Institute of Statistical Research and Training
University of Dhaka

Curriculum

B.S. Honours Program in Applied Statistics and Data Science

Session: 2023–2024

www.isrt.ac.bd/academics/undergraduate

Contents

1 Overview of the Program

1.1 B.S. Honours Program in Applied Statistics and Data Science

The B.S. honours program in Applied Statistics and Data Science is an integrated four-year program offered by the Institute of Statistical Research and Training (ISRT), University of Dhaka. A student with high academic attainment in Secondary School Certificate (S.S.C.) and Higher Secondary School Certificate (H.S.C.) or equivalent levels with Mathematics as a subject of study is eligible for admission. Unless otherwise stated, the regulations for admission of the students and the examinations will be the same as those for the B.S. honours courses in the Faculty of Science. The program includes theoretical and applied courses with intensive computing facilities. However, the program emphasizes applying statistical techniques and data science tools to real-life problems arising in medicine, engineering, business, and social, environmental, and biological sciences. The program is so designed that graduates are equipped to work efficiently and competently in government and non-government organizations, research organizations, service departments, and other related fields.

1.2 The University of Dhaka

The University of Dhaka, established in 1921, is one of the leading universities in Bangladesh for higher education and research. The University assumed a central role in the academic pursuits of the region, including Bangladesh, and has passed through tumultuous times at different periods of our national history and played vital, at times, pioneering roles in all critical junctures in making this nation great. The University of Dhaka started its activities with 3 Faculties, 12 Departments, 60 teachers, 877 students, and three dormitories (Halls of Residence) for the students. The University currently consists of 13 Faculties, 83 Departments, 12 Institutes, 20 residential halls, three hostels, and more than 56 Research Centres. The primary purpose of the University is to create new areas of knowledge and disseminate this knowledge to society through its students.

1.2.1 Vision and Mission of the University

Vision

Create a world-class educational ecosystem that enables individuals to act as dynamic human capital and ethical leaders for a sustainable future.

Mission

MU1 **Transformative Education:** Provide transformative education by enabling students to embrace lifelong learning and fostering a sustainable knowledge-based society through the continuous pursuit of scholarship, humanistic values, and technological innovation.

MU2 **Collaborative Research and Innovation:** Pursue collaborative research and innovation, leveraging partnerships to expand the boundaries of knowledge.

MU3 **Educational Ecology:** Develop an educational ecosystem that fosters excellence, transparency, inclusivity, and accountability.

MU4 **Community Engagement:** Engage with stakeholders and communities to build a just, fair, diverse, and sustainable world.

MU5 **Ethical Responsibility:** Encourage students to become ethically responsible global citizens with a positive societal impact.

MU6 **National Heritage:** Instill a deep sense of national heritage and pride in students, upholding historical roots and global connectivity.

1.3 The Institute of Statistical Research and Training (ISRT)

1.3.1 Overview of the Institute

The Institute of Statistical Research and Training (ISRT), University of Dhaka, is the leading institution for training and research in Applied Statistics and Data Science in Bangladesh. It was founded in 1964 by the Late National Professor Dr. Qazi Motahar Husain, an eminent scientist, academician, and a leading proponent of the statistical sciences in this country. The Institute offers a 4-year B.S. Honours program designed to produce graduates with strong statistical computing skills, sound knowledge of statistical concepts, and the versatility to apply these concepts in areas as diverse as medicine, engineering, business, economics, and the social sciences. The 1-year M.S. program consists of specialized courses in areas ranging from environmental statistics to statistical signal processing, statistical machine learning, and causal inference. It has been designed for students interested in higher studies and research. In addition, the Institute offers Ph.D. and M.Phil. degree programs. Highly experienced faculty members, most of whom have Ph.D. degrees from reputed universities across the world, run these programs.

ISRT boasts an academic environment that is highly competitive and conducive to research. Both students and faculty members benefit from the regular seminars and talks given by researchers from home and abroad on topics of current interest. The Institute has a rich library with well over 15,000 books. It has three state-of-the-art computer labs, cloud computing facilities, and high-speed internet access for graduate and undergraduate students. The aim is to provide a learning environment that stimulates intellectual curiosity, critical thinking, and independent problem-solving skills. The Journal of Statistical Research (JSR), an international journal published bi-annually by ISRT since 1970, is a forum for exchanging research ideas between statisticians in Bangladesh and abroad. Faculty members have research interests in diverse areas such as biostatistics, machine learning, spatial statistics, statistical pattern recognition, public health, Bayesian analysis, and econometrics. They regularly disseminate their research works in peer-reviewed journals and international conferences.

Among its other activities, the Institute frequently organizes short courses and training programs for non-statisticians working in government and non-government organizations who need statistical analysis. In doing so, it has played an active role in promoting and creating awareness about the need for sound statistical practices among people from other disciplines so that they may work more efficiently within their organizations. ISRT also maintains close ties with the Bangladesh Bureau of Statistics (BBS) and other organizations responsible for collecting and disseminating statistical data in Bangladesh. It is frequently called upon to offer its expertise on statistical issues of national interest. Over the years, the Institute has played a significant role in the country's development by producing world-class statisticians for academia and industry and providing statistical expertise on issues of national interest. In addition, the Institute provides statistical consulting services through StatLab primarily for the students and faculty members of the University of Dhaka, aiming to strengthen research on campus by assisting graduate students and faculty members of other disciplines.

1.3.2 Vision and Mission of the Institute

Vision

To take a leading role in producing competent graduates in Applied Statistics and Data Science, conducting cutting-edge research, and creating industrial partnerships to address national and global interests and challenges.

Mission

To pursue excellence in Applied Statistics and Data Science education and research and to provide data-driven solutions to industries and stakeholders for the benefit of society.

- M1 To provide quality education in Applied Statistics and Data Science by ensuring an effective learning environment
- M2 To perform original and impactful research in Applied Statistics and Data Science that would enhance knowledge and contribute to the well-being and advancement of society.
- M3 To provide innovative data-driven solutions to the problems and challenges the industries and other stakeholders face.

1.3.3 Program Education Objectives (PEOs)

- PEO1 To produce graduates with strong theoretical and practical knowledge of Statistics and Data Science in line with the market demand.
- PEO2 To facilitate high-quality research in statistical theory and data science with applications to relevant fields for the betterment of society.
- PEO3 To prepare graduates capable of performing interdisciplinary and collaborative research
- PEO4 To serve the statistical and data science needs of government, industry, and other stakeholders.
- PEO5 To produce graduates with strong leadership, teamwork, and communication skills.
- PEO6 To prepare graduates with ethical and moral values that will help them in their professional lives.

Table 1.1: Mapping of the mission of the Institute (MI) with the mission of the University (MU)

	Missions of the University (MUs)					
MIs	MU1	MU2	MU3	MU4	MU5	MU6
M1	3	2	3	2	2	2
M2	2	3	2	3	2	2
M3	2	3	2	3	2	2

Note: Scale to explain the extent of matching: 3= high, 2= medium, 1=low

Table 1.2: Mapping of the program education objectives (PEOs) with the missions of the University (MUs)

	Missions of the University (MUs)					
PEOs	MU1	MU2	MU3	MU4	MU5	MU6
PEO1	3	2	3	2	2	1
PEO2	2	3	2	2	2	2
PEO3	2	3	2	2	2	2
PEO4	2	2	2	3	2	1
PEO5	1	2	2	3	2	2
PEO6	2	2	2	2	3	2

Note: Scale to explain the extent of matching: 3= high, 2= medium, 1=low

Table 1.3: Mapping of the program education objectives (PEOs) with the mission of the Institute (MI), where the scales are to explain the extent of matching (3=high, 2=medium, and 1=low)

	MIs		
PEOs	M1	M2	M3
PEO1	3	3	1
PEO2	3	3	1
PEO3	2	3	3
PEO4	2	1	3
PEO5	2	2	1
PEO6	2	2	2

1.3.4 Program Learning Outcomes (PLOs)

After completion of the degree program, students will be able to

- PLO1 understand the fundamental concept of statistical sciences
- PLO2 develop a foundation in mathematics required for statistics and data science
- PLO3 develop a strong foundation in statistical theory and methods
- PLO4 possess skills to formulate statistical models and apply advanced data science tools for analyzing data and making evidence-based policy and planning at public and private institutions.
- PLO5 demonstrate computing skills required for statistical research and analytic solutions for government, industry, and other stakeholders.
- PLO6 develop basic knowledge of related subjects and skills to interpret statistical results in interdisciplinary research, and communicate findings to researchers in other disciplines.
- PLO7 develop good oral and written communication skills required for taking up leadership roles and working in a team in their professional life with confidence.
- PLO8 develop strong moral and ethical principles and apply them to professional work for the betterment of society.

The program's learning outcomes reflect all its domains: the fundamental domain, the social domain, the thinking domain, and the personal domain.

Table 1.4: Mapping of the program learning outcomes (PLOs) with the program education objectives (PEOs)

	Program Education Objectives (PEOs)					
PLOs	PEO1	PEO2	PEO3	PEO4	PEO5	PEO6
PLO1	3	3	3	3	1	1
PLO2	3	1	1	1		
PLO3	3	3	2	2		
PLO4	3	3	2	2	2	2
PLO5	2	3	2	1	1	
PLO6	2	2	3	2	3	2
PLO7	2	2	2	2	3	3
PLO8	2	2	2	2	3	3

Note: Scale to explain the extent of matching: 3= high, 2= medium, 1=low

2 Curriculum Framework

2.1 Structure of the Program

The 4-year B.S. Honours program comprising of four academic sessions, each having a duration of 12 months. Each student has to take a total of 140 credits over four academic years. These include 105 credits of theoretical courses, 24 credits of computing courses, eight credits of oral, and three credits for a Capstone project. Out of 140 credits, 30 will be from Mathematics, Economics, and Computer Science courses. In addition, the student has to take a non-credit course on English communication skills in the first academic year.

Moreover, all the courses are divided into two parts: general education (GED) courses and core courses. The GED courses focus on the multidiciplinary topics while the core courses focus on the topics related to Statistics and Data Science, Mathematics, Statistical Computing, Computer Science, viva-voce, and a Capstone project. The breakdown of the credits by type of courses are given in Table 2.1.

Table 2.1: Distribution of 140 credits by type of courses

Type of Courses	Focused Area	Credits
GED Courses	Multidisciplinary	27
Core Courses	Statistics and Data Science	66
	Mathematics	18
	Statistical Computing	12
	Computer Science	06
	Oral	08
	Capstone Project	03
	Total	140

The courses are structured in such as way so that program learning outcomes reflect all its domains: fundamental doamin, social domain, thinking domain and personal domain.

Detail breakdown of the distribution of the courses with credit hours is given in Section 2.3. The credit is defined differently for theoretical and computing courses. For theoretical courses, one credit corresponds to 15 class hours, where each class is of 50 minutes. For computing courses, one credit corresponds to 15 class hours of 50 minutes each for lab work and 15 hours

for practice. Non-credit courses consist of 20 classes, each of which is 1.5 hours.

Table 2.2: Distribution of 12 calender months by different components of the program

Type	Duration (in weeks)
Classes	26
Preparation of final examination	4
Course final examination	6
Result publication	4
Vacations and holidays	12

2.2 Assessment System

2.2.1 Evaluation

A student's performance at each year in a given course will be evaluated by in-course examinations/assignments/performance evaluation in the class/final examinations. Evaluations are aggregated at the end of each academic year. Thirty percent marks of the theoretical and forty percent of the computing courses will be allotted for in-course examinations and attendance.

There will be two in-course examinations for each theoretical, computing, and non-credit course. Thee course teacher may decide to use short questions to evaluate students in in-course examinations. Each in-course assessment will be of one-hour duration for a theory course and the non-credit course, and the average of marks from two exams will be the final mark. However, an in-course assessment will be 1.5 hours for a computing course, and the sum of two marks will be the final mark. The theoretical course final examinations will be 4 hours for 4-credit courses and 3 hours for 3-credit courses. The duration of practical courses' final examinations will be of 4 hours.

Table 2.3: Marks (%) allocation for theoretical, computing, and non-credit courses

Theoretical	Attendance	05
	In-course exam	25
	Final exam	70
Computing	Attendance/assignment	10
	In-course exam	30
	Final exam	60
Non-credit	Attendance	10
	In-course exam	25
	Quizes (5 or 6), assignments, presentations and/or mini projects	65

The performance in the non-credit course will be considered as either Satisfactory or Non-satisfactory based on the cut off point of 40 out of 100

Table 2.4: Marks distribution for attendance

Attendance (%)	Marks (%)
90 and above	5
85 to 89	4
80 to 84	3
75 to 79	2
60 to 74	1
< 60	0

2.2.2 Grading and Grade Point

Grades and grade points will be awarded on the basis of marks obtained in the written, oral and practical examinations.

Table 2.5: Percentage score, letter grade, and grade points

Marks Obtained (%)	Grade	Grade Point
80–100	$A+$	4.00
75–79	A	3.75
70–74	$A-$	3.50
65–69	$B+$	3.25
60–64	B	3.00
55–59	$B-$	2.75
50–54	$C+$	2.50
45–49	C	2.25
40–44	D	2.00
< 40	F	0.00
	I	Incomplete
	W	Withdrawn

Only "D" or higher grades will be counted as credits earned by a student. Grade point average (GPA) will be calculated as the weighted average of the grade points obtained by a student in all the courses completed in a year. GPA will be calculated according to the following formula:

$$\text{GPA} = \frac{\sum(\text{grade points obtained in a course } \times \text{ total credit for that course})}{\text{total credits taken at a given year}}$$

CGPA = cumulative GPA for different years.

Promotion to the Next Academic Year

A student has to attend courses required for a particular year, appear at the annual examinations and score a minimum specified GPA/CGPA for promotion to the next year.

Table 2.6: Minimum required GPA/CGPA for promotion to the next academic year year

Promotion	GPA/CGPA	Grade earned
Year 1 to Year 2	2.00	D
Year 2 to Year 3	2.00	D
Year 3 to Year 4	2.00	D

2.2.3 Minimum Requirements for the Award of the B.S. Honours Degree in Applied Statistics and Data Science

1. Minimum number of required credits must be earned in the maximum period of six academic years starting from the date of 1^{st} year of admission.
2. Must have a CGPA of at least 2.5.
3. A student obtaining an "F" grade in any course (theory or computing) will not be awarded the degree. A student with an "F" grade in any course shall be allowed to retake twice/two times either within three months of publication of the $4th$ year results with special fees or with the following batches.
4. The award of the B.S. (Honours) degree will not depend on the performance on the non-credit course.

Some important policies about the examination system

1. **In-course Examination**
   (a) No make-up test will be arranged for a student who fails to appear in in-course test/tests. Absence in any in-course test will be counted as zero when calculating the average in in-course test for that course. However, a student can apply to the Director if recommended by the respective course teacher. The Director will place the application before the academic committee only if the particular student has met with an accident or her/his parents have expired or s/he has gone through a surgical procedure or any other such situation which the Academic Committee feels can be considered. The make-up test must be held during the course period.
   (b) Course teachers must announce results within four weeks of holding the examinations.
   (c) Marks for in-course assessments must be submitted by concerned course teachers to the Chairman of the Examination Committee and the Controller of Examinations before the final examination.
   (d) Questions for in-course examinations may preferably be of multiple choice (MCQ) types. Students may also be evaluated by giving short questions as decided by the course teacher.
2. **Final Examination**
   (a) The year final examinations will be conducted centrally by the Controller of Examinations as per existing rules.

(b) Attendance in the non-credit course will be added in the calculation of average attendance required to appear in final examination.

(c) Students having 75% or more attendance on average (collegiate) are eligible to appear in the final examination.

(d) Students having 60%-74% attendance are considered to be non-collegiate and will be eligible to sit for the final examination with a penalty (the amount will be fixed by the Dean, faculty of science).

(e) Students having attendance less than 60% will not be allowed to sit for the final examination but may seek readmission in the program.

(f) At the beginning of each academic session, an examination committee is to be constituted for that session by the academic committee of the institute. The Chairman of the Examination Committee will act as a course co-ordinator for that session. The examination committee will have a Chairman, two internal members and an external member.

(g) For theoretical course final examinations, there will be two examiners: the course teacher will be the first examiner and the second examiner will be a faculty from within the department or from any other department of the University of Dhaka relevant to the subject.

(h) Third Examination: Under the double-examiner system and in case of difference of more than 20% of marks, there will be a third examination. Marks of the nearest two examiners (theory and project report) will be averaged for the final mark.

3. **Time Limits for Completion of Bachelor's Degree**

A student must complete the courses of her/his studies for a B.S. (Honours) degree in a maximum period of six academic years.

4. **Improvement**

(a) If a student obtains a grade "$C+$" or lower in any theoretical course in any year, s/he will be allowed to repeat the term-final examination only once with the following batch for the purpose of grade improvement. However, s/he will not be eligible to get a grade better than "$B+$" in such a course. A student failing to improve her/his grade in a course can retain the earlier grade.

(b) Students are not allowed to repeat year final exam for practical courses for the purpose of grade improvment unless he/she obtained "F" grade. Students with only "F" grade in practical courses can give re-take exam to clear "F" for obtaining degree. However, s/he will not be eligible to get a grade better than "$B+$" in such a course.

(c) Grade improvement will not be allowed in those courses in which a student obtains a grade better than "$C+$".

(d) For the purpose of grade improvement, a student will be permitted to repeat term final examinations for a maximum of 8 (eight) credits in a specific year.

(e) A student will be allowed to repeat a maximum of 20 (twenty) credits in her/his four years B.S. Program for grade improvement.

(f) Improvement in the 4^{th} year courses: Students would be allowed to sit for improvement examination in the 4^{th} year courses with the following batch, provided they must do it before the publication of final result by the office of the Controller of Examinations or Issuance of Provisional Certificate by the Controller of Examinations.

(g) No improvement for the non-credit course will be allowed.

5. **Admission to Next Academic Year and Readmission at the Same Year**

(a) A student should take admission to the next academic year within 2 (two) months after publication date of the results of the current year.

(b) A student (if applicable) can take readmission 2 (two) times throughout the program either in the same class or in different classes. In both cases, s/he must complete the degree by 6 (six) years from the time of original admission.

