

054206T4AST

APPLIED STATISTICS LEVEL LEVEL 6

MATH/OS/AS/CR/04/6

Perform Inferential Data Analysis

March/April 2026



**TVET CURRICULUM DEVELOPMENT, ASSESSMENT AND CERTIFICATION COUNCIL
(TVET CDACC)**

WRITTEN ASSESSMENT

TIME: 2 HOURS

INSTRUCTIONS TO CANDIDATE

1. This paper consists of **TWO** sections: **A** and **B**.
2. Answer **ALL** questions in section A and **ANY THREE** (3) questions in section B.
3. Marks for each question are indicated in the brackets.
4. Candidates are provided with a separate answer booklet
5. Do not write on the question paper.

This paper consists of FOUR (4) printed pages

Candidate should check the question paper to ascertain that all pages are printed as indicated and that no questions are missing.

SECTION A (40 MARKS)

Answer ALL questions in this section

1. If a researcher surveys 100 households, inferential analysis can estimate the average income of all households in the county. Define the term inferential data analysis
(2 Marks)
2. Kenya Airways wants to summarize passenger numbers over the past year.
Give THREE measures of central tendency that they can use. (3 Marks)
3. County government has data with a mean revenue of 40 million KSh and a standard error of 4 million KSh. Calculate the 95% confidence interval using $\pm 2SE$. (4 Marks)
4. Before running a regression on Airtel Kenya's sales data, assumptions must be checked.
State FOUR reasons why statistical assumptions should be tested. (4 Marks)
5. Equity Bank wants to model factors affecting loan repayment rates. List FOUR statistical assumptions required for linear regression. (4 Marks)
6. A regression model analyzing KCB Bank's customer deposits shows an Adjusted R^2 of 0.82. State what the adjusted coefficient of determination represents. (2 Marks)
7. After selecting a suitable model for Kenya Airways' passenger data, parameter estimation is needed. Outline THREE steps involved in obtaining parameter estimates. (3 Marks)
8. A study examines factors affecting sales. Outline the difference between independent and dependent variables. (2 Marks)
9. Model output shows Adjusted $R^2 = 0.82$. State what adjusted R^2 represents. (4 Marks)
10. Kenya Pipeline Company tests whether a new pipeline monitoring system reduces leakages. State FOUR purpose of hypothesis testing in this context. (4 Marks)
11. If the coefficient of dataset "η" is 2.5 and the value of the independent variable x is 4, find the predicted value of y. (4 Marks)
12. XYZ Ltd is tracking the number of customer complaints each month.
State FOUR reasons why data should be recorded accurately. (4 Marks)

SECTION B (60 MARKS)*Attempts ANY THREE questions in this section*

13.

- a. The table below shows observed and expected frequencies of defects in a factory process.

Category	Observed (O)	Expected(E)
A	18	20
B	22	20
C	25	20
D	15	20

- Using the dataset, calculate the chi-square (χ^2) test statistic. (6 Marks)
 - Interpret the chi-square result at 5% level of significance. (2 Marks)
 - Interpret the chi-square result. (2 Marks)
- b. Kenya Pipeline Company wants to test whether a new pipeline monitoring system reduces leakages compared to the old system. Explain the steps involved in hypothesis testing for this scenario. (10 Marks)

14.

- a. a sample has a mean of 60 and a standard deviation of 10 with $n=25$
- Calculate the standard error of the mean. (5 Marks)
 - Calculate the 95% confidence interval using $\pm 2SE$. (5 Marks)
- b. Nation Media Group has conducted a regression analysis on advertisement spend and newspaper sales. Explain the key components of a report presenting these inferential analysis results. (10 Marks)

15.

- a. A regression model is given by $y=5+2x$
- Calculate the predicted value of y when $x=6$ (6 Marks)
 - Explain the meaning of the slope parameter. (4 Marks)
- b. Equity Bank wants to determine which factors best predict loan repayment rates. Explain those factors. (10 Marks)

16.

- a. Safaricom Ltd wants to present analysis results on customer subscriptions to management and stakeholders. Explain at least FIVE points on how a PowerPoint presentation of inferential analysis findings should be prepared. (10 Marks)
- b. The table below shows the monthly sales (in KSh '000) and advertising cost (in KSh '000) for a company.

Month	Advertising Cost (x)	Sales (y)
1	10	0
2	15	0
3	20	5
4	25	0
5	30	0

- i. Using the dataset, calculate the simple linear regression equation of sales on advertising cost. (6 Marks)
- ii. Use the model to predict sales when advertising cost is KSh 35,000 and interpret the result. (4 Marks)