COURSE OUTLINE
SZL 303 BIOSTATISTICS

Course Description:

This course explores the meaning of statistics. It introduces students to some basic terms like variable, continuous variable, discrete or discontinuous variables population, sample, histogram, frequency, classes, class interval and frequency distribution; a distribution in statistical terms: mode, median, mean; measuring the spread of a distribution: range, semi-interquartile range, mean deviation, variance, standard deviation; samples and populations: probability and the normal distribution curve, distribution of t, calculating the limits of a mean; and comparing the means of two samples: null hypothesis, alternate hypothesis, differences between standard deviations, limits for standard deviation and variance. The course also examines a comparison of three or more samples: simple analysis of variance; correlation of two variables: scatter diagram, correlation coefficient, regression lines, mean centre; and chi-square test: the 2x2 contingency table. Learners are taken through planning experiments: layout of experiments, controls, precision of measurements, number of replicates, randomization, Latin squares, and interaction.

Practical sessions involve data analyses and problem solving using computers.

Required Learning Materials

Chernick M. R. and Friis R.H. (2003). Introductory Biostatistics for the health sciences. JOHN-WILEY AND SONS INC. New Jersey.

Recommended Reading

1. Cox D.R. (1990) Planning of experiments. WILEY INTERNATIONAL EDITION.
2. Cramer D. (1997). *Basic statistics for social research*. London: ROUTLEDGE.
3. Haber A. and Runyon, R. P. (1977). *General statistics*. London: ADDISON-WESLEY
4. PUBLISHING COMPANY.
5. Wardlaw, A. C (1985) Practical statistics for Experimental Biologists. JOHN WILEY AND SONS LTD.
6. Zar, J.H. (1984). *Biostatistica Analysis* (4th Ed). PRENTICE HALL INTERNATIONAL, INC.

Student Learning Outcomes

Upon successful completion of the course the learner will be able to:

- analyze different experimental designs for generation of qualitative and quantitative data;
- design research hypotheses and generate appropriate data; and
- generated data to appropriate statistical analysis and give relevant interpretation to the output.

Learning Objectives

Upon completion of the course the learner is expected to do the following:

- distinguish qualitative and quantitative data;
- Evaluate strength of different biological experimental designs;
- Evaluate different data analysis methods; and
- Perform statistical analysis using a few computer packages.

Assessment Opportunities

Semester course assessments will include class attendance and active participation;
two continuous written assessment (CAT) and a final examination. Course assessment tasks will follow the guidelines described below:

- all forms of assessment will foster creative critical thinking and tasks that promote application of learning, e.g., project, case study, and other forms of hands-on/application-based activities.
- an assessment task must address at least one learning objective.
- the date for CATs will be communicated.
- assignments will be varied: reading assignments on the course outline will be followed up and at list two times be made into unannounced quizzes that are graded. This will respond to the never-ending question of lack of reading culture by students.

Course Requirements

- regular attendance and participation in class discussions;
- sitting all the required CATS and the final examination;
- adequate preparation for every class;
- completing on time the reading assignments;

Teaching Methodologies

1. Lectures
2. Class discussions
3. Assignments
4. Group discussions

Teaching Equipment

1. LCD Projector
2. Computer
3. Chalkboard

Course Schedule:

Lecture Topic	Learning detailed outline	Reading Assignments
INTRODUCTON TO BIOSTATISTICS	• Meaning of biostatistics • Types of variables • Population and samples ✓ Populations ✓ Samples from populations ✓ Random sampling ✓ Parameters and statistics	Read on types of biological data Stem-and-leaf diagrams Cyclic and circular data
	References used Chernick & Friis (2003) Pages 1 - 46	Follow Up Activities Discuss applications and limitations of Stem-and-leaf diagrams
PRESENTATION OF	• Frequency distribution	Drawing histograms

BIOLOGICAL DATA	distribution	
	References used Chernick & Friis (2003) Pages 46 - 61	Follow Up Activities Of a given set of data, draw a frequency distribution curves and polygons

Lecture Topic	Learning detailed outline	Reading Assignments
PROBABILITY AND STATISTICS	• Probability ✓ Laws of probability o Counting possible outcomes o Probability of an event o Adding probabilities o Multiplying probabilities • Permutation and combinations ✓ Permutations ✓ Combinations	Read of generalised multiplication rule Probabilities with equal outcomes Baye's theorem
	References used Chernick & Friis (2003) Pages 92 - 109	Follow Up Activities Assignment and computer exercise
NORMAL DISTRIBUTION	• Symmetry and Kurtosis • Proportions of normal distribution • The distribution of means • Statistical hypothesis testing • Assessing departures from normality	Read on standard Normal distribution and probabilities
	References used Chernick & Friis (2003) Pages 121 - 122	Follow Up Activities A normally distributed population of maize seed weights has a mean of 63.5g and a standard deviation of 12.2g. a) What proportion of this population is 78.0g or larger? b) What proportion of this population is 78.0g or smaller? c) If there 1000 weights in this population, how many of them are 78.0g or larger? d) What is the probability of choosing at random from this population a weight smaller than 41.0g?