(c) A student (if applicable) may seek readmission and continue studies as a regular student provided s/he has at least 30% attendance in the previous year.

(d) On readmission (if applicable), the student has to retake all courses and examinations. In case s/he does not get the opportunity to repeat the courses due to late admission, marks of in-course assessment and laboratory performance/ assessments in the previous year may be retained by the student. In this case, s/he must retain in-course marks of all previous year courses but not part of the courses.

6. **Academic Awards**

A student can earn the following awards on very successful completion of the degree.

(a) As a recognition of excellent performance, the names of the students may be included in *Dean's Honor Award* or *Dean's Merit Award* in an academic year without appearing at any improvement examination. There will be two categories of awards for graduate students:

i. *Dean's Honor Award*: students with CGPA 3.85 and above (Dean of the faculty of science may change the cutoff).

ii. *Dean's Merit Award*: students with CGPA 4.00.

(b) To be eligible for the ISRT Golden Jubilee award, a student must have a B.S. Honours degree in Applied Statistics and Data Science with the highest CGPA among the students of her/his class without sitting in any improvement examination and have shown outstanding academic merit throughout her/his work for the degree. Note that the student must have completed the degree in four years.

7. **Other General Regulations**

For any matter not covered in the above guidelines, existing rules for Integrated Honours Course of Dhaka University will be applicable.

2.3 Academic Year-wise List of Courses

Table 2.7: List of Courses for the First Academic Year

Course ID	Course Title	Credit Hour
AST101	Introduction to Applied Statistics and Data Science	4
AST102	Elements of Probability	4
MATH103	Foundations of Mathematics	3
MATH104	Differential Calculus	3
MATH105	Integral Calculus	3
MATH106	Linear Algebra	3
ECON107	Economics	4
ENG108	English Communication Skills	Non-credit
CSE130	Programming with C/C++	2
CSE131	SQL for Data Science	2
AST140	Oral I	2
Total		30

Table 2.8: List of Courses for the Second Academic Year

Course ID	Course Title	Credit Hour
AST201	Sampling Distributions	3
AST202	Actuarial Statistics	3
AST203	Statistical Inference I	4
AST204	Design and Analysis of Experiments I	3
AST205	Demography	4
AST206	Sampling Methods I	3
MATH207	Mathematical Methods	3
MATH208	Mathematical Analysis	3
AST230	R for Data Science	2
AST231	Statistical Simulation	2
AST232	Statistical Computing I	2
AST240	Oral II	2
Total		34

Table 2.9: List of Courses for the Third Academic Year

Course ID	Course Title	Credit Hour
AST301	Design and Analysis of Experiments II	4
AST302	Sampling Methods II	3
AST303	Linear Regression Analysis	4
AST304	Epidemiology	3
AST305	Lifetime Data Analysis I	3
AST306	Statistical Inference II	3
AST307	Multivariate Statistics I	3
AST308	Research Methodology	3
AST330	Stata for Data Science	2
CSE331	Python for Data Science	2
AST332	Statistical Computing II	2
AST333	Statistical Computing III	2
AST340	Oral III	2
Total		36

Table 2.10: List of Courses for the Fourth Academic Year

Course ID	Course Title	Credit Hour
AST401	Advanced Probability and Stochastic Processes	3
AST402	Statistics for AI and Machine Learning	4
AST403	Multivariate Statistics II	3
AST404	Econometric Methods	4
AST405	Lifetime Data Analysis II	3
CSE406	Big Data Analytics	3
AST407	Time Series Modelling	3
AST408	Generalized Linear Models	3
AST409	Official Statistics	3
AST430	Statistical Computing IV	2
AST431	Statistical Computing V	2
AST432	Statistical Computing VI	2
AST440	Oral IV	2
AST450	B.S. Capstone Project	3
Total		40

2.4 Sustainable Development Goals (SDGs) and B.S. Honours Program in Applied Statistics and Data Science

The Sustainable Development Goals (SDGs) were adopted by all United Nations member states in 2015 as a universal call for action to end poverty, protect the planet, and ensure that all people enjoy peace and prosperity by 2030. There are 17 key SDGs, which have been designed to bring the world to several life-changing 'zeros', including zero poverty, hunger, AIDS, and discriminations against women and girls. In Bangladesh, Applied Statistics and Data Science graduates can contribute to achieve SDGs during their future employment potentially through their acquired knowledge of Applied Statistics and Data Science. This is because the scientific knowledge of statistics is welcome by all spheres of development issues particularly in policy making, implementation, monitoring and evaluation. Therefore, it is essential to mark the SDGs indicators in the syllabus for B.S. Honors in Applied Statistics and Data Science so that pertinent course instructor(s) can emphasize on relevant topic(s) for the sake of better understanding of the issues by the learners.

Generally, SDG relevant statistics are recorded, updated, monitored, and evaluated as official statistics by different organs of the government. Principally, government agencies under different ministries are in charge of implementing relevant interventions for achieving different SDGs targets, and Bangladesh Bureau of Statistics (BBS) leads the monitoring of the progress towards meeting the targets through conducting surveys and/or using official statistics. However, as an educational institution Institute of Statistical Research and Training (ISRT) can equip Applied Statistics and Data Science graduates with important statistical and computing skills so that they can work for the government and non-government agencies and help to achieve and monitor the SDGs in future.

Many tools that can be used to compute and evaluate SDG relevant elements are being taught in different courses of B.S. Honors in Applied Statistics and Data Science program, specifically in AST101 (Elements of Applied Statistics and Data Science), ECON106 (Economics), AST202 (Actuarial Statistics), AST205 (Demography), AST206 (Sampling Methods I), AST302 (Sampling Methods II), AST304 (Epidemiology), AST305 (Lifetime Data Analysis I), AST308 (Research Methodology), AST402 (Statistics for AI and Machine Learning), AST404 (Econometric Methods), AST405 (Lifetime Data Analysis II), AST406 (Big Data Analytics), AST407 (Time Series Modelling), AST408 (Generalized Linear Models), and AST409 (Official Statistics). Many of the other theoretical courses such as probability theory (AST102), statistical inference (AST203 and AST302) statistical modeling (AST204, AST301, AST303), multivariate statistics (AST307, AST403), and mathematical courses (MATH103-MATH106, MATH207, MATH208), are also useful for preparing students with a strong foundation of statistical methods related to analyzing real life data (e.g. SDG related data).

Besides theoretical courses, the B.S. Honors in Applied Statistics and Data Science offers a number of statistical computing courses, all of which have been designed to prepare students with the expertise required for statistical analysis of data (e.g. estimation and testing, developing predictive model, etc. required for monitoring and evaluating the progress in achieving SDG targets). In addition, each student of B.S. Honors in Applied Statistics and Data Science is required to write a Capstone Project (AST450), and project works are often related to different SDG indicators. The detailed connectivity among SDG indicators and courses taught in B.S. Honors in Applied Statistics and Data Science have been portrayed in Table 2.11. The course instructors are recommended to point to the key words (Table 2.11) picked up

Table 2.11: Connections between SDGs and Courses of B.S. Honours Program in Applied Statistics and Data Science

SDGs	Keywords	Relevant Courses
SDG 1 : End poverty in all its forms everywhere	measuring poverty, zero poverty, poverty line, extreme poverty	101, 106, 206, 308, 450
SDG 2 : End hunger, achieve food security and improved nutrition and promote sustainable agriculture	prevalence of malnutrition among under five children	205, 304, 402, 408, 409
SDG 3 : Ensure healthy lives and promote well-being for all at all ages	reduce neonatal mortality, under five mortality, maternal mortality, death rate	205, 304, 402, 405, 408, 406, 409
SDG 4 : Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all ages	enrollment and dropout rate, participation rate in formal and non-formal education, rate of ICT learning	101, 206, 302, 308, 450
SDG 5 : Achieve gender equality and empower all women and girls	women empowerment, domestic violence, teen marriage	101, 308, 206, 408, 450
SDG 6 : Ensure availability and sustainable management of water and sanitation for all	access to safe drinking water, improved sanitation, hygiene practice	101, 206, 302, 308, 408, 406, 450
SDG 7 : Ensure access to affordable, reliable, sustainable and modern energy for all	access to electricity, population with primary reliance on clean fuels and technology	101, 206, 302, 308, 402, 408, 450
SDG 8 : Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all	GDP, unemployment rate, child labour	101, 106, 206, 308, 402, 406, 407, 450
SDG 10 : Reduce inequality within and among countries	income inequality, poverty line, mapping poverty	106, 206, 308, 450
SDG 13 : Take urgent action to combat climate change and its impacts	climate change, natural disaster	402, 406, 407, 450

from SDG in their teaching module(s) wherever appropriate and emphasize relevant tools for teaching the techniques to estimate, test, and evaluate the SDG indicators thereby complying with national goals in line with the targets set in SDGs.

Overall, the B.S. Honors Program in Applied Statistics and Data Science prepares students to tackle any challenges regarding statistical analysis of data, implementation of plans, monitoring and evaluation of intervention programs in real life scenarios. Therefore, Applied Statistics and Data Science graduates will be able to contribute to achieving SDGs, subject to the employment in the pertinent government organs as well as development organizations in Bangladesh.

3 Detailed Syllabus

3.1 Detailed Syllabus - First Year

AST101: Introduction to Applied Statistics and Data Science	**Credit 4**

Course Description

This course deals with fundamental concepts and methods of statistics as a whole. The impact that statistics has made and will continue to make in Data Science and virtually in all fields of scientific and human endeavors will be illustrated. A brief discussion on the introductory definitions and meaning of terms and terminologies used in Applied Statistics and Data Science will be given. The basic topics in data, variables, their measurements, storage, management and methods for analyzing them will be presented.

Course Objectives

The aim of the course is to provide a thorough theoretical grounding in data summarization and description for finding patterns in data. The primary objective is to provide an introduction to summary measures necessary for the subsequent study of core and specialized courses in statistics and data science, namely, statistical machine learning, data science, probability, regression, generalized linear models, multvariate techniques, time series, econometrics and associated computing software.

Course Learning Outcomes (CLOs)

Given the aims and objectives of the course, the course learning outcomes of the course can be summarized as illustrated here. After completion of the course, students will be able to

- CLO1 develop a deeper understanding of the underlying features of the concepts of data science and modern methods of descriptive statistics
- CLO2 equip students with founding principles and application of data summarization tool kit
- CLO3 develop foundational basis of ithe theories of applied statistics and data science so that they can apply their knowledge and skills to study other courses of the discipline
- CLO4 possess the ability to use fundamental statistical knowledge, methodologies and modern data science tools in a suitable and pertinent way
- CLO5 possess the ability to apply excel software to diverse data summarization techniques

CLO6 contextualise outputs where data are generated from diverse and evolving social, political and cultural dimensions and from diverse digital foot-prints

CLO7 possess the skill to comprehend the balance between the difficulty and correctness of the data summary measures used and the equitability of the delivery of the solution and the ability to interpret the solution in plain language

CLO8 demonstrate presentation skill and independent comprehensive ability and an ability to amalgam knowledge and compliantly apply them to characterise, analyse and solve a wide range of problems.

CLO9 understand statistical good practice and ways to evade possible abuse or misuse of statistical summary.

CLO10 to apply the principles of lifelong learning to any new challenges.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3							
CLO2		3						
CLO3			3					
CLO4				3				
CLO5					2			
CLO6				3				
CLO7			3					
CLO8							2	
CLO9								2
CLO10							2	

Contents

Introduction to data science: data science concept and its life cycle; big data and its sources; explain and demonstrate the value of data in assessing problems and supporting scientific, commercial, social, and artistic problem-solving; distinguish between different types of data that are generated in science, engineering, and design; identify the types of questions that can be asked of data in satisfaction of a particular information goal; Employ strategies for ensuring data quality. Identify aspects of data governance to judge whether and how data can be used in analyses.

Introduction to applied statistics: meaning of statistics; scopes and limitations; concepts of descriptive and inferential statistics; basic concepts: data, sources of data - primary and secondary data; population, sample, parameter, statistic; variables and types of variable: qualitative, quantitative discrete and continuous; scales of measurements; classification of variables by scales of measurements.

Producing data: approaches of producing data; the concept of experimental study and non-experimental study to produce data; introduction to sample survey and questionnaire; the

concept of electronically recorded data and hospital recorded data; the concept of data cleaning and checking before statistical analysis.

Organization and presentation of data: graphical presentation for qualitative and quantitative data; sorting data, grouping qualitative and quantitative data: construction of frequency distribution and relative frequency distribution; graphical presentation of frequency distribution histogram, frequency polygon, ogive.

Concept of distribution: location, scale (spread) and shape, illustration with stem-and-leaf diagram; descriptive measures of data; measures of location; measures of dispersion; moments and their interrelationship; measures of skewness and kurtosis; three- and five-number summary; box-plot and modified box-plot.

Description of bivariate data: bivariate frequency distribution; graphical presentation of bivariate data; contingency table; the concept of association between two variables; percentage table and interpretation of cell frequencies; measures of association for nominal and ordinal variables; measures of association for interval or ratio variables; correlation; the relationship between two variables: simple linear regression; basic issues in inferential statistics.

Data visualization and analysis using spreadsheet (e.g., excel, google sheets, etc.).

Textbooks

1. Weiss N (2007). Introductory Statistics, *7th edition*. Addison Wesley.

Reference Books

1. Mann PS (2020). Introduction to Statistics, *10th edition*. John Wiley & Sons Inc.
2. Newbold P (2004). Statistics for Business and Economics, *3rd edition*. Prentice-Hall.
3. Witten IH, Frank E, and Hall MA (2016). Data Science: An Introduction. John Wiley & Sons Inc.
4. Freeman D, Pisani R, and Purves R (2007). Statistics, *4th edition*. W.W. Norton and Company.

AST102: Elements of Probability **Credit 4**

Course Description

This is an introductory course on probability theory. This course attempts to provide basic concepts of set theory, experiment and sample space, and different approaches of defining probability. It discusses useful laws of probability, conditional probability, Bayes rule, random variables and their distributions, and functions of random valuables. It also covers discussions on certain operators like mathematical expectation and generating function with properties and applications, and thorough discussions on commonly used probability distributions such as binomial, hypergeometric, and negative binomial, Poisson, normal, exponential and gamma distributions.

Course Objectives

To provide basic concepts of sets, counting techniques, and acquaint students with necessary skills for solving probability related problems using appropriate laws. To introduce the notions of random variables. To develop ability to find probability distribution of random variables and of their functions. To introduce operators like generating functions, expectation, etc. for studying the characteristics of distributions. To make familiar with basic probability distributions with possible areas of applications. To prepare the students for learning advance courses where probability theory has a prominent role.

Course Learning Outcomes (CLOs)

Students who successfully complete this course should be able to

CLO1 understand (explain ideas and concept) Students will be able to explain basic concepts of set theory, experiment and sample space, different approaches of defining probability, and useful laws of probability.

CLO2 analyze (draw connection among ideas) Students will be able to analyze and draw connections between concepts such as conditional probability, Bayes rule, random variables, and their distributions.

CLO3 evaluate (justify a stand or decision) Students will be able to justify their decisions by evaluating the probability distributions of random variables and the use of operators like mathematical expectation and generating functions.

CLO4 apply (use information in new situation) Students will be able to apply the concepts and laws of probability to solve problems and find probability distributions of random variables in new situations.

CLO5 create (produce a new or original work) Students will be able to create new insights and original work by using generating functions and studying characteristics of various probability distributions.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2	1	2	1	1
CLO2	2	3	3	3	2	2	1	1
CLO3	2	2	3	3	2	3	1	1
CLO4	2	2	2	3	2	3	1	1
CLO5	2	2	3	3	2	2	2	1

Contents

Combinatorial analysis: basic principles of counting, permutations, combinations; axioms of probability: sample space and events, axioms of probability, sample spaces having equally likely outcomes, probability as a measure of belief; conditional probability and independence: conditional probabilities, Bayes formula, independent events.

Random variables: introduction, discrete random variables, expectation, expectation of a function of a random variable, variance, Bernoulli and binomial random variables, Poisson random variable, other discrete random variables (geometric, negative binomial, hypergeometric); expected value of a sums of random variables; properties of cumulative distribution function; continuous random variables: expectation and variance of continuous random variable, normal random variable, normal approximation to binomial distribution, exponential random variables.

Jointly distributed random variables: joint distribution functions, independent random variables, sums of independent random variables, conditional distributions (discrete and continuous cases); properties of expectation: expectation of sums of random variables, covariance, variance of sums, correlations, conditional expectation, moment generating functions, probability generating function.

Textbooks

1. Ross SM (2009). A First Course in Probability, *8th edition*. Prentice-Hall.

Reference Books

1. Stirzaker D (2003). Elementary Probability, *2nd edition*. Cambridge.
2. Bertsekas DP and Tsitsiklis JN (2008). Introduction to Probability, *2nd edition*. Athena Scientific.
3. Blitzstein JK and Hwang J (2019). Introduction to Probability, *2nd edition*. Springer.

MATH103: Foundations of Mathematics **Credit 3**

Course Description

The Foundation of Mathematics course is meticulously designed to provide a thorough exploration of three fundamental pillars in mathematics: algebra, differential equations, and discrete mathematics. This course covers the theory of numbers, theory of equations, series, formulation of simple applied problems in terms of differential equations and solving approaches, discrete mathematical structures, such as sets, graphs, and combinatorics to prepare students for developing problem-solving skills by applying mathematical concepts to real-world problems, and enhancing analytical thinking and logical reasoning skills.

Course Objectives

The objective of the course is to equip students with a solid understanding of the fundamental concepts, theories, and applications of these mathematical areas so that students can develop a strong foundation for advanced topics in statistics and computing.

