NATIONAL UNIVERSITY OF ENGINEERING

COLLEGE OF ENVIRONMENTAL ENGINEERING

SANITARY ENGINEERING PROGRAM

AA232 – BIOSTATISTICS

I. GENERAL INFORMATION

CODE	: AA232 Biostatistics
SEMESTER	: 3
CREDITS	: 4
HOURS PER WEEK	: 6 (Theory – Practice – Labs)
PREREQUISITES	: AA221 Mathematics II
CONDITION	: Compulsory

II. COURSE DESCRIPTION

This course prepares students in the application of concepts, methods and techniques of the differential and descriptive statistics to describe and analyze data and variable sets. Concepts of probability and random variables are introduced and applied to predict expected future values. Regression techniques are applied to build models relating a data and variable set. Problems of engineering application are solved and specialized software is used too.

III. COURSE OUTCOMES

1. Organize data for their adequate analysis and interpretation and calculate and interpret their fundamental statistical properties (Average value y variance).
2. Explain and determine the random probability of events and variables, as well as their probability density function.
3. Understand and apply random vectors and determine their function of joint probability density.
4. Interpret the concept of sampling distribution and apply it to calculate the probability of an event or variable.
5. Build models of linear regression to represent the relationship between the representative parameters of a data set.
6. Apply chi-square tests to check the probability density function of a data set, as well as the dependence or independence of two variables.

IV. LEARNING UNITS

1. DATA ORGANIZATION AND MAIN STATISTICAL PARAMETERS / 8 HOURS
Statistics / Population, sample and variable / Classification of variables / Methods for organizing and presenting data / Qualitative data / Frequency distribution table / Graphic representations / Measures of central tendency: mean, median, weighted mean / Dispersion methods: variance, standard deviation, coefficient of variation / Measures of quartile, decile and percentile positions / Box plots.

2. PROBABILITY / 8 HOURS
Count methods / Addition and multiplication rules / Permutations and combinations / Probability / Random experiments, sampling space and events / Operations with events / Conditional probability / Total probability / Bayes' theorem / Tree chart / Independent events.

3. RANDON VARIABLES AND PROBABILITY DISTRIBUTION / 12 HOURS
Random variable: Types, expected value and standard deviation / Distribution of probability of a discrete variable: binomial, Poisson and hyper-geometric / Distribution of probability of a continuous variable: uniform, normal, gamma and exponential.

4. RANDOM VECTORS / 4 HOURS

Joint probability density function / Marginal distribution / Expected vector, variance and covariance / expected value of two random functions / Covariance of two random variables.

5. SAMPLING DISTRIBUTION / 4 HOURS

Sampling / Sampling techniques / Probability sampling / Sampling distribution of average value, the variance / Central limit theorem.

6. ESTIMATION AND HYPOTHESIS TESTING / 12 HOURS

Punctual estimation of a random sampling / Punctual estimators: mean, mean proportional, difference of means, variance / Hypothesis testing / Interval estimation. Confidence intervals: mean, mean proportional and variance / Hypothesis testing / Types of error / Correlation coefficient / Prediction intervals / Hypothesis testing on regression coefficients.

7. REGRESSION AND CORRELATION ANALYSIS / 8 HOURS

Simple regression / Scatter plot / Parameter estimation / Total variance decomposition / Determination coefficient / Estimation errors / Correlation coefficient / prediction intervals / Hypothesis testing on regression coefficients.

8. CHI-SQUARE TESTS / 8 HOURS

Setting tests for uniform, binomial and Poisson's distributions / Contingency tables / Independence tests,

V. LABORATORIES AND PRACTICAL EXPERIENCES

Lab 1: Generation of random signals.
Lab 2: Determination and graphs of the probability density function.
Lab 3: Linear and quadratic regression.

VI. METHODOLOGY

This course is carried out in theory, practical and lab sessions. In theory sessions, the instructor introduces concepts, theorems and applications. In practical sessions, several problems are solved and their solution is analyzed. In lab sessions, Minitab simulation software is used to solve problems and analyze their solutions. At the end of the course, students should hand in and expose an integrating paper and project. In all sessions, students' active participation is encouraged.

VII. EVALUATION FORMULA

The average grade PF is calculated as follows:

PF = 0.25 EP + 0.25 EF + 0.10 P1 + 0.10 P2 + 0.07 L1 + 0.07 L2 + 0.07 L3 + 0.09 TF

EP: Mid-Term Exam	EF: Final Exam	TF: Final paper
P#: Quizzes	L#: Final paper	

VIII. BIBLIOGRAPHY

1. **MENDENHALL, WILLIAM**
   Statistics and probability for engineering (Spanish)
   Prentice Hall Editorial (2012)

2. **ALVAREZ, JOSÉ AND TORRES LUIS**
   Probability and Statistics (Spanish)
   Alfa Omega Editorial (2012)

3. **MONTGOMERY, DOUGLAS AND RUNGER GEORGE**
   Probability and Statistics Applied to Engineering (Spanish)
   McGraw Hill Editorial (2010)