

**COURSE CODE: MST 1102**

**COURSE TITLE: Introduction to Statistics**

**COURSE CREDITS: 4      Lecture hours 4      Tutorial/lab hours 1**

**COURSE DESCRIPTION:**

Introduction to statistics is designed to assist the students in acquiring a good intuitive grasp of statistics, specifically in terms of what it is, how and when to apply various statistical techniques to various situations, to interpret results and to draw meaningful conclusions from data. The teaching/learning process is enhanced by computer laboratory session where students will use non-programmable software to obtain statistical output for interpretation and presentation. The recommended package is SPSS which has a large menu system that simplifies the task of instructing the software. This will be useful especially to beginners.

In today's world, the use of simple but powerful statistical tools is critical to successful business practices. Introduction to statistics will give students exposure to a wide range of statistical tools and concepts required in various fields such as the environmental sciences, quality management, as well as in various business fields and other functional areas. The ability to collect, organize, manage, interpret, and present data, is a powerful asset to possess. These skills will be developed through the study of various statistical methods and techniques in this course. Additionally, the student will begin to develop a systematic approach to problem solving and statistical decision making.

This course also serves as a core basic level course for all students in the Agriculture, Biology, Chemistry, Medicine and Forestry programs. This course seeks to introduce students to the principles, concepts and practices of statistics of applied sciences. This course entails descriptive and inferential statistics, analysis of result, and interpretation of computer output and presentation of findings. After completing the course students will be able to identify data types,

manage them and present them in a clear and logical format. They will be able to use statistics in their every daily life.

This course is introductory in nature and will equip students with the intuitive bridge necessary in understanding statistics in other more advanced statistics courses. Assessment is designed to encourage students to work continuously with the course materials. Active learning will be achieved through weekly assignments and problem sheets allowing continuous feedback and guidance on problem solving techniques in tutorials and lectures. Assessment will be based on assignments and in-course tests followed by a comprehensive final examination.

**PRE-REQUISITES/CO-REQUISITES/FOLLOW-ON COURSES:**

**PRE-REQUISITES:** University matriculation

**CO-REQUISITES:** MTH1101 – Algebra

**FOLLOW-ON COURSES:** MST2102 - Probability, Distribution Theory & Inference

**STUDENT LEARNING OUTCOMES:**

After taking this course, students will be able to:

1. Distinguish between different data types.
2. Differentiate among statistical terminologies
3. Organize, manage and present data
4. Organize data for appropriate statistical analysis
5. Produce and interpret graphical summaries of data
6. Demonstrate proficiency in analyzing data using the appropriate statistical method.
7. Generate a wide variety of statistical output using a statistical software.

**COURSE CONTENT:**

Definition & scope of statistics. Importance of statistics in the sciences. Characteristic of data. Variable: quantitative and qualitative variable. Measurement scales: Nominal, Ordinal, Interval and Ratio scales. Raw data, categorized data. Discrete & continuous variable, Accuracy and precision of data.

Definition and explanation of terms used in statistics: population, sample, replication and randomization, parameter, statistics, estimate, mean and variance, parametric vs. non-parametric statistics.

Data Presentation: Tabular, graphical, Frequency distribution, histogram & Frequency polygon, Cumulative frequency distributions, Relative frequency distribution, Stem and leaf display, box and whisker plot, bar chart, line graphs & pie chart. Group data presentation: Class interval, class boundary and class width. Graphic presentation of frequency distribution: Computer aided presentations.

Statistics of location/ Central tendency. Measures of central tendency: mean, median and mode. Relationship between mean, median and mode. Computer aided calculations. Measures of Dispersion: variance, standard deviation, Coefficient of Variation, methods for computing mean and standard deviations, standard errors. Computer aided calculations

Probability distribution. Probability concepts: events, space, independent and dependent events. Relative frequencies and subjective probabilities, complement and independent trials. Conditional Probability, disjointed and independent events, rules of addition and multiplication, Discrete Random variables: Expectation and variance; Binomial random variables; Mean, variance and probability of binomial random variable.

Continuous Random Variable: Normal distributions; probabilities of a normal distribution. Finding probabilities using a standard normal table. Finding the proportion under the curve. Sampling distribution of a sample mean; Sampling distribution for sample proportion. Assumptions of parametric statistics.