Learning Outcomes

After completion of the course, students are expected to

CLO1 learn and understand the theory of numbers, differential equations and discrete mathematics in detail

CLO2 understand and apply different mathematical structures, such as sets, graphs, and combinatorics, to solve problems in computer science, cryptography, and other related fields

CLO3 develop problem-solving skills by applying mathematical concepts to real-world problems

CLO4 to enhance analytical thinking and logical reasoning skills.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	3	3					
CLO2	2	2	2	3	3			
CLO3	1	1	1	3	3	1		
CLO4			2	3	2	2		

Contents

Part A (Basic Algebra)

Theory of numbers: unique factorization theorem; congruences; Euler's phi-function; inequalities: order properties of real numbers; Weierstrass', Chebysev's and Cauchy's inequalities; inequalities involving means; complex numbers: field properties; geometric representation of complex numbers; operations of complex numbers; summation of algebraic and trigonometric finite series; theory of equations: relations between roots and coefficients; symmetric functions of roots; Descartes rule of signs; rational roots;

Beta and gamma function and their properties; incomplete beta and gamma function ; Dirichlet's theorem.

Part B (Differential Equations)

Formulation of simple applied problems in terms of differential equations; equations of the first order and their solutions; singular solutions; geometric applications; linear equations with constant coefficients; method of undetermined coefficients; variation of parameters, simple cases of linear equations with variable coefficients.

Part C (Discrete Mathematics)

Combinatorics: Counting principles. Inclusion-exclusion principle. Pigeonhole principle. Generating functions. Recurrence relations. Graphs, structure, and symmetry of graphs, trees and connectivity, Eulerian and Hamiltonian graphs and diagraphs, directed graphs, and planar graphs.

Algorithms on graphs: Introduction to graphs, paths and trees. Shortest path problems: Dijkstra'a algorithm, Floyd-Warshall algorithm and their comparisons. Spanning tree problems. Kruskal's greedy algorithm, Prim's greedy algorithm and their comparison.

**In addition to the compulsory question, students should answer at least one question from each of the part.*

Textbooks

1. Ayres F (1995). Theory and Problems of Modern Algebra. McGraw-Hill.
2. Ross SL (1980). Introduction to Ordinary Differential Equations, *4th edition.* Wiley.
3. Rosen KH (2018). Discrete Mathematics and Its Applications, *8th Edition.* McGraw Hill.

Reference Books

1. Linda Gilbert (2009). Elements of Modern Algebra, *8th Edition.* Brooks/Cole.
2. Stanley J. Farlow (2007). An Introduction to Differential Equations and Their Applications. Dover Publications.

MATH104: Differential Calculus **Credit 3**

Introduction

The course is designed to provide a comprehensive foundation in the concepts and techniques of calculus. Key topics include basic functions, their graphs, and properties, as well as curve sketching, limits, continuity, differentiation, and relative extrema. Applications of differentiation and the Taylor Series are also covered. Additionally, the course expands calculus from the two-dimensional world of single-variable functions into the three-dimensional world and beyond, addressing multivariable functions.This course aims to equip students with a robust understanding of both single-variable and multivariable calculus, preparing them for advanced mathematical and analytical challenges.

Objectives

This course aims to develop both computational proficiency and a deep geometric understanding of calculus, equipping students with essential mathematical skills for advanced studies in statistics. Emphasizing the ability to perform complex calculations accurately and visualize geometric interpretations of functions and surfaces. Students will learn to apply these techniques to real-world problems and statistical analyses, including curve fitting and differential equations.

Learning Outcomes

Upon completion of the course, students will be able to

CLO1 execute differentiation for single-variable and multivariable functions, and apply these techniques to solve real-world problems

CLO2 understand and analyze the geometric interpretations of calculus, including the behavior of functions, shapes of curves, and properties of surfaces in three-dimensional space.

CLO3 employ optimization techniques, including the use of Lagrange multipliers, to solve complex problems involving maxima and minima in various contexts.

CLO4 cultivate a logical and systematic approach to problem-solving, enhancing their ability to tackle advanced mathematical challenges in academic and professional settings.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	2	3	1	3	3			
CLO2	2	3	1	3	2			
CLO3	2	3	1	3	2			
CLO4	2	3	3	3	2			

Contents

Functions and their graphs: polynomial and rational functions, logarithmic and exponential functions, trigonometric functions and their inverses, hyperbolic functions and their inverses, composition of functions.

Limit and Continuity of Functions: Definition. Basic limit theorems, limit at infinity and infinite limits. Continuous functions. Properties of continuous functions on closed and bounded intervals.

Differentiability and related theorems: Tangent lines and rates of change. Definition of derivative. One-sided derivatives. Rules of differentiation. Successive differentiation. Leibnitz theorem. Related rates. Linear approximations and differentials. Rolle's theorem, Lagrange's and Cauchy's mean value theorems. Extrema of functions, problems involving maxima and minima. Concavity and points of inflection. Indeterminate forms. L'Hospital's rule.

Power series expansion: Taylor's theorem with general form of the remainder; Lagrange's and Cauchy's forms of the remainder. Taylor's series. Maclaurin series. 5. Applications: Physical, Biological, Social Sciences, Business and Industry.

Vector-valued functions: Introduction to Vector-Valued Functions, Calculus of Vector-Valued Functions, Tangent lines to graphs of vector-valued functions. Arc length from vector view point. Arc length parameterization.

Curvature: Unit Tangent, Normal and Binormal Vectors, Curvature of plane and space curves: Curvature from intrinsic, Cartesian, Parametric and Polar equations. Radius of curvature. Centre of curvature.

Partial Differentiation: Functions of several variables, Graphs of functions of two variables, Limits and continuity, Partial derivatives, Differentiability, linearization and differentials. The Chain rule. Partial derivatives with constrained variables, Directional Derivatives and Gradients, Tangent Planes and Normal Vectors, Extrema of functions of several variables, Lagrange multipliers. Taylor's formula for functions of two variables.

Textbooks

1. Anton H, Bivens I, and Davis S (2012). Calculus, *10th edition*. Laurie Rosatone.

Reference Books

1. Stewart J (2015). Calculus: Early Transcendentals, *9th edition*. Cengage Learning
2. Swokowski E, Olinick M, and Pence DD (1996). Calculus, *6th edition*. Brooks Cole.

MATH105: Integral Calculus **Credit 3**

Course Description

This course introduces students to Integral Calculus, focusing on the techniques of integration and their applications to solving physical problems. Key topics include finding areas under curves, lengths of curves, volumes, and surface areas using integration methods. Students will gain a comprehensive understanding of integral calculus concepts and develop the skills needed to apply these techniques to various mathematical and real-world scenarios.

Course Objectives

The course aims to equip students with a robust understanding of Integral Calculus and its applications. By the end of the course, students will be able to interpret definite integrals geometrically as areas under curves and construct them as limits of Riemann sums. They will understand differentiation and anti-differentiation as inverse operations, as outlined in the Fundamental Theorem of Calculus, and will be proficient in evaluating integrals using various techniques. Additionally, students will be able to recognize and solve different types of improper integrals, calculate areas between curves, volumes of solids of revolution, surface areas, and arc lengths using integration. They will also develop the ability to solve problems involving multiple integrations in rectangular, cylindrical, and spherical coordinate systems, and apply appropriate models.

Learning Outcomes

Upon completion of the course, students will be able to

CLO1 interpret the definite integral geometrically as the area under a curve and construct definite integrals as the limit of Riemann sums.

CLO2 select and apply appropriate models and techniques to define and evaluate line and surface integrals.

CLO3 solve problems involving multiple integrations using rectangular, cylindrical, and spherical coordinate systems.

CLO4 calculate areas between curves, volumes of solids of revolution, surface areas, and arc lengths using integration.

CLOs	Program Learning Outcomes (PLOs)							
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	1	3	2	3	1			
CLO2	2	3	2	3	3			
CLO3	2	3	2	3	3			
CLO4	1	3	1	3	3			

Contents

The anti derivatives (indefinite integral): Techniques of integration; integration by parts; integration by substitution; integration of rational functions, Integration by reduction.

Definite integration: using antiderivatives. Definite integration using Riemann sums. Fundamental theorems of calculus. Basic properties of integration.

Applications of integration: Plane areas. Solids of revolution. Volumes by cylindrical shells. Volumes by cross-sections. Arc length and surface of revolution.

Improper integrals: Improper integrals of different kinds. Gamma and Beta functions.

Integrals in polar coordinates: Area and arc length in polar coordinates.

Double Integrals: Double Integrals over Nonrectangular Regions, Double Integrals in Polar Coordinates, Surface Area; Parametric Surfaces and Applications of Double Integrals.

Triple integrals: Volume as a triple integral, Triple Integrals in Cylindrical and Spherical Coordinates, Centers of Gravity Using Multiple Integrals and Applications of Triple Integrals, Change of Variables in Multiple Integrals; Jacobians.

Topics in vector calculus: Vector Fields, Gradient, Divergence, curl and their physical meanings Line Integrals, Green's Theorem, Surface Integrals, The Divergence Theorem, Stokes' Theorem, Applications of Surface Integrals; Flux.

Textbooks

1. Anton H, Bivens I, and Davis S (2012). Calculus, 10th edition. Laurie Rosatone.

Reference Books

1. Stewart J (2015). Calculus: Early Transcendentals, *9th edition*. Cengage Learning
2. Swokowski E, Olinick M, and Pence DD (1996). Calculus, *6th edition*. Brooks Cole.

MATH106: Linear Algebra **Credit 3**

Course Description

Linear algebra can be used in pretty much any application that deals with more than one random variable at a time rather than dealing with a random vector. Particularly, linear

algebra will be heavily used in multivariate statistics course(s). Some matrix algebra will also be very convenient for potential studies in Markov chains and stochastic processes. Linear regression is a very common use of linear algebra as well.

Course Objectives

To solve systems of linear equations using various methods, including Gaussian and Gauss-Jordan elimination and inverse matrices. To perform matrix algebra, determinants, and their properties. To understand and apply properties of real vector spaces and subspaces, including linear independence and dependence. To find eigenvalues and eigenvectors and use them in applications.

Learning Outcomes

Upon completion of MATH 106, students will be able to

- CLO1 solve systems of linear equations using Gaussian and Gauss-Jordan elimination and inverse matrices
- CLO2 Perform matrix algebra, determinants, and understand their properties (Concept).
- CLO3 Understand and apply properties of real vector spaces and subspaces, including linear independence and dependence
- CLO4 Find eigenvalues and eigenvectors and apply them in various scenarios (Application).
- CLO5 Create orthogonal and orthonormal bases using the Gram-Schmidt process and solve application problems

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	1	2	3	3	3			
CLO2	3	3	2					
CLO3	3	3	3					
CLO4	2			3	2			
CLO5			2	3	3	2		

Contents

Matrices, vectors and their operations: basic definitions and different types of matrices, matrix operations (addition, multiplication), trace of a matrix, determinant and adjoint of a square matrix, properties of determinants, inverse of matrix, properties of inverse, Kronecker product and related operations.

System of linear equations: Gaussian elimination, Gauss-Jordan elimination, homogeneous linear systems, null spaces and the general solution of linear systems, rank and linear systems, generalized inverse of a matrix, generalized inverses and linear systems.

Vector spaces and subspaces: vector addition and scalar multiplication, linear spaces and subspaces, intersection and sum of subspaces, linear independence and dependence, basis and

dimension, inner product, norms and orthogonality, orthogonal projections, Gram-Schmidt orthogonalization.

Eigenvalues and eigenvectors: eigenvalue equation, characteristic polynomial and its roots, Eigenspaces and multiplicities, diagonalizable matrices, computation of eigenvalues and eigenvectors.

Singular value and Jordan decompositions: singular value decomposition, SVD and linear systems, computing the SVD, Jordan canonical form.

Quadratic forms: matrices in quadratic forms, positive and nonnegative definite matrices, congruence and Sylvester's law of inertia, nonnegative definite matrices and minors, some inequalities related to quadratic forms, simultaneous diagonalization and the generalized eigenvalue problem.

Textbooks

1. Anton H and Rorres C (2013). Elementary Linear Algebra, *11th edition*. Wiley.

Reference Books

1. Strang G (2023). Introduction to Linear Algebra, *4th edition*. Wellesley-Cambridge.
2. Banerjee S and Roy A (2014). Linear Algebra and Matrix Analysis for Statistics. Chapman and Hall/CRC

ECON107: Economics **Credit 4**

Course Description

This course provides an introduction to a broad range of economic concepts, description of economic laws and theories, and analytical reasoning behind those. It considers both microeconomics: the supply and demand-based choices made by individual units (households and firms), and macroeconomics: the broader analysis of the economy. The use of supply and demand models will be the fundamental tools and the trade-offs and choices will be considered through comparison of costs and benefits measures. Production and market structure will be analyzed at the firm level as well. Macroeconomic issues concerning the interaction of goods and services markets, labor supply and economic resources at an aggregate level (region or country level) will be discussed.

Course Objectives

The objectives of the course are to introduce students to principal microeconomic theories and demonstrate how these laws can be used to explain economic choices made by households and firms; to describe how macroeconomic concepts can be used to analyze economy as a whole; to help students learn how to interpret economic models, diagrams and tables and use them to analyze different economic phenomena; and to illustrate how government policies influence microeconomic choices and macroeconomic upshots.

Learning Outcomes

After completion of the course, students are expected to

CLO1 understand theories and principles in microeconomics including consumer theory, production cost, price theory, market structure, trade-offs, comparative advantage, factor markets, public goods, externalities and market failure.

CLO2 learn macroeconomic concepts and principles including productivity, inflation, employment and government economic policies.

CLO3 acquire quantitative skills used in economic modelling and analysis.

CLO4 understand fiscal and monetary policies of Bangladesh and how they are related to relevant SDG goals.

CLO5 demonstrate the skills to apply and explain macroeconomic and microeconomic principles to analyse different individual, firm or state-level economic issues.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3							
CLO2	3							
CLO3	2	3	2	1				
CLO4	1		2	2	3	2		1
CLO5	1		1	3	1	3	3	1

Contents

Definition and scope of economics; theory of demand and supply; demand schedule; supply schedule; equilibrium of demand and supply; elasticity of demand and supply: measurement of elasticity; price elasticity of demand and supply.

Demand and consumer behavior; utility theory; equi-marginal principle; indifference curve analysis: consumers surplus; individual and market demand; derivation of demand curve; theory of production: production function; total, average and marginal product; law of diminishing returns; factors of production; pricing of factors of production; division of labor; localization of industries; returns to scale; law of variable proportion; isoquants; Cob-Douglas and CES production function; theory of cost; fixed and variable cost; total and marginal costs; least cost rule; opportunity cost.

Market structure: perfect and imperfect competition; pricing under monopoly, oligopoly and monopolistic competition; short-run and long-run equilibrium analysis; income and wealth: factor incomes vs. personal incomes, role of government, wealth; fundamentals of wage determination, the supply of labor, determinants of supply, empirical findings, wage differentials; basic concepts of interest and capital, prices and rentals on investments, rates of return and interest rates, present value of assets, real vs. nominal interest rates.

Key concepts of macroeconomics: objectives and instruments of macroeconomics; measuring economic success, tools of macroeconomic policy; real vs. nominal GDP, "Deflating" GDP by a price index; consumption, investment, NDP, GNP, price indexes and inflation; consumption and saving: consumption function, saving function; investment: determinants of investment, revenues; theories of economic growth: four wheels of growth, human and natural resources,

capital; theories of economic growth: classical dynamics of Smith and Malthus, neoclassical growth model.

Index number: characteristics and uses, problems in the construction, classification; methods: unweighted, weighted: Laspeyre's, Paasche's, Dorbish and Bowley's, Fisher's, Marshall and Edgeworth's, Kelly's and the chain index numbers; test of accuracy, base shifting, splicing, deflating of index numbers; application of consumer price index number.

Textbooks

1. Samuelson PA and Nordhaus WD (2009). Economics, *19th edition*. McGraw Hill.

Reference Books

1. Mankiw NG (2015). Principles of Economics, *7th edition*. Cengage Learning.
2. Dowling ET (2011). Introduction to Mathematical Economics, *3rd edition*. McGraw-Hill Education.
3. Newbold P, Carlson W and Thorne B (2012). Statistics for Business and Economics, *8th edition*. Pearson.

ENG108: English Communication Skills **Non-credit**

Course Description

English is the language of our international communication in all areas. Having a good command of English helps the graduate to have more opportunities in life, first of all, their career. Strong communication in English involves four modes: reading, writing, speaking, and listening. Different people have naturally differing aptitudes for these skills.

Course Objectives

The aim of this course is to develop the student's ability to use English effectively for the purpose of practical communication. The course aims to expand vocabulary, increase proficiencyin reading, writing and listening and develop greater understanding of grammatical rules and usage. It therefore has four components: Writing, Reading, Listening and Speaking.

Learning Outcomes

After completion of this course students are expected to learn the followings:

CLO1 Reading Skills: Develop critical reading abilities to identify, retrieve, and understand relevant information and implied meanings.

CLO2 Effective Writing: Communicate clearly and accurately, summarizing, paraphrasing, and applying logical reasoning with proper grammar and vocabulary.

CLO3 Enhanced Listening: Improve the ability to identify and understand facts, ideas, opinions, and connections, employing good listening practices.

CLO4 Clear Speaking: Convey information and opinions effectively, using appropriate grammar, vocabulary, and pronunciation.

CLOs	Program Learning Outcomes (PLOs)							
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1						3	3	1
CLO2						3	3	1
CLO3						3	3	1
CLO4						3	3	1

Contents

Reading: Reading comprehension and application of reading strategies; Skimming and scanning texts to identify and retrieve facts, details, important points and themes; Recognizing ideas, opinions and attitudes in a range of texts such as letters, brochures, forms, imaginative writing, reports, academic papers and creative essays; Developing critical reading skills (e.g., paying attention to different perspectives exposed in texts); Extending reading opportunities to build vocabulary; Learning differences between formal and informal writing.

Writing: Structure of a paragraph, writing process of narrative and descriptive paragraphs; Learning how to summarize, paraphrase, synthesize and apply logical reasoning; Sentence variety: simple, compound and complex sentences; Getting familiar with/revising basic grammar through literary texts, newspaper articles, short excerpts.; Common sentence errors: Subject/verb agreement, incorrect verb forms, shift in tense, shift in point of view, unclear or missing referent, lack of pronoun agreement, etc.; Punctuation errors; Academic referencing; Carrying out simple writing tasks (e.g., writing essays on a given topic, writing emails/letters, etc.) in an appropriate and accurate form of English.

