## MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

**P.O. Box 972-60200 – Meru-Kenya**
**Tel: +254(0) 799 529 958, +254(0) 799 529 959, + 254 (0) 712 524 293,**
**Website: info@must.ac.ke  Email: info@must.ac.ke**

### University Examinations 2024/2025

FIRST YEAR FIRST SEMESTER EXAMINATION FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN FOOD SCIENCE

### AFT 6102: BIOSTATISTICS AND DATA ANALYSIS

**DATE: JANUARY 2025**                                             **TIME: 3 HOURS**

**INSTRUCTIONS:** *Answer question **one** and any other **two** questions*

### QUESTION ONE (30 MARKS)

(a) Explain the meaning of the following terms as used in biostatistics          (3 marks)

    (i)       Response variable

    (ii)     ANOVA

    (iii)    Model

(b) i. Highlight categorical data and give an example.                      (2 marks)

    ii. Describe the parametric methods in data analysis                      (3 marks)

    iii. Given that the variance of random variable $X_i$ is given by Var $(X_i) = \delta^2$. Prove that the standard error (standard deviation of sampling mean $\bar{X}$ is given by s.e. $= \dfrac{\delta}{\sqrt{n}}$

                                    (5 marks)

(c) i. Describe how the normal distribution can be used to approximate the binomial density.

(2 marks)

ii.    On average 2% of drugs dispensed are expired. In a sample of 300 units of such drugs, determine the probability that at least 2 units will be expired.  (3 marks)

(d) The weight distribution of male and female pigs reared by a firm are as follows.

| County | Sample size | Mean weight (kgms) | Standard deviation (kgms) |
|--------|-------------|--------------------|---------------------------|
| Males | 205 | 95.61 | 4.16 |
| Females | 315 | 83.35 | 5.56 |
| | | | |

(i)    Determine the combined mean                                    (2 marks)

(ii)   Determine the sex that has more variability in weight           (5 marks)


## QUESTION TWO (20 MARKS)

In a study to evaluate the effects of different processing methods on the texture of a food product, a food scientist wants to test whether there is a significant difference in the average texture scores (measured by a standard texture analyzer) between five different processing treatments. The treatments are:

- Treatment A: Steam cooking
- Treatment B: Boiling
- Treatment C: Baking
- Treatment D: Grilling
- Treatment E: Microwaving

Each treatment was tested with three replications, and the texture scores (in arbitrary units) are as follows:

| Treatment | Replication 1 | Replication 2 | Replication 3 |
|---|---|---|---|
| Treatment A (Steam cooking) | 35 | 38 | 37 |
| Treatment B (Boiling) | 40 | 42 | 41 |
| Treatment C (Baking) | 45 | 44 | 46 |
| Treatment D (Grilling) | 33 | 35 | 32 |
| Treatment E (Microwaving) | 50 | 49 | 48 |

Perform a One-Way ANOVA to test whether there is a significant difference in the mean texture scores between the five treatments. (14 marks)

b). Discuss the three principles of experimental designs. (6 marks)

**QUESTION THREE (20 MARKS)**

a). Suppose a researcher is studying the effect of two different preservatives (Preservative A and Preservative B) on the shelf life of a particular type of bread. Two separate batches of bread are prepared, one with Preservative A and the other with Preservative B, and the shelf life (in days) of 15 loaves from each batch is recorded.

The data for shelf life (in days) for each group is as follows:

- **Preservative A:** 8, 10, 12, 9, 11, 10, 8, 12, 9, 10, 11, 9, 8, 10, 12
- **Preservative B:** 6, 7, 7, 8, 6, 7, 6, 8, 7, 6, 8, 7, 6, 7, 8

Using a two-sample t-test with a significance level of 0.05, answer the following:

i. State the null and alternative hypotheses for this test. (2 marks)

ii. Calculate the mean and standard deviation for both samples. (4 marks)

iii. Perform the t-test and calculate the t-statistic and degrees of freedom. (5 marks)

iv. Read the critical value of the t statistic. (1 mark)

v. Interpret the results of the t-test and state whether you can reject the null hypothesis. (2 marks)

b). Highlight three probabilistic sampling techniques (6 marks)

## QUESTION FOUR (20 MARKS)

In a food science study, researchers conducted an experiment to examine the relationship between the type of packaging used (Plastic, Glass, or Paper) and the shelf life of a product (Short, Medium, or Long). The data collected is as follows:

| Packaging Type | Short Shelf Life | Medium Shelf Life | Long Shelf Life | Total |
|---|---|---|---|---|
| Plastic | 30 | 50 | 20 | 100 |
| Glass | 25 | 60 | 15 | 100 |
| Paper | 45 | 40 | 15 | 100 |
| Total | 100 | 150 | 50 | 300 |

i. Perform the Chi-square test for independence at a 5% significance level. Show all your calculations and interprete your results. (7 marks)

ii. Explain how this test can be done in SPSS (2 marks)

b) A firm manufacturing milk coolant reported that on average its coolant will have a lifespan of 6 years with a standard deviation of 0.4 years. In a certain period they sold 2500 coolants and 500 of them were replaced for breaking down before the warrant period.

     (i)     If the hospital buys a coolant from the firm, determine the probability that its lifespan will be,

        - Less than 4.8 years (2 marks)

        - Between 5 and 7 years. (3 marks)

     (ii)    Determine the firms warrant period for its vaccine coolants. (3marks)

     (iii)   Determine the 95% confidence interval for the mean lifetime of the firm vaccine coolants. (3 marks)

## QUESTION FIVE (20 MARKS)

In a study of food science, a researcher is interested in understanding how various factors affect the shelf life of a certain food product. The researcher collects data on 25 different food samples and measures the following covariates:

1. **Temperature** (in degrees Celsius)
2. **pH level** (on a scale from 0 to 14)
3. **Moisture content** (percentage)

Using multiple regression analysis, the researcher develops a model to predict the shelf life (in days) based on these three covariates.

i. Describe how the goodness of fit of the model can be assessed. What statistical test or measures would you use to evaluate whether the model fits the data well? (4 marks)

ii. Discuss how you would test the significance of the coefficients for each of the three covariates (Temperature, pH level, and Moisture content). Explain how to interpret the results of hypothesis tests for each coefficient, assuming a sample size of 25. (5 marks)

iii. Explain how this analysis can be done using SPSS software (5 marks)

b. A sample of 10 loaves of bread is tested, and the average shelf life is found to be 10 days, with a standard deviation of 2.5 days.Using a 95% confidence level, calculate the confidence interval for the true mean shelf life of this bread. (6 marks)