Hypothesis testing of a single mean. Type 1 and Type II errors. Introduction to t distribution. Introduction to confidence intervals. Confidence intervals for population mean. Finding sample size for estimating a population mean. One sample analysis: One tailed and two tailed hypothesis tests.

Comparing two groups. Hypothesis testing of two means. Comparing two independent means; comparing paired means. Constructing confidence Interval for the difference of two means. Computer aided calculation and its interpretation using statistical software.

Comparing multiple means. Analysis of Variance - F-Ratio.

Introduction to regression. Regression equation, slope and intercept. Coefficient of determination. Simple linear regression using software. Linear Regression. Transformation of

data. Correlation and correlation coefficient. Computer aided calculation and its interpretation using statistical software.

Non-Parametric comparisons of two populations: Association tests and Goodness of fit. Chi - square distribution. Proportion and count data, Contingency tables.

Presentations of research data of group assignment. Computer aided calculation and its interpretation using statistical software.

### **Laboratory Content Guide**

Understanding the software system and how it is organized. Reading in and imputing data. Producing and saving and exporting results. Transferring results to a word document. Reviewing variables and metadata in the software system. Labelling variables and data values. Setting missing value codes. Recoding data. Transforming data.

#### *Data Tabulation:*

Frequency and relative frequency distributions, cross-tabulations (contingency tables)  
Cumulative frequency distributions

#### *Diagrams:*

Stem and leaf display, bar charts, pie charts, box and whisker plots, histograms & frequency polygon, frequency and cumulative frequency curves, line graphs.

#### *Measures of central tendency & Dispersion:*

Mean, median and mode, minimum, maximum, variance, standard deviation, standard errors

Estimating and comparing means and proportions. Generating test results and producing confidence intervals with different levels of confidence.

Paired t-test, independent samples t-test (also assuming that the variances are equal)

Generating results for Analysis of variance

Computing correlation coefficients

Computing regression equations and conducting hypothesis tests about regression equations and coefficients.

Producing cross-tabulations. Conducting tests of association and estimating odds ratios.

### **METHOD OF TEACHING:**

1. Interactive learning: Interactive Lectures
2. Collaborative learning: group assignments and presentations
3. Active and experiential learning: Field practical demonstration and exercises
4. Active and interactive learning: Computer aided analysis

## 5. Computer Laboratory sessions

This course will be delivered by a combination of lectures, tutorials and computer laboratory. There would be 4 lecture hours and 1 tutorial/lab hour per week. Tutorials are practical sessions which may include problem solving sessions and/or supervised laboratory time. Group work activity will be used in online activities to engender team building and to assist with the problem solving. The tutorial will be interspersed with the lectures by having students discuss exercises, revise material as needed, and cover new content each day. Course materials such as exercises, assignments, solutions will be posted on the University's online system.

Additional problems will be given during lectures and tutorials but will not be marked. However, students will need to do some of these, as well as the assignments, in order to learn the material properly and to adequately prepare for examinations.

### **METHODS OF ASSESSMENT:**

Coursework:

Two Tests	20%
Lab assignments and quizzes	10%
Group lab project	10%
<b>Course Work total</b>	<b>40%</b>
<b>Final Exams (one 2-hour paper)</b>	<b>60%</b>

### **RECOMMENDED READINGS:**

Chase, W., & Brown, F. (1999) General Statistics, 4<sup>th</sup> Edition, John Wiley & Sons Inc.

Devore, J. (2015). Probability and Statistics for Engineering and the Sciences. 9<sup>th</sup> Edition, Cengage Learning.

Triola, M. F. (2014). Elementary Statistics. 12<sup>th</sup> Edition. Pearson Education.

Weiss, N. A., & Weiss, C. A. (2012). Introductory statistics. 9<sup>th</sup> Edition, Pearson Education.

Clewer, A.G. & Scarisbrick, D.H. (2001) *Practical Statistics and Experimental Design for Plant*

*and Crop Science*. John Wiley & Sons Ltd. New York.

Gardener, M. (2012) *Statistics for Ecologists Using R and Excel: Data Collection, Exploration, Analysis and Presentation*. Pelagic Publishing.