Listening: Understanding simple information presented in a variety of forms, e.g. news, weather, dialogue, interviews and telephone conversations; Learning to develop and apply basic skills of a good listener (e.g. grasping the main idea of a text, taking notes); Identifying and retrieving some facts from materials e.g. a formal talk, monologues, lectures, etc.; Understanding what is implied but not actually spoken e.g., gist, purpose and intention;

Speaking: Basics of everyday communication; Carrying out a range of speaking activities such as introducing oneself, participating in short debates, engaging in conversation in different environments e.g., while shopping, meeting friends, travelling, visiting a doctor, answering the telephone, etc.; Conducting a sustained conversation with a sense of audience and purpose; Effective use of communication strategies to share knowledge or participate in group/pair discussions; Learning and practicing basic pronunciation skills (IPA symbols, diphthongs, monophthongs, long and short vowels); Developing presentation skills.

Textbooks

1. Bailey S (2003). Academic Writing: A Handbook for International Students. Routledge.
2. Soars L and Soars J (2003). Headway Intermediate. Oxford University Press.

Reference Books

1. Langan J (1984). College Writing Skills with Readings. Connect.

2. Roach P (2009). English Phonetics and Phonology: A Practical Course, *4th edition*. Cambridge University Press.
3. O'Connor JD (1980). Better English Pronunciation. Cambridge University Press.
4. Cambridge IELTS

CSE130: Programming with C/C++ **Credit 2**

Course Description

Programming with C/C++ course mainly focuses on the programming language C. The course introduces machine level language, origin of C and applications of C programming in statistics.

Course Objectives

This module covers introductory ideas to connect C programming with statistics through programming code execution. To understand how to analyze descriptive statistics and data analysis using C programming.

Learning Outcomes

After completion of the course, students are expected to

CLO1 be introduced with the C programming language.

CLO2 know techniques that facilitate the development of structured computer programs, such as algorithms, pseudocode, flowchart, etc. Understand how to control C programs, write C functions, use arrays.

CLO3 how to connect C programming with statistics through programming code execution.

CLO4 understand how to analyze descriptive statistics and data analysis using C programming.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1					3			
CLO2					3			
CLO3					2	1		
CLO4					3	1		

Contents

Introduction to programming: algorithm, flowchart, code (program); levels of programming: machine level, assembly level and high level language; execution of code: translator, compiler, interpreter, assembler; steps of execution: compilation, link, run.

An overview of C: the origins of the C language, compilers versus interpreters; variables, constants, operators, and expressions: data types, declaration of variables, assignment statements, constants, operators, expressions; program control statements: C statements, conditional statements, loop statements, labels; functions: the return statement, function arguments, arguments to main(), returning pointers, pointers to functions; arrays: single-dimension arrays, passing single dimension arrays to functions, two- and multi-dimensional arrays, arrays and pointers, allocated arrays, array initialization.

Applications of C programming in data analysis: frequency distributions, data summary, e.g. mean, median, maximum, minimum, matrix operations, calculation of different rates, fitting simple linear regression and sorting a vector.

Textbooks

1. Dietel PJ and Deitel HM (2010). C How to Program, *7th edition*. Pearson.

Reference Books

1. Perry G and Miller D (2014). C Programming Absolute Beginner's Guide, *3rd Edition*. Que.

CSE131: SQL for Data Science **Credit 2**

Course Description

In today's data-driven world, business enterprises and industries can access more data than ever. However, they need skilled professionals who can manage and analyze data using various tools and techniques to make sense of all this data. Structured Query Language (SQL) is a widely used tool in data science for managing and querying data stored in relational databases. This course is designed to introduce students to SQL and its application in data science.

Course Objectives

To teach students the basics of SQL through hands-on exercises and real-world examples. To cover advanced topics such as SQL joins, subqueries, and aggregation functions. To provide students with a comprehensive understanding of SQL and its pivotal role in data science.

Learning Outcomes

Upon completion of CSE-131, students will be able to

CLO1 understand the basic concepts of SQL, such as tables, queries, and basic SQL syntax (Concept).

CLO2 retrieve and manipulate data from databases using SQL queries, including filtering, sorting, and joining data from multiple tables (Application).

CLO3 perform basic data analysis using SQL, including calculating summary statistics and aggregating data (Compute, Application).

CLO4 understand how SQL fits into the data science workflow (Concept).

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2	1	2			
CLO2	1		2	3	3			
CLO3				3	3	2		
CLO4	3							

Contents

Introduction to databases, SQL and MySQL, creating a MySQL database, MySQL Data types, table creation, Populating and modifying tables, When Good Statements Go Bad, query mechanics.

The select, from, where, group by, order by, and having clauses.

SQL joins, subqueries, common table expressions, aggregation functions, window functions, date and time functions, and user-defined functions (UDFs).

Calculating summary statistics, univariate and bivariate frequency distributions, correlation, and regression analysis.

Textbooks

1. Beaulieu A. (2020). Learning SQL: Generate, manipulate, and retrieve data, *3rd edition*. O'Reilly.

Reference Books

1. Teate RM (2021). SQL for data scientists: A beginner's guide for building datasets for analysis. Wiley.

AST140: Oral I **Credit 2**

Each student must be examined by a committee of selected members at the end of the academic year.

3.2 Detailed Syllabus - Second Year

AST201: Sampling Distributions **Credit 3**

Course Description

This course provides a comprehensive introduction to sampling distributions-a fundamental concept in statistical theory that underpins the process of making inferences and drawing conclusions from sample data. Sampling distributions play a critical role in statistical analysis by providing a bridge between sample data and population parameters. Contents will involve deriving the theoretical properties of different distributions, examining the inter-relationship between them and focusing on their applications in statistical inference.

Course Objectives

This course's main objective is to acquaint students with the main concepts and uses of probability and sampling distributions. More specific objectives include demonstrating the derivation of sampling distributions of sample statistics, describing properties of common sampling distributions and illustrating the role and use of sampling distributions in inferential statistics.

Learning Outcomes

After completion of the course, the students are expected to

- CLO1 learn about different generating function techniques and their role in mathematical derivation of distributions of functions of random variables
- CLO2 understand the theoretical and practical concepts of probability and sampling distributions
- CLO3 gain insights into the central limit theorem and its implications
- CLO4 learn about mathematical properties of common sampling distributions and inter-relationships
- CLO5 understand the application of sampling distribution in statistical inference.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	3	3					
CLO2	3	1	3					
CLO3			2					
CLO4	2	3	2					
CLO5				2	2	3	2	2

Contents

Generating function techniques: moment generating function, cumulant generating function, probability generating function, characteristic function; finding distributions of functions of random variables: change of variable technique, distribution function technique, and moment-generating function technique; probability integral transformation; statistic and sampling distribution; law of large numbers; central limit theorem; exact distribution of sample mean; chi-square distribution and its properties; F-distribution and its properties; t-distribution and its properties; non-central chi-square, F and t distributions: definition and derivation; concept of order statistics, distributions of single order statistics, and joint distribution of two or more order statistics.

A brief review of some probability distributions and their properties: uniform, normal, exponential, gamma, beta, log-normal, Cauchy; definition of truncated distribution; definition of compound and mixture distribution; family of distributions: Pearsonian distribution.

Textbooks

1. Robinson EA (2011). Probability Theory and Applications. Springer.

Reference Books

1. Zehna PW (1970). Probability Distributions and Statistics. Allyn and Bacon.
2. Jones O, Maillardet R and Robinson A (2009). Introduction to Scientific Programming and Simulation using R. Chapman & Hall/CRC.
3. Arnold BC, Balakrishnan N and Nagaraja HN (2008). A First Course in Order Statistics. Society for Industrial and Applied Mathematics.

AST202: Actuarial Statistics **Credit 3**

Course Description

This course aims at providing an introduction to principles of actuarial science and mathematics involved in it. Topics include role of insurance in economy, different instruments involved in actuarial computations including interest and discount rates, pricing and valuation of actuarial products including annuities, amortization schedules, sinking funds, life assurances and premiums, among others.

Course Objectives

The course will enable understanding of the fundamental concepts of actuarial science and appreciate its role in economy. It will build a concrete knowledge of underlying theories behind actuarial computations. Finally, it will help to understand how statistical and mathematical models are used in pricing and valuing actuarial products and their real life applications.

Learning Outcomes

Upon completion of this course the students should be able to

CLO1 understand the fundamental concepts of actuarial science and appreciate its role in economy,

CLO2 obtain concrete knowledge of underlying theories behind actuarial computations.

CLO3 understand how statistical and mathematical models are used in pricing and valuing actuarial products and their applications in various real-life actuarial problems.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	3	3					
CLO2	2	3	1	2	3			
CLO3	2	3	1	3	3	3	1	1

Contents

The meaning of actuarial science; role of insurance in the economy; role of an actuary.

Fundamentals of theory of interest: definition of simple interest and compound interest and their comparisons; accumulated value factors and present value factors; effective and nominal rates of interest and their interrelationship; effective and nominal rates of discount; relation between interest and discount; equations of value and use of the time diagram in solutions of problems in interest; problems involving unknown length of investment and unknown rate of interest; annuity; different types of annuities certain; present and accumulated values of immediate annuity and annuity due; present value of deferred annuities and variable annuities; capital redemption policies; amortization schedules and sinking funds.

Actuarial mathematics: discrete life annuity and its applications; present values of different life annuities; life assurance; present values of various life assurances in terms of commutation functions; related problems; premiums; different types of premiums; net premiums; office premiums; prospective policy values.

The basic deterministic model: cash flows; an analogy with currencies; discount functions; calculating the discount function; interest and discount rates; constant interest; values and actuarial equivalence; regular pattern cash flows; balances and reserves; basic concepts; relationship between balances and reserves.

Stochastic interest-rate models: stochastic interest-rate models I; basic model for one stochastic interest rate; independent interest rates; stochastic interest-rate models II; dependent annual interest rates; modelling the force of interest; what one can do with these models.

Textbooks

1. Kellison SG (1991). The Theory of Interest, *2nd edition.* McGraw-Hill/Irwin.
2. Promislow SD (2011). Fundamentals of Actuarial Mathematics, *2nd edition.* John Wiley & Sons.

Reference Books

1. Shailaja R. Deshmukh (2010). Actuarial Statistics: An Introduction Using R, *3rd edition.* Orient Blackswan.

AST203: Statistical Inference I **Credit 4**

Course Description

This course deals with fundamental concepts and techniques of statistical inference including estimation and tests of simple and composite hypotheses. A brief revision will also be given of some basic topics in probability theory as well as random variables.

Course Objectives

The aim of the course is to provide a thorough theoretical grounding in statistical inference. The primary objective is to provide an introduction to mathematical statistics necessary for the subsequent study of specialized courses in statistics, bio-statistics, actuarial science and econometrics. The impact that statistics has made and will continue to make in virtually all fields of scientific and other human endeavors is considered.

Learning Outcomes

After completion of the course, students are expected to

CLO1 know the most common parametric and nonparametric distributions

CLO2 explore the concepts of the likelihood principle, the maximum likelihood, least square and the method of moments

CLO3 develop the skill to handle a parametric hypothesis testing problem

CLO4 illustrate inferential theory and implement the asymptotical concepts in practices

CLO5 improve the computing ability using confidence interval and goodness of fit notions

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2	1			
CLO2	2	2	3		2	1		
CLO3	3	1	2	2	3	2		
CLO4	2	2	3	2	3	2	1	1
CLO5				3	3	3	2	2

Contents

Basic Concepts: Fundamental ideas of statistical inference; parametric and non-parametric inference; estimators, statistics, parameters, theory and reality; sampling distributions and uses in inference; Estimation: Estimation of parameters and fitting of probability distributions; parameter estimation: method of moments, method of least squares, method of maximum likelihood - properties of maximum likelihood estimators, Bayes estimator; properties of a good point estimator; method of evaluating estimators: mean squared error; best unbiased

estimator - efficiency and the Cramer-Rao lower bound; sufficiency - sufficient statistics, properties of sufficient statistics, exponential family and factorization theorem, the Rao-Blackwell theorem, minimal sufficient statistics, complete statistics.

Interval estimation: introduction; method of constructing confidence interval - pivotal quantity, exact and approximate confidence interval; large- and small-sample confidence intervals; selecting the sample size; simultaneous confidence region.

Test of Hypothesis: Elements of statistical hypothesis test; approaches to hypothesis testing: Neyman-Pearson approach, Fisher approach, and Jeffreys' approach; common tests based on normal distribution: one-sample settings, two-sample (independent and paired) settings, more than two sample settings; the duality of confidence intervals and hypothesis tests; exact and large sample test; evaluating statistical test procedure - the power of tests, optimal test - Neyman-Pearson lemma; most powerful tests; composite hypotheses; generalized likelihood ratio test; uniformly most powerful tests; unbiased tests; goodness-of-fit tests - probability plots, test for normality, chi-square goodness of fit test, Kolmogorov-Smirnov test; statistical tests applied to categorical data problems: introduction, Fisher's exact test, the chi-square test of homogeneity, chi-square test of independence.

Textbooks

1. Hogg RV, Tanis EA and Zimmerman DL (2015). Probability and Statistical Inference, *9th edition*. Pearson.

Reference Books

1. Casella G and Berger RL (2001). Statistical Inference, *2*nd edition. Cengage Learning.
2. Hogg RV, McKean J and Craig AT (2010). Introduction to Mathematical Statistics, *7th edition*. Pearson.
3. Mood AF, Graybill FA. and Boes DC (1973). Introduction to the Theory of Statistics, *3rd edition*. McGraw-Hill.

AST204: Design and Analysis of Experiments I | **Credit 3**

Course Description

The design of an experiment involves considering the set of treatments selected for comparison, specifying the units to which the treatments are to be applied, determining the rules for allocating treatments to units, and specifying the measurements to be made on each unit. Various designs are now standard in a wide range of practical situations in agriculture, medicine, engineering, and other fields. This course covers methodological and practical issues in the design and analysis of experiments. Topics covered include an introduction to the design of experiments, completely randomized design, randomized complete block design, Latin square design, and balanced incomplete block design.

Course Objectives

The course aims to help students understand the methodological issues of experimental design. Students will develop the skills to write hypotheses that can be tested using experiments and

to analyze and interpret data obtained from experiments.

Learning Outcomes

After completing the course, students are expected to

CLO1 identify the response variable, factor(s) of interest, and nuisance factors, and choose the most suitable experimental design for the given problem.

CLO2 set up the relevant null and alternative hypotheses, obtain an ANOVA table, and construct appropriate (orthogonal) contrasts.

CLO3 carry out multiple comparisons, check model assumptions graphically and formally, and apply a suitable transformation when needed.

CLO4 determine the appropriate sample size, analyze data, and state conclusions in the language of the problem posed.

CLO5 demonstrate skills in designing and implementing a standard experimental design and explaining it to different stakeholders in various sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	2	1				
CLO2	2	2	3	2				
CLO3	2	2	3	1				
CLO4	1	1	2	2	3			
CLO5				3	1	3	3	

Contents

Introduction to design of experiments: strategy of experimentation; some typical examples of experimental design; basic principles; guidelines for designing experiments.

Experiments with a single factor: the analysis of variance; analysis of fixed effects model; estimation of model parameters; unbalanced data; model adequacy checking; regression model, comparisons among treatment means, graphical comparisons of means, contrasts, orthogonal contrasts, multiple testing, Scheffe's method, comparing pairs of treatment means, comparing treatment means with a control; Determining sample size; operating characteristic curve, specifying standard deviation increase, confidence interval estimation method; discovering dispersion effects; regression approach to analysis of variance; least squares estimation of the model parameters, general regression significance test.

Randomized blocks, Latin squares, and related designs: the randomized complete block designs (RCBD); statistical analysis of RCBD, model adequacy checking; estimating model parameters; Latin square design; Graeco-Latin square design; balanced incomplete block design (BIBD); statistical analysis of BIBD; least squares estimation of BIBD; recovery of intra-block information in the BIBD.

Textbook

1. Montgomery DC (2019). Design and Analysis of Experiments, *10th edition*. Wiley.

Reference Books

1. Dean AM, Voss AM, and Draguljić (2017). Design and Analysis of Experiments, *2nd edition*. Springer.
2. Bailey R (2008). Design of Comparative Experiments. Cambridge.

AST205: Demography **Credit 4**

Course Description

Understanding a nation's population patterns, socioeconomic growth, and public health all depend heavily on its demographic data. To analyze this data, this course defines demographic life events and describes statistical analysis methodologies.

Course Objectives

The course is intended to introduce the basic ideas of demography and state the importance of demographic studies. It helps to understand well-known demographic theories and different measures of demographic events. Students will perceive and analyze current demographic situation of Bangladesh.

Learning Outcomes

Upon successful completion of the course, the students are expected to

CLO1 know demographic data sources and collection procedures

CLO2 measure different fertility, mortality and nuptiality statistics

CLO3 learn migration statistics and effect on population growth migration statistics and effect on population growth

CLO4 compute and interpret different types of fertility data and use them in real life scenarios

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3		3					
CLO2	3	2	1		3			1
CLO3		2	3	2	2			1
CLO4	1	2	2	1	1			1

Contents

The basic concept of demography; role and importance of demographic/population studies; sources of demographic data: census, vital registration system, sample surveys, population registers, and other sources, especially in Bangladesh.

Errors in demographic data: types of errors and methods of testing the accuracy of demographic data; quality checking and adjustment of population data; post-enumeration check (PEC) and detection of errors and deficiencies in data and the needed adjustments and corrections.

Fertility: basic measures of fertility, crude birth rate, age-specific fertility rates (ASFR), the general fertility rate (GFR), the total fertility rate (TFR), the gross reproduction rate (GRR), and net reproduction rate (NRR), child-woman ratio; the concept of fecundity and its relationship with fertility. Fertility trends and patterns in Bangladesh.

Demographic theory: transition theory and the present situation in Bangladesh; Malthus' theory and its criticism. Mortality: basic measures of mortality: crude death rate (CDR), age-specific death rates (ASDR), infant mortality rate, child mortality rate, neonatal mortality rate; standardized death rate its need and use; direct and indirect standardization of rates; commonly used ratios: sex ratio, child-woman ratio, dependency ratio, and density of population. Mortality trends and patterns in Bangladesh.

