Abbes Laghrour University Khenchela — *Khenchela, 09 May 2026*
Faculty of Natural and Life Sciences — *Duration: 1 hour 30 minutes*
Department of Natural and Life Sciences
2nd Year (Agronomy, Biology, Biotechnology, Ecology, Food Sciences)

Final Exam of Biostatistics

Part I (12 pts): Single-answer questions (16 questions = 0,75 points per question)

Q1. To represent the percentage distribution of plant species in a forest, the most suitable graph is:
A) Histogram
B) Bar chart
C) Pie chart
D) Boxplot
E) Scatter plot

Q2. An individual's blood type (A, B, AB, O) is a variable:
A) Discrete quantitative
B) Ordinal qualitative
C) Nominal qualitative
D) Continuous quantitative
E) Binary

Q3. In this series of masses (g): 42, 45, 47, 48, 210. The mean is 78.4 g while the median is 47 g. What property of the mean does this situation illustrate?
A) It is always above the median
B) It is insensitive to extreme values
C) It is strongly influenced by extreme values
D) The mean and the median always coincide
E) It can take several values

Q4. Standard deviation is preferred to variance to characterize dispersion because:
A) It is always below the mean
B) It is expressed in the same unit as the variable
C) It does not depend on extreme values
D) It is expressed in the square of the unit of the variable
E) It can be applied to qualitative variables

Q5. The effect of nitrogen dose applied on the yield of a cereal crop was studied. The independent variable is:
A) yield
B) The species of cereal
C) The dose of nitrogen applied
D) Rainfall
E) The duration of the vegetative cycle

Q6. The scatter plot is mainly used for:
A) Visualize the linear relationship between two quantitative variables
B) Show the temporal evolution of a variable
C) Comparing percentages between groups
D) Represent the distribution of a qualitative variable between three groups
E) Represent the frequency of a single categorical variable

Q7. The coefficient of variation (CV) is particularly useful for:
A) Calculate the median of a normal distribution
B) Compare the relative dispersion of two series measured in different units
C) Determine the degree of a correlation
D) Testing the independence of two qualitative variables
E) Estimating the covariance of two variables

Q8. The median of a statistical series:
A) Represents the most frequent value in the series
B) Separates the ordered series into two equal halves
C) Is equal to the mean for any distribution
D) It cannot be calculated for continuous variables
E) None of these answers are correct

Q9. The number of bacterial colonies in a petri dish is a variable:
A) Nominal qualitative
B) Continuous quantitative
C) Discrete quantitative
D) Ordinal qualitative
E) Binary

Q10. A researcher counts the number of pollinating insects out of 50 flowers (values: 0, 1, 2, 3, 4...). The most appropriate representation is:
A) Pie Chart
B) Histogram
C) Stick chart
D) Cumulative curve
E) Box plot

Q11. Regarding the Pearson correlation coefficient (r), which statement is correct?
A) $r = 0$ means that the two variables are correlated
B) Can be greater than 1 if the correlation is strong
C) $r = -0.90$ indicates a strong negative correlation
D) Can only take positive values
E) Depends on the unit of the variables

Q12. The interquartile range (IQR) is defined as:
A) $Q3 - Q1$
B) Max − Min
C) $Q2 - Q1$
D) $Q3 - Q2$
E) None of these answers are correct

Q13. The variation of egg hatching rate in urban pigeons according to rainfall was studied. The dependent variable is:

A) Rainfall
B) The species/breed of pigeons
C) Egg hatching rate
D) Incubation period
E) The height of the nest

Q14. The severity of disease stage (mild, moderate, severe, critical) is a variable:

A) Continuous quantitative
B) Discrete quantitative
C) Nominal Qualitative
D) Ordinal qualitative
E) Binary

Q15. For a series with high extreme values (outliers), which position or dispersion parameter is the most robust (which remains uninfluenced)?

A) Standard deviation
B) Median
C) Mean
D) Variance
E) Range

Q16. Which of the following position parameters can be calculated for a nominal qualitative variable?

A) Arithmetic mean
B) Median
C) Mode
D) 1st quartile
E) Standard deviation

Part II *(8pts)****: Exercises*** *(8 questions = 1 point per question)*

Exercise 1

In an entomological study, the number of larvae of a pest (X) found on each of the 50 tomato plants in an experimental plot was counted. The results are summarized in the following table.

x_i (number of larvae)	0	1	2	3	4	5	6
n_i (number of tomatoes)	5	12	15	10	5	2	1

Calculate the position and dispersion parameters in the table on the response page.

Exercise 2

A clinical researcher studies the relationship between systolic blood pressure (X, in mmHg) and cholesterol level (Y, in mg/dL) in the only seven patients diagnosed with a rare genetic disorder in the country.

Patient	1	2	3	4	5	6	7
X (mmHg)	14	10	20	16	22	12	18
Y (mg/dL)	10	12	6	8	5	11	7

Answer these questions in the table on the response page.

- Calculate the covariance
- Calculate the correlation coefficient r
- Determine the simple linear regression equation $Y = aX + b$

Good luck

Khenchela, 09 May 2026

Final Exam of Biostatistics – Response page

2nd Year (Agronomy, Biology, Biotechnology, Ecology, Food Sciences)

***Part I** (12 points)*

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
C	C	C	B	C	A	B	B

Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
C	C	C	A	C	D	B	C

***Part II** (8 points)*

***Exercise 1** (5 points)*

Parameter	Formula	Result
Mean	$\bar{x} = \frac{\sum(ni.xi)}{n} = 108/50$	$\bar{x} = 2.16$ larvae
Median	$Me = (Modality\ of\ n/2 + Modality\ of\ (n/2+1))/2$ $= (2+2)/2$	Me = 2 larvae
Interquartile range	$IQR = Q3 - Q1 = 3 - 1$	IQR = 2 larvae
Standard deviation	$\delta = \sqrt{v} = \sqrt{\frac{\sum ni(xi-\bar{x})^2}{n-1}} = \sqrt{94}, 72/49$	$\delta = 1.93$ larvae
Coefficient of variation	$CV = \frac{\delta}{\bar{x}} . 100$ $= (1.93 / 2.16) \times 100$	CV = 64.35%

***Exercise 2** (3 points)*

Parameter	Formula	Result
Covariance	$Cov\ (x,y) = \frac{\sum (xi - \bar{X})(yi - \bar{Y})}{N} = -68/7$	Cov (x,y) = −9.71
Correlation coefficient	$r(x;y) = \frac{cov\ (x,y)}{(\delta x \times \delta y)} =$ -9.71/ 4 ×2.44	r = −0.995 → strong negative correlation
Linear regression equation	$a = \frac{Cov\ (x,y)}{v_x} \rightarrow$ a = −0.607 $b = \bar{Y} - a \cdot \bar{X} \rightarrow$ b = 18.14	Y = −0.607X + 18.14