Lecture Topic	Learning detailed outline	Reading Assignments
BIONOMIAL AND POISSON DISTRIBUTION	• Binomial distribution • Poisson distribution	Normal approximation to Poisson distribution
	References used Chernick & Friis (2003) Pages 109	Follow Up Activities A total of 24 potted plants were used in an experiment designed to test nitrogen fixation potential of a Rhizobia strain on three different plant species. Twelve (12) potted plants of *Sesbania* sp., eight (8) of *Calliandra* sp. and four (4) of *Acacia* sp. How many different sequences of these species are possible in a greenhouse setup
STANDRAD ERROR AND CONFIDENCE INTERVAL	• Standard error • Confidence interval	Variance and standard deviation for grouped data Calculating the combined mean and variance of several samples
	References used Chernick & Friis (2003) Pages 133 - 161	Follow Up Activities The time taken for cessation of bleeding was recorded for a large number of persons whose fingers had been pricked. The mean time was found to be 1.407 min and the standard deviation was 0.588 min. In an effort to determine whether pressure applied to the upper arm increases bleeding time, six persons had pressure equal to20 mmHg applied to their upper arms and had their fingers pricked. For these six persons, the times taken for bleeding to stop were 1.15, 1.75, 1.32, 1.28, 1.39, and 2.50min. a) State appropriate Null hypothesis for this experiment b) Give a 95% confidence interval for the mean bleeding time under pressure for the six persons c) Draw a conclusion as to whether pressure increases bleeding time or not.

Lecture Topic	Learning detailed outline	Reading Assignments
HYPOTHESIS TESTING	• Null hypothesis and alternative hypotheses • The standard format for hypothesis testing	General formulation of hypothesis and hypothesis testing
	References used Chernick & Friis (2003) Pages 182 - 191	Follow Up Activities Assignment and computer exercise
THE *t*- DISTRIBUTION	• One group of observations (or one sample test) • Two independent group of observation ✓ Variances not known ✓ Variances known	Confidence intervals of means using *t* distribution Comparison of variances
	References used Chernick & Friis (2003) Pages 193 – 195	Follow Up Activities Assignment and computer exercise

Lecture Topic	Learning detailed outline	Reading Assignments
Mid-term revision and examination	• Testing topics covered in weeks 1, 2, 3 and 4	Revise topics covered during week 1,2, 3 and 4

Lecture Topic	Learning detailed outline	Reading Assignments
THE CHI-SQUARE DISTRIBUTION	• Testing the Goodness of fit • Test of independence ✓ The 2 x 2 Contingency table ✓ The 2 x 4 Contingency table	Goodness of fit to prescribed probabilities Yate's correction
	References used Chernick & Friis (2003) Pages 232 - 239	Follow Up Activities Assignment and computer exercise
ANALYSIS OF VARIANCE	• One-way (Single factor) ANOVA ✓ Equal replication (Sample size) ✓ Unequal replications (sample size) • Two-way (factor) ANOVA • Multiple comparisons ✓ The Tukey test ✓ The Newman-Keuls test • Assumptions of ANOVA • Data transformation ✓ Logarithmic ✓ Square root ✓ Arcsine (angular	Read on Experimental designs Sample size determination.
	References used Chernick & Friis (2003) Pages 296 – 301	Follow Up Activities Assignment and computer exercise

Lecture Topic	Learning detailed outline	Reading Assignments
SIMPLE LINEAR REGRESSION	• Simple linear regression equation • Simple linear regression analysis	Read on why and when to choose regression analysis How to handle outliers
	References used Chernick & Friis (2003) Pages 252-259	Follow Up Activities Assignment and computer exercise
SIMPLE LINEAR CORRELATION	Simple linear correlation	Product-moment correlation coefficient Testing the significance of *r*
	References used Chernick & Friis (2003) Pages 252 - 277	Follow Up Activities Assignment and computer exercise

Lecture Topic	Learning detailed outline	Reading Assignments
NON-PARAMETRIC STATISTICS	• The Mann-Whitney ✓ The two-tailed Mann-Whitney test ✓ The one-tailed Mann-Whitney test with tied ranks • The median test • Kruskal-Wallis test	Read on Non-parametric analysis
	References used Chernick & Friis (2003) Pages 308 - 324	Follow Up Activities Assignment and computer exercise
COMPUTER PROGRAM AND REVISION FOR EXAMINATION	Introduction to data analysis in excel and R	Follow Up Activities Assignment and computer exercise