Nuptiality: marriage, types of marriage, age of marriage, age at marriage and its effect on fertility, celibacy, widowhood, divorce, and separation, their effect on fertility and population growth.

Migration: definition, internal and international migration; sources of migration data; factors affecting both internal and international migration, laws of migration; the impact of migration on origin and destination, its effect on population growth, age and sex structure, labor supply, employment and unemployment, wage levels, and other socio-economic effects; migration of Bangladeshis abroad and its impact on the overall economic development of the country.

Graduation of data: meaning and its need, techniques of graduation, graduation of age distribution; life table: its concept, structure, and calculation; complete life table (life table by single year of age) and abridged life table, multiple decrement life tables, working life table, different life table functions and inter-relationships among them, use of life table, etc. Model life tables, Coale and Demeny regional model life tables.

Stable and stationary populations.

Textbooks

1. Siegel JS and Swanson DA (2004). The Methods and Materials of Demography, *2nd edition.* Emerald.
2. Shryock HS, Siegel JS and Larmon EA (1975). The Methods and Materials of Demography, volume I and II. U.S. Department of Commerce Publication.

Reference Books

1. Poston Jri. DL, and Bouvier LF (2016). Population and Society: An Introduction to Demography, *2nd edition.* Cambridge University Press.

AST206: Sampling Methods I **Credit 3**

Course Description

Sampling refers to the statistical techniques used in survey research in many areas such as business, marketing, environmental science, political polls, and many areas of social science and public health. This course introduces different sampling techniques and approaches to designing a sample survey and methods for analyzing survey data from both a theoretical and applied perspective. This course covers topics related to general procedures in sample surveys, some basic sampling techniques, sample size determination, and methods of estimation of parameters of interest under each sampling scheme.

Course Objectives

The objectives of the course are to make students familiar with different sampling techniques, sample size determination, and estimating the parameters of interest and their use in practice, and to prepare students to demonstrate their skills for designing all basic sample surveys, analyzing the data, and explain them to various stakeholders and other researchers in diverse areas.

Learning Outcomes

After completion of the course, students are expected to

- CLO1 understand the concept of a sample survey and the general steps of conducting a survey
- CLO2 know in detail about different sampling techniques and their use in practice to design a survey
- CLO3 understand the methods of estimating the parameters of interest under different sampling schemes.
- CLO4 develop skills to determine the sample size, calculate sampling weight, and apply them in practice
- CLO5 demonstrate skills for designing and implementing a standard sample survey and explaining it to different stakeholders in many sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2	1				
CLO2	2	1	3	2				
CLO3	1	1	3	1	2			
CLO4	1		2	2	2			1
CLO5	1		1	3	1	3	3	1

Contents

Introduction: concept of sampling, and definition of related terms; role of sampling theory, requirements of a good sampling design, steps in a sample survey, probability and nonprobability sampling, selection (draw-to-draw) and inclusion probability, sampling weight, with and without replacement sampling, characteristics of estimate: bias, mean square error and variance (precision), errors in sample survey and census, sample size determination: basics and complex scenarios.

Simple random sampling (SRS): sample selection, estimation: mean, total, proportion, ratio of two quantities, unbiasedness and variances/standard errors (SEs) of the estimators, estimators of the SEs, confidence interval (normal approximation); finite population correction, estimation over subpopulation, computation: inclusion probabilities and sampling weights.

Systematic sampling: motivation, use and challenges, sample selection, different estimators and their unbiasedness and variances, estimator of the variances, comparison with SRS, sampling from population with linear trend or periodic variation.

Stratified random sampling: concept, reasoning and needs in heterogeneous population, number and formation of strata, sample selection, estimators (total, mean, proportion), variances of the estimators, estimators for the variances, different allocation techniques, comparison with SRS, design effect and its uses, poststratification, quota sampling.

Auxiliary information in estimation: ratio estimators (total, mean), different properties: unbiasedness, variance (approximate), estimated variance, confidence interval, comparison with mean per unit estimates, conditions for best linear unbiased ratio estimator, application in stratified sampling, unbiased ratio-type estimates; product estimator; regression estimator: linear regression estimate and its properties (unbiasedness, variance and estimated variance) under preassigned b and estimated b, comparison with mean per unit estimate, application in stratified sampling, relative merits and demerits.

Cluster sampling: motivation and reasoning, formation and size of clusters; cluster sampling with equal sized clusters: estimators and their various properties (unbiasedness, variance and estimated variance), comparison with SRS and systematic sampling, optimum cluster size, stratification in cluster sampling: estimation and comparison with simpler sampling designs.

Special sampling designs: capture-recapture sampling: implementation, Peterson and Chapman estimators for population size and their variances, Hypergeometric and Multinomial models for estimating population abundance; ranked set sampling: sample selection and estimation.

Textbooks

1. Thompson SK (2002). Smapling. *2nd edition.* Wiley.
2. Cochran WG (1977). Sampling Techniques, *3rd edition.* Wiley.

Reference Books

1. Lohr SL (1998). Sampling: Design and Analysis. Duxbury.
2. Levy PS and and Lemeshow S (2008). Sampling of Populations: Methods and Applications, *4th edition.* Wiley.
3. Rao PSRS (2000). Sampling Methodologies with Applications, *1st edition.* Chapman &

Hall/CRC.

4. Thompson SK (2012). Sampling, *3rd edition*. Wiley.

MATH207: Mathematical Methods **Credit 3**

Course Description

The objectives of the course are to develop an understanding of the core ideas and concepts of numerical methods, to enable students to apply rigorous analytic and highly numerate approaches, and to develop scientific computer programs and to equip students with the ability to apply Fourier series and Fourier integrals to significant applied problems, particularly in telecommunication, which are closely related to Laplace transform.

Course Objectives

The objectives of the course are: To develop an understanding of the core ideas and concepts of numerical methods. To enable students to apply rigorous analytic and highly numerate approaches, and to develop scientific computer programs. To equip students with the ability to apply Fourier series and Fourier integrals to significant applied problems, particularly in telecommunication, which are closely related to Laplace transform.

Learning Outcomes

Upon completion of the course, students will be able to

CLO1 Understand and explain the theoretical foundations and concepts of numerical methods, including interpolation, numerical integration, and solutions of algebraic and transcendental equations

CLO2 Analyze and solve real-life problems using Fourier series and Laplace transform techniques, and draw connections among these ideas

CLO3 Evaluate the effectiveness of different numerical methods and justify the use of specific techniques for solving mathematical problems

CLO4 Apply rigorous analytical approaches and develop scientific computer programs to implement numerical methods

CLO5 Create and present solutions to complex applied problems using Fourier and Laplace transforms, and demonstrate their application in telecommunication and other fields

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2						1
CLO2	3	2	2	2	2			1
CLO3	2	2	3	3	2			1
CLO4	2	2	2	3	2	1		1
CLO5	2	3	3	3	2	2		1

Contents

Interpolation and inverse interpolation: uses of Newton's forward and backward interpolation formula; Lagrange's formula; numerical integration: Simpson's rule; Weddle's rule; trapezoidal rule; Gauss's quadratic formulae and proper examples from the applications to econometrics, meteorology and biomedicine; Euler's formula of summation and quadrature.

Solution of numerical algebraic and transcendental equations; equations in one unknown; finding approximate values of the roots; finding roots by repeated application of location theorem; method of interpolation or of false position; solution by repeated plotting on a large scale; Newton-Raphson method; Newton-Raphson method for simultaneous equations.

Fourier series: periodic function; Fourier series process of determining the Fourier coefficients; Dirichlet conditions; odd and even functions; half range Fourier sine or cosine series; Parseval's identity; differentiation and integration of Fourier series.

Laplace transform: introduction; definition of integral transformation; definition of Laplace transform; Laplace transform of some elementary functions; sufficient conditions for the existence of Laplace transform; some important properties of Laplace transform; initial and final value theorem; Laplace transforms of some special functions.

Inverse Laplace transform: definition of inverse Laplace transform; Lerch's theorem; some important properties of the inverse Laplace transform; partial function decompositions; definition of convolution; convolution theorem; Heaviside's expansion formula; evaluation of integrals; application of Laplace transform.

Introduction to Taylor's and Laurent series.

Textbooks

1. Burden RJ and Faires JD (2010). Numerical Analysis, *9th edition*. Brooks Cole.

Reference Books

1. Kreyszig E (2011). Advanced Engineering Mathematics, *10th edition*. Wiley.
2. Jeffrey A (2001). Advanced Engineering Mathematics. Academic Press.

MATH208: Mathematical Analysis **Credit 3**

Course Description

Graduate students in applied statistics need to know mathematical analysis. Real numbers, sets, intervals, sequences and series, limits, differentiability, continuity, integration, measure theory, and probability are all covered in this subject. Prerequisites for this study include a basic mathematical science course called calculus.

Course Objectives

The course has great value for a student who wishes to go beyond the routine techniques to solve standard problems, and who wants to extend ideas to a new context. It develops the ability to analyze mathematical situations properly and precisely. It also helps the students to learn how to work comfortably with concepts that initially seemed so mysterious.

Learning Outcomes

Upon successful completion of the course, the students are expected to

CLO1 Develop the ability to analyze mathematical situations properly and precisely.

CLO2 Learn to work comfortably with concepts that initially seemed so mysterious.

CLO3 Know real numbers, sets, intervals; sequences and series of real numbers and their convergence and divergence; Metric space and topological space and their properties; different properties of limits, continuity, differentiability and Rieman integral of a function.

CLO4 Basic concepts of measure theory and probability measure and to relate them in real data scenarios.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	3	2		3			
CLO2	3	3	1	1	3			
CLO3	1	3	1	1	3			1
CLO4	3	3	3	1	1			2

Contents

The real number system; axioms and completeness and its consequences; Dedekind cut, sets, compact sets; simple operation on them.

Sequence of functions of one and several variables; limit; continuity; continuous functions; uniform continuity; differentiation and integration; infinite series of constants and functions; convergence and divergence; power series: differentiation and integration of power series; Taylor expansion with remainder or in infinite series.

Metric and topological spaces; limit points; open and closed sets; interior and exterior points; boundary points; continuous mapping and Cauchy sequences.

Measure and integrals on abstract sets on real lines; Cramer measurability: fundamental definitions; auxiliary lemma; fundamental theorems; measurable functions; Lebesgue measure on a real line, plane; integrals; Riemann-Steiljes integrals.

Textbook

1. Rudin W (1976). Principles of Mathematical Analysis. McGraw-Hill.

Reference Books

1. Pugh CC (2015). Real Mathematical Analysis. Springer.
2. Rudin W (2017). Principles of Mathematical Analysis, *3rd edition*. McGraw-Hill.

AST230: R for Data Science **Credit 2**

Course Description

Any scientific task without the knowledge of software is difficult to imagine and complete in the current scenario. R is a free software that is capable of handling mathematical and statistical manipulations. It has its own programming language as well as built-in functions to perform any specialized task. Students intend to learn the basics of R software in this course.

Course Objectives

The objective of the course are to teach the basics of R programming language that includes understanding different types of R objects, writing functions, reading/writing data of different formats, basic programming tools, and creating different types of plots.

Learning Outcomes

After completion of the course students are expected to

CLO1 be able to learn the basics of R programming language that includes understanding different types of R objects

CLO2 understand the concepts of R functions, reading/writing data of different formats

CLO3 able to use different data analysis tools, such as joining two data tables, subsetting from a data table, obtaining summary statistics, etc.

CLO4 to create different types of plots using R functions

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	1				2			1
CLO2	1			3	2			1
CLO3	1			3	2	1	1	1
CLO4	2			3	3	1	1	1

Contents

Introduction to R: History and overview of R programming language, R objects, data structure (e.g. lists, data frames, etc.), reading and writing data files, subsetting R objects, vectorized operations, control structures, functions (both in-built and custom), simulation, and calling C function from R.

Exploratory data analysis with R: managing data with different tidyverse packages (e.g. dplyr, ggplot2, etc.), exploratory graphs (grammar of graphics), and generating summary statistics.

Application of R in optimizing non-linear functions using Newton-Raphson iterative procedure, numerical integration and differentiation.

Textbooks

1. Wickham H and Grolemund G (2017). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly.

Reference Books

1. Wickham H (2019). Advanced R, *Second Edition*. Chapman & Hall/CRC.

AST231: Statistical Simulation **Credit 2**

Course Description

This course is designed to introduce computational methods and simulation techniques which that will be helpful for understanding theoretical concepts and practical applications of probability and sampling distributions, and statistical inference including point and interval estimation and test of hypotheses. R programming is used to demonstrate related simulations and computations.

Course Objectives

The main objective of this course is to introduce students to the theory of simulation and illustrate to them the theoretical concepts of sampling distributions and statistical inference using computer-simulated data. The course also aims at introducing students to computational methods and statistical software (R programming) required for fitting probability distributions, conducting point and interval estimation, and performing test of hypotheses using real-life datasets.

Learning Outcomes

Upon completion of the course, students are expected to

CLO1 acquire detailed theoretical and practical knowledge of common techniques of simulating data

CLO2 write computer (R programming) codes for simulating data from different probability distributions and different resampling techniques.

CLO3 use simulated data to demonstrate theoretical concepts related to sampling distributions and statistical inference.

CLO4 apply computational methods required for fitting probability distributions, conducting point and interval estimation, and performing test of hypothesis using real-life data.

CLO5 demonstrate skills to provide data-drive solution to inferential problems, interpret findings from computational analysis and explain them to researchers from various disciplines and to practitioners working industries

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2					
CLO2	1	1	2	2	3			
CLO3	1	1	2	2	3			
CLO4	1		2	2	3	3		1
CLO5	1		1	2	2	3	3	1

Contents

Concepts of simulation and its uses in statistics; random number generations: congruential generators, seeding; random variate generations: inversion method (direct method), rejection method (indirect method); Monte-Carlo integration: hit-and-miss method, improved Monte-Carlo integration; variance reduction: antithetic sampling, importance sampling, control variates. Use of simulated, repeated sampling and real datasets for illustrating theoretical and practical aspects of fitting probability distributions, sampling distributions, point and interval estimation, and tests of hypotheses.

Textbooks

1. Jones O, Maillardet R and Robinson A (2009). Introduction to Scientific Programming and Simulation using R. Chapman & Hall/CRC.

Reference Books

1. Robinson EA (2011). Probability Theory and Applications. Springer.

AST232: Statistical Computing I **Credit 2**

Course Description

This computing course will discuss the applications of the statistical methods discussed in the courses design of experiment and demography. In addition, this course demonstrates the use of statistical software R to apply the models related to different experimental designs to

analyze the data in the field of agriculture, industry, and medicine. Moreover, this course helps students to learn, analyze and interpret real-life demographic data.

Course Objectives

The course's main objective is to teach students to use statistical software (e.g., R) and analyze experimental design and demographic data. The specific objective is to provide hands on learning by incorporating practical exercises and real-world examples to equip students with the knowledge and skills needed to excel in conducting scientific research or working in industry.

Learning Outcomes

Upon completion of the course, students will be able to

- CLO1 have a solid understanding of the use of statistical software, e.g., R, and methods to analyze data in the area of experimental design and demography
- CLO2 learn how to apply the models related to different experimental designs (CRD, RBD, BIBD) to analyze the data in the field of agriculture, industry, and medicine
- CLO3 learn to analyze and interpret demographic data, gaining insights into population trends, age-sex structures, fertility, mortality, migration, population projections and population growth rates
- CLO4 be equipped with the skills to write programs to obtain statistical results from data in the relevant field and draw meaningful conclusions from statistical analyses, which are essential for informed decision-making in various fields for public policy and planning.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	1	2	1	3	3			
CLO2	1	2	1	3	3	3		
CLO3	1	1	1	3	3	3		
CLO4	1	1	1	3	3	2	2	2

Contents

Problems reelated to design of experiments (completly randomized, randomized block, Latin square, split-plot and nested designs) and demography, calculation of different measures related to fertility and mortality, and lifetable.

AST240: Oral II **Credit 2**

Each student must be examined by a committee of selected members at the end of the academic year.

3.3 Detailed Syllabus - Third Year

AST301: Design and Analysis of Experiments II **Credit 4**

Course Description

This course explores the concepts and techniques involved in factorial design. It covers how to design factorial experiments, conduct them, and analyze the resulting data. Additionally, the course addresses experiments with random factors, as well as nested and split-plot designs.

Course Objectives

The main objectives of the course are to teach the basic ideas of factorial design; design experiments involving up to 2 and 3 factors with k levels; analyze the data from such experiments; interpret the results of any analysis; and design and analyze experiments involving random factors. Additionally, the course aims to teach the methods involved in designing and analyzing experiments with k factors, each having two levels and three levels.

Learning Outcomes

After completion of the course, students are expected to

- CLO1 understand the concept of factorial designs and design experiments involving up to two or three factors with k levels.
- CLO2 understand the concept of blocking and confounding in designs involving k factors with two levels.
- CLO3 understand the design and analysis of fractional factorial experiments for k factors, each with two levels, as well as experiments involving random factors.
- CLO4 develop skills to analyze the data produced from various experimental designs.
- CLO5 demonstrate skills in designing and implementing proper factorial experiments and explaining them to various stakeholders across different sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2	1				
CLO2	3	1	3	1				
CLO3	3	1	3	1	2			
CLO4	1		1	2	1			
CLO5			1	3	1	2	3	1

Contents

Introduction to factorial designs: basic definition and principles; advantage of factorials; two-factor factorial design; statistical analysis of fixed effects model, model adequacy checking,

estimating model parameters, choice of sample size, assumption of no interaction in a two-factor model, one observation per cell; general factorial design; fitting response curve and surfaces; blocking in a factorial design.

2^k factorial design: introduction; 2^2 design; 2^3 design; general 2^k design; a single replicate in 2^k factorial design; blocking in a 2^k factorial design; confounding in 2^k factorial design; confounding in 2^k factorial design in two blocks; confounding in 2^k factorial design in four blocks; confounding in 2^k factorial design in 2^p blocks; partial confounding.

