

## DEPARTMENT OF MATHEMATICS AND ACTUARIAL SCIENCE

UNIVERSITY ORDINARY EXAMINATION

2023/2024 ACADEMIC YEAR

**THIRD** YEAR **SECOND** SEMESTER EXAMINATION FOR BACHELOR OF SCIENCE IN APPLIED STATISTICS WITH PROGRAMMING

AMS 316 – STATISTICAL GENETICS

DURATION: 2 HOURS

**INSTRUCTIONS TO CANDIDATES:**

1. Answer question one and any other two questions.

2. Mobile phones are not allowed in the examination room.

3. You are not allowed to write on this examination question paper.

**SECTION A: ANSWER ALL QUESTIONS IN THIS SECTION**

**QUESTION ONE (30 MARKS)**

a. Define the following terms
    i.    DNA    (1mark)
    ii.    Nucleotide    (1mark)

b. Explain the main function of the cell.    (3marks)

c. Describe the meiosis process    (3marks)

d. Rabbit's ears can either be short or floppy. Where short ears are dominant over floppy ears. There are 653 individuals in the population. 104 rabbits have floppy ears and 549 have short ears. Find:
    i.    The frequency of the dominant and recessive alleles    (2marks)
    ii.    The frequency of individuals with dominant, heterozygous and recessive genotypes.    (3marks)

e. Suppose we collect data from a random sample or 218 individuals resulting in 68 AA genotypes, 116 Aa genotypes and 34 aa genotypes. Is this locus in Hardy-Weinberg equilibrium for the sampled population?    (5marks)

f. The following genotype and phenotype counts were observed from a backcross experiment

| Genotype | Observed Count |
|---|---|
| $AB/ab$ | 163 |
| $Ab/ab$ | 74 |
| $aB/ab$ | 23 |
| $ab/ab$ | 89 |

    i.    Estimate $r_{AB}$    (2marks)
    ii.    Compute a 99% confidence interval for $r_{AB}$    (2marks)
    iii.    Conduct a test for the hypothesis that the two loci are not linked.    (3marks)

g. A sample of 20 resulted in counts 8,7 and 5 for genotypes AA, Aa and aa respectively. Use Fisher's exact test to check if this locus in in Hardy-Weinberg equilibrium. (5marks)

**SECTION B – ANSWER ANY TWO QUESTIONS IN THIS SECTION**

**QUESTION TWO (20 MARKS)**

a. Discuss the various genetic mutations that can sometimes occur and some common human syndromes that are the result of these DNA structural changes.    (12marks)

b. Describe the backcross experiment    (8marks)

**QUESTION THREE (20 MARKS)**

a. Explain the concept of DNA sequence analysis.    (2marks)

b. Consider a DNA sequence of 50 bases given below:
    CCACTGCGTAGGTTGGCGTGATAGCGGAATGAAGAATTAGGCCAGCCGGA
    i.    Regarding the sequence two bases at a time, perform a test of independence of conservative bases.    (10marks)
    ii.    Formulate the transition probability matrix from one base to another.    (4marks)
    iii.    What is the long-term composition of the base pairs A, C, T and G in this locus?    (4marks)

**QUESTION FOUR (20 MARKS)**

a. Describe micro array experiments (3marks)

b. The data below is for gene expression levels of patients undergoing chemotherapy for leukaemia.

| Young (<50 years old) | 1.8 | 2.2 | 1.5 | 1.3 | 1.7 | 1.9 | 1.1 | 2.1 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Old (≥50 years old) | 2.1 | 1.4 | 2.0 | 1.9 | 1.8 | 2.1 | 2.0 | 1.9 | 1.6 | 2.2 |

Does age affect gene expression levels?

Test at $\alpha = 0.05$ (7marks)

c. A researcher is interested in studying the effectiveness of chemotherapy for different types of leukaemia.
Consider the accompanying data at expression levels and perform an appropriate test at hypothesis at $\alpha = 0.05$

| AML | CML | ALL | CLL |
|---|---|---|---|
| 1.4 | 1.1 | 1.5 | 1.7 |
| 1.2 | 0.9 | 1.3 | 2.5 |
| 1.2 | 1.6 | 1.5 | 1.8 |
| 1.3 | 1.9 | 2.4 | 2.2 |
| 1.1 | 1.5 | 2.1 | |
| | 0.8 | 1.6 | |
| | 1.3 | | |

(10marks)