



**FORMAT FOR COURSE CURRICULUM**

**Course Title: Statistical Methods & SAS**

**Course Code: BIOF612**

**Units: 03**

L	T	P/S	SW/F W	TOTAL CREDIT UNITS
3	0	0	0	3

**Course Objectives:**

Theory: The main objective of this course is to enable students to make meaningful interpretations of the biological data using statistical methods and then process those using SAS.

**Pre-requisites:** Basic knowledge of computational biology, basic mathematics.

**Student Learning Outcomes:**

- The student will apply statistical knowledge to analyze Clinical data.
- The student can discriminate between good and worst trails and can calculate the differences between various parameters.

**Course Contents/Syllabus- Theory:**

	Weightage (%)
<b>Module I</b>	<b>10</b>
<b>Descriptors/Topics: Basics of Statistics</b> Introduction, Data analysis steps, types of biological variables, probability, basic concepts of hypothesis testing, confounding variables, Overview on statistical modeling in biology.	
<b>Module II</b>	<b>20</b>
<b>Descriptors/Topics: Descriptive Statistics</b> Statistics of central tendency-mean, median and mode, Statistics of dispersion, standard error of mean, confidence limits	
<b>Module III</b>	<b>20</b>
<b>Descriptors/Topics: Estimation Theory</b> Biased and unbiased estimator, Confidence interval: population mean and proportion	
<b>Module IV</b>	<b>20</b>

<p><b>Descriptors/Topics: Inference- Tests of hypotheses</b></p> <p>Formulation of Hypothesis Simple and Composite, Type I and Type II errors Power of a test, Significance of a test, P-value Testing, Z, t-test and F-test, Chi-square. Comparison of two samples, Correlation and Linear Regression</p>	
<p><b>Module V</b></p>	<b>20</b>
<p><b>Descriptors/Topics: Basics of Statistics Softwares</b></p> <p>SAS, SPSS and Stata, Inference of population structure and Gibbs sampling, Inference of regulatory motifs and the EM algorithm, Predicting disease phenotypes using microarrays, Modeling DNA methylation and epigenetics-basics.</p>	
<p><b>Module VI</b></p>	<b>10</b>
<p><b>Descriptors/Topics: Introduction of SAS Environment</b></p> <p>SAS syntax, variable naming rules, and entering your own data. Creating a SAS data set from raw data, Grouping Data values, Reading and combining SAS data sets. Relating information from multiple sources, Data translation tools, Manipulating and formatting date values, , Data Step and Proc Step, Manipulating the Data- Converting the numeric data to character and vice-versa. Using Logical Operators and Where Conditions, Merging of Datasets.</p>	

**Pedagogy for Course Delivery:**

The class will be taught using theory and practical based method. In addition to assigning the practical problems, the course instructor will spend considerable time in understanding the concept of statistics. The instructor will cover the statistics fundamentals used to analyze the data using SAS techniques.

Lectures: 41  
 Tutorial:  
 Presentation/ Seminar: 3  
 Class Test: 1  
 Total: 45

**Assessment/ Examination Scheme:**

Theory L/T (%)	Lab/Practical/Studio (%)	Total (%)
<b>100</b>	<b>NA</b>	<b>100</b>

**Theory Assessment (L&T):**

Continuous Assessment/Internal Assessment						End Term Examination
Components (Drop down)	Class Test 1	Viva	Home Assignment		Attendance	
Weightage (%)	15	5	5		5	70

**Text :**

- W.J. Ewens and G.R. Grant, “Statistical Methods in Bioinformatics” An Introduction. Springer-Verlag 2001. ISBN: 978-0387400822
- Kleinbaum DG, Kupper LL and Muller KE, “Applied Regression Analysis and Other Multivariate Methods” PWS-KENT Publishing Company, Boston (Second Edition). ISBN: 978-0495384960.

**References:**

- 1 J.L. Devore, “ Probability and Statistics”, (fifth edition), Thomson Asia, 2002