Two-level fractional factorial designs: one-half fraction of 2^k design; one-quarter fraction of 2^k design; general 2^{k-p} fractional factorial design; resolution III designs; resolution IV and V designs.

Three-level and mixed-level factorial and fractional factorial designs: 3^k factorial design, confounding in 3^k factorial design, fractional replication of 3^k factorial design, factorials with mixed levels.

Response surface methods: introduction to response surface methodology; method of steepest ascent; analysis of second-order response surface; experimental designs for fitting response surfaces; mixture experiments; robust designs.

Experiments with random factors: random effects model; two-factor factorial with random factors; two-factor mixed model; sample size determination with random effects; rules for expected mean squares; approximate F tests; approximate confidence intervals on variance components; modified large-sample method; maximum likelihood estimation of variance components.

Nested and split-plot designs: two-stage nested designs; statistical analysis, diagnostic checking, variance components; general m-staged nested design; designs with both nested and factorial factors; split-plot design; split-plot designs with more than two factors; split-split-plot design, strip-split-plot design.

Textbooks

1. Montgomery DC (2019). Design and Analysis of Experiments, *10th edition*. Wiley.

Reference Books

1. Dean AM, Voss AM, and Draguljić (2017). Design and Analysis of Experiments, *2nd edition*. Springer.
2. Bailey R (2008). Design of Comparative Experiments. Cambridge.

AST302: Sampling Methods II | **Credit 3**

Course Description

The main objectives of this course is to acquaint students with the methodologies associated with the advanced sampling techniques. This course introduces advanced sampling methods used in sample survey. It covers sampling of unequal clusters, two-stage sampling, multistage sampling, methods for estimating variance in complex surveys, and non-sampling errors.

Course Objectives

To acquaint students with the methodologies associated with the advanced sampling techniques and their application in real survey.

Learning Outcomes

Upon completion of the course, students are expected to be able to

CLO1 understand the concept of a different sampling methods

CLO2 gain knowledge about different multistage sampling, methods for estimating variance in complex surveys and non-sampling error and their practical uses in large scale survey.

CLO3 understand the methods of estimating the population characteristics of interest under different sampling scheme.

CLO4 develop skills to determine the required sample size and calculate parameters under different sampling scheme.

CLO5 demonstrate skills for designing and implementing a standard sample survey and explaining it to different stakeholders in many sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	3	1				
CLO2	2	1	3	2				
CLO3	1	1	3	1	2			
CLO4	1		2	2	2			1
CLO5	1		1	3	1	3	3	1

Contents

Probability proportional to size (PPS) sampling: motivating examples, with replacement (WR) sampling: cumulative measure of size method and Lahiri's method, Hansen-Hurwitz (H-H) estimator (unbiasedness, variances, estimated variance), comparison with SRS, optimum measure of size, PPS without replacement (WOR) sampling: challenges and solutions, initial probabilities, normalizing probabilities, inclusion probabilities and their relation with the sample size, Horvitz-Thompson (H-T) estimator (unbiasedness, variance, estimated variance), different methods of PPSWOR: Brewer's method, Durbin's method, Des Raj method, Murthy's method, Rao- Hartley-Cochran method, Multinomial distribution for PPSWR sampling, H-T estimator in case of PPSWR sampling.

Sub-sampling, sub-sampling of unequal sized clusters: different estimators and their variances, two-stage sampling: design, estimators (total, mean), variances and their unbiased estimators, three stage sampling: design, estimators (total, mean), variances and their estimators, general framework (two-stage and three-stage) for estimating population total, different sampling designs at different stages, determination of sample sizes in two and three-stage sampling, optimum sampling and sub-sampling fractions, use of information from pilot survey.

Concept of double sampling and its necessity, application in stratified sampling, and in Ratio and Regression estimators, repeated sampling from the same population: sampling on two and more than two occasions.

Complex survey: definition and challenges involved in complex surveys, approaches of variance estimation (VE), replication methods for VE: random group method, balanced repeated replication (balanced half-sample replication) method, Jackknife method and Bootstrap method, implementation of replication methods in complex sampling designs, post-stratification.

Non-sampling errors: sources of the errors, effects of nonresponse, inference on population proportion in presence of nonresponse, types of nonresponse, Call-backs and its effects, Hansen and Hurwitz approach for nonresponse, Politz-Simmons adjustment for bias reduction, mathematical model for errors of measurement, mechanism of nonresponse, imputation and its different techniques.

Special sampling designs: multiplicity, network sampling: design and estimation (multiplicity and Horvitz-Thompson estimators for population total, and their different properties), adaptive sampling: adaptive cluster sampling (ACS) and related concepts used in ACS, Hansen-Hurwitz and Horvitz-Thompson estimators for population total, and their different properties.

Textbooks

1. Cochran WG (1977). Sampling Techniques, *3rd edition*. Wiley.

Reference Books

1. Lohr SL (1998). Sampling: Design and Analysis. Duxbury.
2. Thompson SK (2012). Sampling, *3rd edition*. Wiley.
3. Kirk MW (2007). Introduction to Variance Estimation, *2nd edition*.Wiley.

AST303: Linear Regression Analysis — **Credit 4**

Course Description

This introductory course provides an overview of different types of regression and describes the use of linear regression. This course covers the theory underlying regression analysis, simple and multiple linear regression, classical estimation and testing methods, and residual analysis. In addition, it includes the formulation, analysis, interpretation and validation of linear regression models as well as the practical application of a statistical package to demonstrate how the theory can be applied to address a particular research question.

Course Objectives

Students are expected to learn the techniques of linear regression models and they should understand applications of the linear regression models in real life situations.

Course Learning Outcomes (CLOs)

After completion of this course, students are expected to

CLO1 understand (explain ideas and concepts) Students will be able to explain the theory underlying regression analysis, including simple and multiple linear regression, and classical estimation and testing methods.

CLO2 analyze (draw connection among ideas) Students will be able to analyze residuals and interpret the results of linear regression models.

CLO3 evaluate (justify a stand or decision) Students will be able to evaluate the validity and accuracy of linear regression models using statistical testing methods.

CLO4 apply (use information in new situation) Students will be able to formulate and apply linear regression models to solve real-life problems using statistical packages such as STATA or R.

CLO5 create (produce a new or original work) Students will be able to create and validate new linear regression models to address specific research questions.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2	1	2	1	1
CLO2	2	3	3	3	2	2	1	1
CLO3	2	2	3	3	2	3	1	1
CLO4	2	2	2	3	2	3	1	1
CLO5	2	2	3	3	2	2	2	1

Contents

Measures of association for quantitative data: correlation and inference concerning correlation; regression and model building, motivating examples, uses of regression.

Simple linear regression model: model for $E(Y|x)$, least squares estimation, assumptions related to errors, maximum likelihood estimation (MLE) of model, sampling distribution of MLEs of the model parameters, inferences concerning the model parameters (confidence intervals and t-test), confidence interval estimate of the $E(Y|x)$ (confidence band).

Model accuracy and diagnostics: goodness of fit test (F-test, coefficient of determination, R^2); prediction and prediction interval for a new Y at specific x, residual analysis: definition, normal probability plot, plots of residuals versus fitted values, residuals versus x, other residual plots, statistical tests on residuals; detection and treatment of outliers; concept of lack of fit and pure error, test for lack of fit, transformations as solution to problems with the model, weighted least squares.

Matrix representation of simple linear regression model, inference and prediction.

Multiple linear regression models: formulation of multiple regression models, estimation of the model parameters: least squares estimation, maximum likelihood estimation, sampling distributions of the MLEs, confidence interval and hypothesis testing for concerning model parameters; model accuracy and diagnostics: goodness of fit test (F test, R^2, adjusted R^2), prediction of a new observation; extra sum of squares principles and its application in testing general linear hypothesis, checking all assumptions concerning model and use of remedy

measures when assumptions are not valid, detection and treatment of outliers, influential observations.

Polynomial regression model: introduction; polynomial models in one variable: basic principles, piecewise polynomial fitting; polynomial models in two or more variables; orthogonal polynomials.

Indicator variables: the general concept of an indicator variable, use of the indicator variables in linear regression, models with only indicator variables, idea of regression models with an indicator response variable.

Variable selection and model building: the model building problem, consequences of model mis-specification, criteria for evaluating subset regression models, computational techniques for variable selection.

Validation of regression models: concept, cross validation.

Textbooks

1. Weisberg S (2013). Applied Linear Regression, *4th edition*. Wiley.

Reference Books

1. Montogomery DC, Peck EA, and Vining GG (2012). Introduction to Linear Regression Analysis, *5th edition*. Wiley and Sons.
2. Draper NR and Smith H (1999). Applied Regression Analysis, *3rd edition*. Wiley.

AST304: Epidemiology **Credit 3**

Course Descriptions

This course serves as an introduction to epidemiology, the study of the distribution and determinants of health-related states or events in populations, and the application of this study to control health problems. Epidemiology is a cornerstone of public health, providing essential tools for understanding disease patterns and developing effective interventions. Topics include measures of disease occurrence, study designs, exposure-disease relationship, control of extraneous factors and interaction.

Course Objectives

The course aims to equip students with the knowledge, skills and abilities necessary to understand disease patterns, identify risk factors, and contribute to evidence-based decision-making in public health and healthcare settings. It also emphasizes critical thinking, problem-solving, and ethical conduct in epidemiological research and practice.

Learning Outcomes

After completion of the course, the students are expected to

CLO1 understand the basic principles and concepts of epidemiology

CLO2 learn key epidemiological study designs and methods

CLO3 develop skills in analyzing and interpreting epidemiological data

CLO4 learn strategies to minimize bias and confounding in study design and analysis

CLO5 learn the application of epidemiological findings in developing public health interventions and policies

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1						2		
CLO2	1	2						
CLO3	1	2		3	3	2		
CLO4						2		
CLO5							2	3

Contents

Introduction: disease processes, statistical approaches to epidemiological data, study design, binary outcome data, causality.

Measures of disease occurrence: prevalence and incidence, disease rates, hazard function; review of simple random samples, probability, conditional probabilities, and independence of two events.

Measures of disease-exposure association: relative risk, odds ratio, relative hazard, risk, attributable risk.

Study designs: population-based studies, cohort studies, case-control studies, case-cohort studies; Assessing significance of 2×2 tables obtained from chohort designs, case-control designs.

Estimation and inference for measures of assocaition: odds ratio, sampling distribution and confidence interval for odds ratio, relative risk, excess risk, attributable risk.

Confounding and interaction: causal inference, counterfactuals, confounding variables, control of confounding variables by stratification, causal graphs, controllong confounding in causal graphs; Cochran-Mantel-Haenszel test, summary estimates and confidence intervals for odds ratio and relative risk after adjusting for confounding factors.

Interaction: multiplicative and additive interaction, interaction and counterfactuals, test of consistency of association across strata, overall test of association, a test for trend in risk.

Introduction to matching, types of matching, and analysis of matched studies.

Textbooks

1. Jewell NP (2003). Statistics for Epidemiology. Chapman and Hall.

Reference Books

1. Timothy L. Lash. Tyler J. Vander Weele, Sebastien Haneuse, and Kenneth J. Rothman (2021). Modern Epidemiology, *4th edition*. Wolters Kluwer.
2. Fleiss JL, Levin B, and Paik MC (2003). Statistical Methods for Rates and Proportions, *3rd Edition*. Wiley.

AST305: Lifetime Data Analysis I | **Credit 3**

Course Description

This course deals with the analysis of time-to-event data (also known as survival or failure time data), which are commonly encountered in scientific investigations. It is being extensively used in medicine, clinical trials, biological and epidemiological studies, engineering, economics, and social sciences. This course provides an opportunity for students to learn lifetime probability distributions that are useful for modeling tine-to-event data. Topics include lifetime distributions and non-parametric and parametric approaches for analyzing time-to-event data.

Course Objectives

The course's primary focus is to make students familiar with the situation where censored data appear and to be able to analyze such data by implementing appropriate methods and models. Secondly, to introduce statistical theory and methodology for analyzing lifetime data obtained from censored samples, emphasizing statistical lifetime distributions, types of censoring, graphical techniques, and non-parametric/parametric estimation.

Learning Outcomes

After successful completion of the course, students are expected to

CLO1 understand the basic features of censored lifetime data, survival function, and hazard functions

CLO2 to learn different properties of lifetime distributions

CLO3 to understand the use of location-scale and log-location scale families of ditributions in the analysis time-to-event data and learn its properties

CLO4 understand different non-parametric methods for estimating survivor and hazard functions and corresponding inferences

CLO5 understand parametric estimation methods for survival and hazard functions and corresponding inferences

CLO6 be able to communicate conclusions obtained from time-to-event data analysis with researchers of related fields and policymakers

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3			1				
CLO2	3			1				
CLO3	3		2	2				
CLO4	3	2	3	2	2	1		1
CLO5	3	3	3	2	2	1		1
CLO6	2				2	3	3	2

Contents

Basic concepts and models: lifetime distributions - continuous models, discrete models, a general formulation; some important models - exponential, Weibull, log-normal, log-logistic, gamma distributions, log-location-scale models, inverse Gaussian distributions, mixture models; regression models. Observation schemes, censoring, and likelihood: right censoring and maximum likelihood; other types of incomplete data; truncation and selection effects; information and design issues. Nonparametric and graphical procedures: nonparametric estimation of survivor function and quantiles; descriptive and diagnostic plots; estimation of hazard or density functions; methods of truncated and interval censored data; life tables. Inference procedures for parametric models: inference procedures for exponential distributions; gamma distributions; inverse Gaussian distributions; grouped, interval censored, or truncated data; mixture models; threshold parameters; prediction intervals.

Inference procedure for log-location-scale distributions: inference for location-scale distributions; Weibull and extreme-value distributions; log-normal and log-logistic distributions; comparison of distributions; models with additional shape parameters; planning experiment for life tests.

Textbooks

1. Lawless J (2003). Statistical Models and Methods for Lifetime Data, *2nd Edition*. Wiley.

Reference Books

1. Kalbfleisch J and Prentice R (2003). The Statistical Analysis of Failure Time Data, *2nd edition*. Wiley.
2. Collett D (2014). Modelling Survival Data in Medical Research, *3rd edition*. Chapman & Hall/CR.

AST306: Statistical Inference II **Credit 3**

Course Description

To infer means to make general statements based on specific observations. This is a task undertaken in almost every quantitative field of research. This course deals with the advanced statistical inferential methods. It emphasizes concepts, methods and theory covering the broad topics of hypothesis testing, data reduction, likelihood inference in exponential families, and criteria for comparison of estimators. Successful completion of this course will provide a foundation for understanding probability-based statistical inference material presented in other courses.

Course Objectives

The objective of this course is to provide a thorough theoretical grounding and working knowledge of statistical inference. It aims to teach fundamental materials which will be of use in specialized courses in statistics such as actuarial science and biostatistics. At the end of this course, one should be able to perform hypothesis tests, explain the principles of data reduction, judge the quality of estimators as well as choose appropriate methods of inference to tackle real problems.

Learning Outcomes

After completion of the course, students are expected to

CLO1 Provide a thorough theoretical grounding and working knowledge of statistical inference, including understanding the concepts of sufficiency, ancillarity, and completeness, and their roles in designing estimators.

CLO2 Choose appropriate methods of inference to tackle real problems.

CLO3 Judge the quality of estimators and compare between estimators based on different criteria.

CLO4 Perform hypothesis tests with optimal properties and design hypothesis tests, and choose appropriate methods of inference to tackle real problems.

CLO5 Perform hypothesis tests with optimal properties and design hypothesis tests.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2	1	1	2	1
CLO2	3	2	3	3	2	2	2	1
CLO3	2	2	3	3	2	2	2	1
CLO4	2	3	3	3	2	2	2	1
CLO5	2	2	3	3	2	2	2	1

Contents

Statistical inference: parametric, nonparametric, and semiparametric inference.

Approximate and computationally intensive methods for statistical inference: the general problem of inference; likelihood functions; maximum likelihood estimation; optimization techniques-Newton type methods; EM algorithm-simple form, properties, uses in analyzing missing data, fitting mixture models and latent variable model; restricted maximum likelihood (REML) method of estimation; Multi-stage maximization; Efficient maximization via profile likelihood; confidence interval and testing hypothesis in these complex cases.

Bayesian method of inference: prior and posterior distribution, different types of prior, credible intervals and testing hypothesis; analytical approximations - asymptotic theory, Laplace approximation; numerical integral methods - Newton-Cotes type methods; Monte-carlo methods; simulation methods - Markov chain Monte Carlo.

Exact tests: test for single proportion and comparison of two proportions.

Resampling techniques: bootstrap-confidence intervals, test, parametric bootstrap, advantages and disadvantages of parametric bootstrap; jackknife-confidence interval, test, and permutation test.

Nonparametric inference and robustness: introduction, inference concerning cumulative distribution function (cdf), quantiles and statistical functionals: empirical cdf, quantiles, estimating statistical functionals, influence functions, testing statistical hypothesis-one sample settings, two or more sample settings; tolerance limit; empirical density estimation- histograms, kernel, kernel density estimation.

Textbooks

1. Mukhopadhyay N (2020). Probability and Statistical Inference. CRC Press.

Reference Books

1. Rice J (2013). Mathematical Statistics and Data Analysis., *3rd Edition*. Cengage Learning.
2. Hogg RV, McKean J and Craig AT (2019). Introduction to Mathematical Statistics, *8th Edition*. Pearson.

AST307: Multivariate Statistics I **Credit 3**

Course Description

The course provides a deeper knowledge within multivariate statistics, both in theory and applications. Multivariate distributions with emphasis on the multivariate normal distribution and its properties are considered. The course also provides deeper studies of the inference theory in multivariate analysis, for example concerning the mean and covariance matrices. The concepts that are more thoroughly treated in the course are: multivariate analysis of variance, profile analysis, multivariate regression analysis. Matrix algebra is used through the course as a necessary tool for the inference as well as computer sessions. An overview about necessary concepts from matrix algebra useful for the course will be given.

Course Objectives

Multivariate data analysis helps in the reduction and simplification of data as much as possible without losing any important details. As MVA has multiple variables, the variables are grouped and sorted on the basis of their unique features. The variables in multivariate data analysis could be dependent or independent. It is important to verify the collected data and analyze the state of the variables. In multivariate data analysis, it is very important to understand the relationship between all the variables and predict the behavior of the variables based on observations. It is tested to create a statistical hypothesis based on the parameters of multivariate data. This testing is carried out to determine whether or not the assumptions are true.

Learning Outcomes

After completion of the course, students are expected to

CLO1 appreciate the range of multivariate techniques available,

CLO2 be able to summarize and interpret multivariate data

CLO3 have an understanding of the link between multivariate techniques and corresponding univariate techniques

CLO4 be able to use multivariate techniques appropriately, undertake multivariate hypothesis tests, and draw appropriate conclusions.

CLO5 know how to analyse several real life data in business, economics, Demography, Public Health, and medical data etc. using through multivariate techniques.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3							
CLO2	3		2					
CLO3	2	3	3	3	3	1		
CLO4	1	2	3	2	2	2	1	
CLO5	2	1	3	1	3	2	2	

Contents

Preliminaries of multivariate analysis: applications of multivariate techniques; the organization of data; data display and pictorial representations; distance.

Random vectors and random sampling: some basics of matrix and vector algebra; positive definite matrices; a square-root matrix; random vectors and matrices; mean vectors and covariance matrices; matrix inequalities and maximization; the geometry of the sample; random sample and expected values of sample means and covariance matrix; generalized variance; sample mean, covariance, and correlation as matrix operations; sample values of linear combinations of variables.

The multivariate normal distribution: the multivariate normal density and its properties; sampling from a multivariate normal distribution and maximum likelihood estimation; sampling distribution and large sample behavior of sample mean vector and sample variance-covariance matrix; assessing the assumption of normality; detecting outliers and data cleaning; transformation to near normality.

Inferences about a mean vector: the plausibility of mean vector as a value for a normal population mean; Hotelling T^2 and likelihood ratio tests; confidence regions and simultaneous comparisons of component means; large sample inference about a population mean vector; inferences about mean vectors when some observations are missing; time dependence in multivariate data.

Comparisons of several multivariate means: paired comparisons and a repeated measures design; comparing mean vectors from two populations; comparison of several multivariate population means (one-way MANOVA); simultaneous confidence intervals for treatment effects; two-way multivariate analysis of variance; profiles analysis; repeated measures designs and growth curves.

Multivariate linear regression models: the classical linear regression model; least squares estimation; inferences about regression model; inferences from the estimated regression function; model checking; multivariate multiple regression; comparing two formulations of the regression model; multiple regression model with time dependent errors.

Textbooks

1. Johnson RA and Wichern DW (2008). Applied Multivariate Statistical Analysis, *6th edition.* Prentice-Hall.

Reference Books

1. Srivastava MS (2002). Methods of Multivariate Statistics. Wiley.
2. Anderson TW (2003). An Introduction to Multivariate Statistical Analysis. *3rd edition.* Wiley.

AST308: Research Methodology | **Credit 3**

Course Description

The research methodology course is fundamental to teaching students the theoretical and practical skills required to conduct research in various fields. The course introduces students to the fundamental principles of research design, data collection, data analysis, and report writing. It will cover a range of research approaches, including quantitative, qualitative, and mixed methods.

Course Objectives

The course aims to provide students with hands-on experience in designing and conducting research studies, particularly to develop students' knowledge and understanding of how to formulate research questions, develop hypotheses, choose appropriate research methods, and

collect and analyze data and write a report for scientific publication. In addition, the course is designed to provide students with the knowledge and skills necessary to conduct valid, reliable, and ethical research.

Learning Outcomes

After completion of the course, students are expected to

- CLO1 formulate research questions, develop hypotheses, choose appropriate research methods, and collect and analyze data
- CLO2 learn about the various data collection techniques, such as surveys, interviews, observations, and statistical techniques for analyzing data
- CLO3 apply their knowledge through practical exercises
- CLO4 become competent in planning, conducting, evaluating, and presenting a research project
- CLO5 learn ethics-ethical issues, citation, acknowledgment, plagiarism, reproducibility, and accountability

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	1		1	3	3	1	2	3
CLO2	1		1	3	3	1	1	3
CLO3				3	1		2	2
CLO4	1		1	3	3	1	3	2
CLO5		2		2	3	1	3	3

Contents

Foundations of Research: Meaning, concept, motivation, and objectives of the research; Types of research – descriptive vs. analytical, applied vs. fundamental, quantitative vs. qualitative, conceptual vs. empirical, concept of the applied and basic research process. Criteria and steps of good research.

Language of research: Variables and attributes; concepts and constructs; theory and models; proposition and hypotheses: relational terminologies: independent and dependent variables, intervening variables, extraneous variables, moderating variables confounder variables.

Problem Identification and Formulation: Defining and formulating the research problem, the importance of literature review in defining a problem, identifying gap areas from literature and research database, research question, and formulation of research hypothesis.

Research Design and Methods: Research methods vs. methodology; features of a good research design, concept, types, and uses of exploratory, descriptive, and experimental research designs. Qualitative and quantitative research approach, mixed methods design. Concept of measurement, problems in measurement in research - validity, and reliability. Sampling: use

of sampling techniques in research design; brief review of sampling and sample size determination.

Techniques of data collection: Qualitative approaches – focus group discussion (FGD), in-depth interview (IDI), key informant interview (KII) and their guidelines and checklist; concepts of ethnography, content analysis, and discourse analysis, Quantitative approach – the concept of a survey, mode of a survey – face-to-face interview, telephone interview, online/email, etc. Survey tools – Questionnaire, schedule, structured, semi-structured, open-ended, and close-ended questions. Data collection apps: survey CTO, Kobo Toolbox, etc.; Field Implementation – a pilot study, monitoring the data collection, quality control, and data validity.

Monitoring and Evaluation (M&E): Concept of monitoring and evaluation, objectives of M&E; performance monitoring versus performance evaluation. Key steps of M&E; M&E design: Baseline, ongoing, and end-line evaluation; evaluation criteria–relevance, effectiveness, efficiency, impact, and sustainability.

Human resources and budget in research: allocation of human resources, preparing a budget for research work.

Data Analysis: Steps of data preparation and analysis; Guidelines for using secondary data for research: when, how, and why;

Dissemination of Research findings: Different Steps in Writing Report, Layout of the Research Report; Effective presentation, preparing articles for peer review.

Research Ethics and Scholarly Publishing: Ethics-ethical issues, ethical committees (human and animal); IPR- intellectual property rights and patent law, commercialization, copyright, royalty; scholarly publishing concept and research paper design, citation, acknowledgment, plagiarism, reproducibility, and accountability.

Experiential Learning Project: Group project and field work – writing a research proposal, developing and implementing field data collection, data analysis, and report writing and presentation.

Textbooks

1. Kothari CR and Garg G (2019). Research Methodology: Methods & Techniques. New Age.

Reference Books

1. Gertler PJ, Martinez S, Premand P, Rawlings LB and Vermeersch CMJ (2017) . Impact Evaluation in Practice, *2nd edition*. World Bank Group, Washinngton DC.
2. Coninck JD, Chaturvedi K, Haagsma B, Griffioen H and Glas MVD (2008). Planning, monitoring and evaluation in development organizations: sharing training and facilitation experiences. Sage.

AST330: Stata for Data Science **Credit 2**

Course Description

This course is designed to provide students with a comprehensive understanding of using Stata-a powerful statistical software package-for data science applications. Students will learn how to leverage Stata's capabilities to explore, analyze, and visualize data, with a focus on practical skills relevant to data science and statistical analysis.

Objectives

The objectives of the course are to empower students with practical Stata skills and analytical competencies essential for conducting data-driven research and making informed decisions in various academic, professional and research settings. The course emphasizes hands-on learning of data entry, management, analysis and effective communication of findings.

Learning Outcomes

After completion of the course, the students are expected to

- CLO1 gain proficiency in using Stata for data entry, manipulation, cleaning and transformation
- CLO2 develop skills in data visualization and presentation of results
- CLO3 apply Stata for reproducible research
- CLO4 develop problem-solving skills by tackling data analysis challenges
- CLO5 apply statistical reasoning and Stata skills to address research questions and make data-driven decisions

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1				2	3			
CLO2	1		1			3	1	
CLO3				2	3			
CLO4				2	3			
CLO5				3			3	2

Contents

Introduction to Stata: different windows and files, help file and searching for information; data entry, reading both stata and and other format of data file, combining Stata files; exploring data: example commands-browse, edit, list, sort, describe, assert, codebook; data management: creating a new data set specifying subsets of data with in and if qualifiers, generating and replacing variables, using functions based on egen command, converting numeric and string formats, creating new categorical and ordinal variables, reshaping or collapsing data, weighting observations, creating random data and random samples; graphs: example

commands- histograms, scatterplots, line plots, connected-line plots, two-way plots, box plots, combining graphs; exploratory data analysis: summary statistics and tables: example commands - summarize, tabstat, table; frequency tables and two-way cross tabulations, multiple tables and multi-way cross tabulations, tables of means, medians and other summary statistics.

Textbooks

1. Hamilton LC (2006). Statistics with Stata, Thomson Brooks/Cole.

Reference Books

1. Acock AC (2010). A Gentle Introduction to Stata, *6th edition*. Stata Press.

CSE331: Python for Data Science | **Credit 2**

Course Description

Python has become one of the most popular programming languages for data science due to its simplicity, versatility, and robust data analysis and machine learning capabilities. As organizations continue to digitize and generate increasingly large amounts of data, open-source tools like Python can enable data manipulation and exploration, machine learning, and statistical analysis in a scalable manner. This course provides a gentle introduction to programming in Python and its applications in data science and analytics.

Objectives

To give students a working knowledge of Python programming. To teach programming fundamentals, basic data structures, writing functions, and importing and exporting data in different formats. To cover data manipulation and data visualization using Python. To prepare students for more advanced courses in data science and machine learning.

Learning Outcomes

Upon completion of CSE-530, students will be able to

- CLO1 Understand the fundamentals of Python programming, including basic syntax, data types, control structures, and functions (Concept, Computing).
- CLO2 Perform data analysis using NumPy and pandas, including data manipulation and exploratory data analysis (Application, Computing).
- CLO3 Create effective data visualizations using matplotlib and seaborn (Application, Computing).
- CLO4 Produce dynamic statistical reports using interpreting and documenting visualizations and analyses of real-world problems. (Communication, Collaboration).

Contents

Fundamentals of Python: Installing Python and Jupyter Notebook; the basic syntax of a Python program, Python data types; expressions and variables; lists, tuples, sets, and dictionaries; writing conditions, loops, and functions.

Data analysis with NumPy and pandas: installing NumPy and pandas, NumPy arrays; indexing, slicing, and iterating NumPy arrays; arithmetic and matrix operations with NumPy; pandas objects– DataFrame, Series, and Index; data indexing and selection; handling missing data; combining and joining datasets, aggregation and grouping, exploratory data analysis.

Data visualization with matplotlib and seaborn: Bar plots, histograms, density plots. boxplots and scatterplots.

Textbooks

1. McKinney W. (2022). Python for data analysis: Data wrangling with Pandas, NumPy, and Jupyter, *3rd edition*. O'Reilly.
2. VanderPlas J. (2016). Python data science handbook: Essential tools for working with data. O'Reilly.

Reference Books

1. Grus, J. (2019). Data science from scratch: First principles with python. O'Reilly.

AST332: Statistical Computing II | **Credit 2**

Course Description

This statistical computing course provides a comprehensive understanding of applications of the methods covered in the courses design and analysis of experiments, epidemiology, and lifetime data analysis to real-life problems, equipping students with the tools and knowledge to address real-life problems across various fields. By achieving these skills, students are expected to make data-driven decisions and contribute to solving the complex challenges faced by society today.

Course Objectives

This course aims to equip students with the skill to use statistical software (e.g., R, Stata, or Python) to apply the methods discussed in the courses design and analysis of experiments, epidemiology, and lifetime data analysis. Through these skills, students are expected to tackle problem-solving challenges while working in industry, agency, or academia, make data-driven decisions, and contribute to planning and policy.

Learning Outcomes

After completion of the course, the students are expected to

CLO1 understand how to use R, Stata or Python to linear models used in factorial experiments

CLO2 to be able to perform analysis related to epidemiological studies

CLO3 to be able to obtain non-parametric and parametric estimates of survival and cumulative hazards functions

CLO4 able to interpret the statistical results for decision-making processes related to healthcare, engineering, finance, and more.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1				3	2	2		2
CLO2				3	2	2		2
CLO3				3	2	2		2
CLO4					2	2	3	2

Contents

Analysis of factorial designs, screening experiments, split-plot designs, and response surfaces. Estimating odds ratio, risk ratio, and risk difference for different types of study designs and corresponding inference. Estimating the survival function and cumulative hazards function using non-parametric and parametric methods.

AST333: Statistical Computing III **Credit 2**

Course Description

This computing course is based on applications of regression models, multivariate methods, and statistical inference to real-life problems. By mastering these techniques for solving real-life problems, students will be equipped to analyze complex datasets, uncover meaningful insights, and contribute to data-driven decision-making processes in academia, industry, and research.

Objectives

The objectives of this computing course are to teach students the applications of some statistical methods (regression models, multivariate methods, and statistical inference) using statistical software (R, Python, or Stata) and interpret the results. More specifically, the idea is to engage students in hands-on exercises and projects that involve real-life datasets, providing practical experience in applying regression models, multivariate methods, and statistical inference techniques to address real-world problems.

Learning Outcomes

After completion of the course, the students are expected to

- CLO1 understand the applications of linear regression models to real-life problems
- CLO2 perform parametric and nonparametric hypothesis tests and corresponding confidence intervals
- CLO3 effectively utilize statistical computing tools and software packages such as R, Stata, or Python to implement regression models, multivariate methods, and statistical inference techniques
- CLO4 manipulate and analyze large datasets, visualize results, and present their findings clearly and concisely.

CLOs	Program Learning Outcomes (PLOs)							
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1				3	3	2	2	
CLO2				3	3	2	2	
CLO3				3	3	2		
CLO4				3	3	2		

Contents

Problems related to fitting simple and multiple linear regression models, related inference, model diagnostics, and model selection; Exploratory analysis of multivariate data, multivariate tests, MANOVA and multivariate regression, hypothesis tests, confidence intervals, and bootstrap methods, and nonparametric tests.

AST340: Oral III **Credit 2**

Each student must be examined by a committee of selected members at the end of the academic year.

3.4 Detailed Syllabus - Fourth Year

AST401: Advanced Probability and Stochastic Processes	**Credit 3**

Course Description

This module covers two important parts, one is advanced probability and another is stochastic processes. The module begins with a brief overview of modern probability and convergence of random variables and then moves on to the stochastic process to describe Markov chains, Poisson process, continuous time Markov chain and renewal process.

Course Objectives

To develop a deep understanding of the advanced concepts in probability theory and stochastic processes. To apply these concepts to solve real-world problems in various domains. To prepare students for research and professional practice in fields requiring sophisticated stochastic modeling.

Course Learning Outcomes (CLOs)

Upon completion of the course, students will be able to

- CLO1 demonstrate advanced knowledge of probability theory and stochastic processes, including probability spaces, random variables, and convergence of random variables
- CLO2 apply probability theory to model and analyze random experiments
- CLO3 solve complex probabilistic problems using Markov chains, Poisson processes, and renewal processes
- CLO4 develop skills to derive and prove theorems and lemmas related to probability theory and stochastic processes.
- CLO5 demonstrate skills to interpret the results obtained from stochastic modeling and provide recommendations.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	3					
CLO2			3	3	2	2		
CLO3			3	3	2	2		
CLO4	1	3	3					
CLO5						3		

Contents

Modern probability: events as sets, field, sigma field, probability measure, Borel field and extension of probability measure, measure theoretic approach of random variables; probability

space.

Convergence of random variables: modes of convergence: almost sure, rth mean, in probability, in distribution, their interrelationship; law of large numbers, strong and weak laws of large numbers, limiting distribution; central limit theorem; law of iterated logarithm; martingale.

Markov chains: introduction, Chapman-Kolmogorov equations, classification of states, limiting probabilities, gambler's ruin problem, mean time spent in transient states, branching processes, time reversible Markov chains, Markov chain Monte Carlo methods, Markov decision processes, hidden Markov chains.

Poisson process: exponential distribution, properties, convolutions of exponential random variables; counting processes, Poisson process, interarrival and waiting time distributions, further properties of Poisson processes, generalizations of the Poisson process, nonhomogeneous Poisson process, compound Poisson process, conditional or mixed Poisson processes.

Continuous-time Markov chains: introduction, continuous-time Markov chains, birth and death processes, transition probability function, limiting probabilities, time reversibility.

Introduction to queueing theory: Classical M/M/1 queue, global and local balance, performance measures, Poisson arrivals see time averages (PASTA) property, M/M/1/S queueing systems, blocking probability, performance measures, multi-server systems – M/M/m, performance measures, waiting time distribution of M/M/m, performance measures of M/M/m/m with finite customer population, Erlang loss systems.

Renewal theory and its applications : introduction, distribution of renewals, limit theorems and their applications, renewal reward processes, regenerative processes, semi–Markov processes, Markov renewal processes.

Textbooks

1. Grimmett G and Stirzaker D (2001). Probability and Random Processes, *3rd edition*. Oxford.

Reference Books

1. Ross S (2010). Introduction to Probability Models, *10th edition*. Elsevier.
2. Jones PW and Smith P (2018). Stochastic Processes: An Introduction, *3rd Edition*. Routledge.

AST402: Statistics for AI and Machine Learning | **Credit 4**

Course Description

The course provides an introduction to the field of Artificial Intelligence and a broad but thorough introduction to the methods and practice of basic statistical machine learning and its core methods, models, and algorithms. This course is about providing students of applied statistics and data science with detailed knowledge of how the basic Machine Learning methods work for artificial intelligence and how statistical models can be brought to bear in computer systems not only to analyze large, high-dimensional, and big data sets but also

to let computers perform tasks, that traditional methods of computer science are unable to address.

Course Objectives

After completing the course students will be familiar with the concept of Artificial Intelligence, Big Data, and their areas. It describes philosophies and techniques of Artificial Intelligence, including intelligent agents, and search. It also describes several models for supervised and unsupervised machine learning and regularized modeling techniques, assess the strengths and weaknesses of each of these models, know the underlying mathematical relationships within and across statistical learning algorithms, develop and implement optimization methods for training of statistical models, design decision and optimal control problems to improve the performance of statistical learning algorithms, design and implement various statistical machine learning algorithms in real-world applications, evaluate the performance of various statistical machine learning algorithms, demonstrate a working knowledge of dimension reduction techniques. identify and implement advanced computational methods in machine learning.

Learning Outcomes

After completion of the course, students are expected to

- CLO1 gain understanding of Artificial Intelligence (AI), Machine Learning, Big Data, and their application areas.
- CLO2 be able to identify and describe several models for supervised and unsupervised machine learning, including regularized modeling techniques for AI.
- CLO3 develop skills to analyze, interpret, and predict data using various machine-learning methods based on various resampling techniques, with a particular focus on AI.
- CLO4 understand and explain the underlying mathematical relationships within and across statistical learning algorithms.
- CLO5 develop the ability to conduct comprehensive data analyses for high-dimensional data, including fitting, diagnosing, and validating various machine learning models, to address real-world data challenges effectively for AI.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2		2		
CLO2	3	3	2	3		2		
CLO3	2	3	1	2	1	2	1	
CLO4	2	3	2	1				
CLO5	3	3	2	2	1	1	1	

Contents

Introduction to Artificial Intelligence—Definition and History of AI, AI Applications and Impact on Society, Overview of AI Subfields, Components of an AI system. Problem Solv-

ing and Search—Problem Formulation, Uninformed Search Strategies: Breadth-First Search, Depth-First Search, Informed Search Strategies: Best-First Search, A* Algorithm.

Overview of basic concepts and techniques of statistical machine learning, Data science concepts and their application areas. Overview of supervised and unsupervised learning.

Overview of Big Data, including its definition, sources, characteristics, and the challenges associated with processing and analyzing large datasets. Big data applications: examples and use cases. The Hadoop framework and its various components include HDFS, MapReduce, YARN, and Hive.

Statistical learning: Statistical learning and regression, the curse of dimensionality and parametric models, assessing model accuracy and bias-variance trade-off, classification problems, and K-nearest neighbors.

Linear regression: Model selection and qualitative predictors, interactions, and nonlinearity.

Classification: Introduction to classification, logistic regression, and maximum likelihood, multivariate logistic regression and confounding, case-control sampling and multiclass logistic regression, linear discriminant analysis and Bayes theorem, univariate linear discriminant analysis, multivariate linear discriminant analysis, and ROC curves, quadratic discriminant analysis and naive bayes.

Resampling methods: Estimating prediction error and validation set approach, k-fold cross-validation, cross-validation- the right and wrong ways, the bootstrap, more on the bootstrap. Linear model selection and regularization: Linear model selection and best subset selection, forward stepwise selection, backward stepwise selection, estimating test error using Mallow's Cp, AIC, BIC, adjusted R-squared, estimating test error using cross-validation.

Shrinkage methods: Ridge regression, Lasso, adaptive Lasso, Group Lasso, elastic net and adaptive elastic net, bridge, tuning parameter selection for ridge regression, lasso, and elastic net, total variation regularization.

Dimension reduction and unsupervised machine learning techniques: Techniques to reduce the dimension of data, principal components and principal components regression and partial least squares. Clustering (K-means, Fuzzy C-means, Hierarchical Clustering).

Introduction to Decision trees and Classification trees.

Artificial Intelligence and Statistical machine learning case studies in R and Python.

Textbooks

1. Stuart J. Russell and Peter Norvig (2021). Artificial Intelligence: A Modern Approach, *4th edition*, Pearson Education Inc.
2. James G, Witten D, Hastie T and Tibshirani R (2013). An Introduction to Statistical Learning: with Applications in R. Springer.

Reference Books

1. Hastie T, Tibshirani R and Friedman J (2009). The Elements of Statistical Learning: Data Mining, Inference and Prediction, *2nd edition.* Springer.
2. Russell S and Norvig, P (2020). Artificial Intelligence: A Modern Approach, *4th edition.* Pearson Education.

AST403: Multivariate Statistics II **Credit 3**

Course Description

Multivariate analysis skills are recognized as part of the key requisites for statistical analysts. The complexity of most real-world phenomena requires investigators to collect and analyze observations on many different variables rather than a single variable. The need for statistical techniques to extract information from multivariate data thus becomes essential and crucial for data analysts. This course focuses on multivariate methods based on normal theory. It provides students with working knowledge on how to analyze data and solve problems involving measurements of p variables on each of n subjects.

Course Objectives

The objective of this course is to provide students with experience in using multivariate techniques for analyzing research data. The aim is to teach students how to select appropriate methods for multivariate data analysis and interpret the results effectively.

Learning Outcomes

After completion of the course, students are expected to

CLO1 perform a principal components analysis, and interpret the principal component scores.

CLO2 understand the concept of factor analysis, and different methods to estimate the parameters of a factor model; understand factor rotation, and interpret the rotated factor loadings.

CLO3 perform a canonical correlation analysis, and interpret the canonical variate scores;

CLO4 carry out cluster analysis, know how to partition the data into clusters of known composition.

CLO5 demonstrate skills for implementing appropriate methods for multivariate data analysis, interpret the results, and explaining it to different stakeholders in many sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	2	2	2			
CLO2	3	2	2	2	1			
CLO3	3	1	2	2	1			
CLO4	3	2	1	1				
CLO5				3	1	3	3	1

Contents

Principal components: population principal components, summarizing sample variations by principal components, graphing the principal components, large sample inference.

Factor analysis: the orthogonal factor models, methods of estimation (maximum likelihood estimates and principal factor analysis), selection of loadings and factor (factor rotation, varimax rotation, quartimax rotation, oblimin rotations), factor scores, structural equations models.

Canonical correlation analysis: canonical variates and canonical correlations, sample canonical variates and sample canonical correlations, large sample inference.

Discrimination and classification: separation and classification of two populations, classification of two multivariate normal populations, evaluating classification functions, Fisher's discriminant function, classification with several populations, Fisher's method for discriminating several populations.

Clustering: similarity measures, hierarchical clustering methods, nonhierarchical clustering methods; fuzzy clustering, determination of number of clusters: Gap statistics and its several modifications, several cluster validity indices, cluster's homogeneity test; multidimensional scaling.

Textbooks

1. Johnson RA and Wichern DW (2008). Applied Multivariate Statistical Analysis, *6th edition*. Prentice-Hall.

Reference Books

1. Srivastava MS (2002). Methods of Multivariate Statistics. Wiley.
2. Anderson TW (2003). An Introduction to Multivariate Statistical Analysis. *3rd edition.* Wiley.

AST404: Econometric Methods | **Credit 4**

Course Description

This course covers a range of econometric methods required to conduct empirical economic research and understand applied econometric results. Topics include models for panel data, simultaneous equations models, models with lagged variables, and limited dependent variables.

Course Objectives

The objectives of the course are to introduce students to the basic principles of econometric analysis; to help students gain theoretical understanding of the statistical models needed for econometric research including their underlying assumptions, advantages and limitations and to help them understand how to apply different econometric tools to analyze real-world economic problems and interpret findings.

Learning Outcomes

On successful completion of this course the students are expected to

CLO1 understand the basic principles of econometric analysis and general steps for building econometric models.

CLO2 gain theoretical understanding of the statistical models needed for econometric research including underlying assumptions, sources and consequences of model misspecifications and comparative advantages and disadvantages of different models.

CLO3 learn about methods for estimating econometric models from data.

CLO4 apply different econometric tools to analyze real-world economic phenomena and problems.

CLO5 demonstrate skills to summarize, interpret and explain results of econometric analyses.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2	1				
CLO2	2	1	3	3				
CLO3	1	1	3	1	2			
CLO4	1		2	2	2			1
CLO5	1		1	3	1	3	3	2

Contents

Econometric modeling, data and methodology; specification analysis and model building: bias caused by omission of relevant variables, pretest estimation, inclusion of irrelevant variables, model building; testing non-nested hypotheses, encompassing model, comprehensive approach-J test, Cox test; model selection criteria.

Models for panel data: fixed effects: testing significance of group effects, within- and between-groups estimators, fixed time and group effects, unbalanced panels and fixed effects; random effects: GLS, FGLS, testing for random effects, Hausman's specification test.

Simultaneous equations models: illustrative systems of equations, endogeneity and causality; problem of identification: rank and order conditions for identification; limited information estimation methods: OLS, estimation by instrumental variables (IV), Two-Stage Least Squares (2SLS), GMM Estimation, limited information maximum likelihood and the k class of estimators, 2SLS in nonlinear models; system methods of estimation: Three-Stage Least Squares (3SLS). full-information maximum likelihood, GMM estimation, recursive systems and exactly identified equations; comparison of methods-Klein's Model I; specification tests; properties of dynamic models: dynamic models and their multipliers.

Models with lagged variables: lagged effects in a dynamic model, lag and difference operators; simple distributed lag models: finite distributed lag models, infinite lag model: geometric lag

model; Autoregressive Distributed Lag (ARDL) models: estimation of the ARDL model, computation of the lag weights in the ARDL model, stability of a dynamic equation, forecasting; Vector Autoregressions (VAR): model forms, estimation, testing procedures, exogeneity, testing for Granger causality, impulse response functions, structural VARs, application: policy analysis with a VAR.

Limited dependent variable: truncated distributions, moments of truncated distributions, truncated regression model; censored data: censored normal distribution, censored regression (Tobit) model, estimation, issues in specification; censoring and truncation in models for counts, application: censoring in the Tobit and Poisson regression models.

Textbooks

1. Greene WH (2011). Econometric Analysis, *7th edition*. Prentice Hall.

Reference Books

1. Gujarati DN (2010). Basic Econometrics, *5th edition*. McGraw-Hill.
2. Wooldridge JM (2010). Introductory Econometrics: A Modern Approach, *5th edition*. Cengage Learning.

AST405: Lifetime Data Analysis II | **Credit 3**

Course Description

Lifetime data (also known as survival or failure time data) are commonly observed in health science, engineering, and social sciences. This course deals with different approaches of modeling lifetime data. The theoretical backgrounds and related practical applications of both parametric and semi-parametric approaches of analysing lifetime data are discussed in the course. A brief introduction to competing risks and frailty models is also provided.

Course Objectives

The course primarily focuses on familiarizing students with regression models commonly used to analyze time-to-event (lifetime) data. Both the accelerated failure time (AFT) model and Cox's proportional hazards (PH) model will be discussed in detail in the course. Students will also learn about the methods for analyzing competing risks models and correlated time-to-event data.

Learning Outcomes

After completion of the course, students are expected to

CLO1 understand the concepts related to regression models for time-to-event data in general

CLO2 understand the theories related to accelerated failure time and proportional hazards model

CLO3 develop skills to interpret the results obtained from regression models for time-to-event data.

CLO4 understand the concepts of correlated time-to-event data and frailty models

CLO5 understand the concepts of multiple mode of failures and different approaches of competing risks models

CLO6 develop skills to explain regression model fits to different stakeholders

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	3	2		1		
CLO2	3	2	3	1		1		
CLO3	2		1	2	1	2	1	
CLO4	3	1	2	1				
CLO5	3	1	2	1		1		
CLO5	2				1	2	2	1

Contents

Parametric regression models: log-location-scale (Accelerated Failure Time) regression models, proportional hazards regression models; graphical methods and model assessment; inference for log-location-scale models; extensions of log-location-scale models; hazard-based models.

Semiparametric multiplicative hazards regression models: methods for continuous multiplicative hazards models - estimation and tests for regression parameter vector, comparison between two or more lifetime distributions, justification, and properties of the likelihood function, adjustment of tied lifetimes, estimation of baseline cumulative hazard function, stratification, time-varying covariates, model checking.

Competing risks: introduction, hazard, and cumulative incidence function, modeling cause-specific hazards and cause-specific incidence.

Frailty models: introduction to frailty, gamma frailty distributions, parametric frailty model, semi-parametric frailty model, and shared frailty model.

Textbooks

1. Lawless J (2003). Statistical Models and Methods for Lifetime Data, *2nd Edition*. Wiley.

Reference Books

1. Kalbfleisch J and Prentice R (2003). The Statistical Analysis of Failure Time Data, *2nd edition*. Wiley.
2. Collett D (2014). Modelling Survival Data in Medical Research, *3rd edition*. Chapman & Hall/CR.
3. Therneau TM and Grambsch PM (2000). Modeling survival data: extending Cox model. Springer.

CSE406: Big Data Analytics **Credit 3**

Course Description

This course is designed to introduce key computational concepts, tools and techniques for curating, managing, and analyzing data of large volume, various types and different frequencies. This course assumes basic exposure to the concepts of artificial intelligence, machine learning algorithms, deep learning and computer programming.

Course Objectives

The principal aim of this course is to introduce Big Data and its characteristics and challenges to students, and to teach them appropriate tools for managing and analyzing such large-scale data. Further objectives include helping students understand applications of Big Data in different fields and also ethical issues related to Big Data.

Learning Outcomes

After completion of the course, students are expected to

- CLO1 understand and describe the characteristics and challenges of Big Data.
- CLO2 identify and utilize appropriate tools for managing large-scale data.
- CLO3 understand and explain the underlying mathematical relationships within and across artificial intelligence tools and to design and implement those tools for real-world applications.
- CLO4 develop computational skills for analyzing data of large volume, various types, and different frequencies.
- CLO5 recognize and evaluate ethical issues related to the use of Big Data.

	Program Learning Outcomes (PLOs)						
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7
CLO1	2	3	3	3		2	
CLO2	2	2	2	3		2	
CLO3	3	3	1	2	2	2	1
CLO4	2	2	3	3		1	
CLO4	1	1	1	1	1	1	3

Contents

Introduction to Big Data: definition, characteristics and applications of Big Data in various fields.

Data pre-processing: data collection and extraction – scraping data, data cleaning- handling missing values, noisy data and outliers; redundancy and correlation analysis, tuple duplication, conflict detection and resolution. Structured and unstructured data and databases - relational

and NoSQL databases. Data reduction– overview, Wavelet transformation', Attribute Subset Selection, Data Cube Aggregation; Data Transformation and Discretization.

Introduction to Big Data Analytics: techniques to address data analysis issues related to data volume (Scalable and Distributed analysis), data velocity (High-Speed Data Streams), Data Variety (Complex, Heterogeneous, or Unstructured data), and Data Veracity (Data Uncertainty).

Database management essentials for Big Data organization and manipulation: Introduction to data organization (lists, queues, priority queues, trees, graphs, hash). Basic graph models and algorithms for searching, shortest path algorithms, flow networks, matching. Processing and streaming Big Data, introduction to data architecture software including MapReduce, Hadoop distributed file system, Spark, Terradata, and how these tools work.

Data Analysis and Visualization Techniques: Descriptive statistics, probabilistic modeling of Big Data (e.g., graphical models, latent variable models, hidden Markov models.) Bayesian Inference (e.g., variational inference, expectation propagation, sampling.) Bayesian Machine Learning (e.g., Bayesian linear regression). Fundamentals of data visualization, Infographics, layered grammar of graphics. Introduction to Modern, mosaic plots, parallel coordinate plots, introduction to GGobi data visualization system, linked plots, brushing, dynamic graphics, model visualization.

Big Data Ethics and Privacy: Ethical considerations in data collection and analysis, privacy and security concerns in Big Data, legal and regulatory frameworks for Big Data.

Textbooks

1. Balusamy B, Nandhini AR, Kadry S, and Gandomi AH (2021). Big Data: Concepts, Technology, and Architecture. Wiley.

Reference Books

1. Li K-C, Jiang H, Yang LT, and Cuzzocrea A (2015). Big Data: Algorithms, Analytics, and Applications. Chapman & Hall/CRC.
2. Erl T, Khattak W, and Buhler P (2016). Big Data Fundamentals: Concepts, Drivers & Techniques. The Prentice Hall.

AST407: Time Series Modelling | **Credit 3**

Course Description

This course serves as an introduction to time series theory. Its objective is to equip students with various classical time series models, including deriving their properties, inference methods, and forecasting techniques for analyzing time series data. From a computational standpoint, the course aims to demonstrate theory with real datasets. Conclusions and proofs are provided for some basic formulas and models to enable students to understand the principles of time series theory.

Course Objectives

Upon completion of the course, students should become acquainted with the main concepts of time series theory and methods of analysis. They should understand the differences between cross-sections and time series and grasp the fundamental advantages and necessity of appropriate forecasting methods in particular environments. Additionally, students should be proficient in using the statistical software R to analyze time series data.

Learning Outcomes

After completing the course, students are expected to

- CLO1 Understand the concept of time series analysis and be able to choose the most suitable analysis technique for the given data.
- CLO2 Have detailed knowledge of different time series techniques and the general steps for conducting and forecasting time series data.
- CLO3 Understand the methods for estimating the parameters of interest under different time series techniques.
- CLO4 Develop skills to forecast future observations and apply them in practice.
- CLO5 Demonstrate skills in designing and implementing a time series technique and explaining it to stakeholders in various sectors.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	2	1				
CLO2	2	1	3	2				
CLO3	2	2	3	1	1			
CLO4	1	1	2	3	1			
CLO5				1	2	3	3	

Contents

Introduction and examples of time series; simple descriptive techniques: time series plots, trend, seasonal effects, sample autocorrelation, correlogram, filtering.

Probability models: stochastic processes, stationarity, second-order stationarity, white noise model, random walks, moving average (MA) processes, autoregressive (AR) processes, ARMA processes, seasonal ARMA processes, the general linear process; properties, estimation and model building, diagnostic checking.

Forecasting: naive procedures, exponential smoothing, Holt-Winters, Box-Jenkins forecasting, linear prediction, forecasting from probability models.

Non-stationary time series: non-stationarity in variance-logarithmic and power transformations; non-stationarity in mean; deterministic trends; integrated time series; ARIMA and seasonal ARIMA models; modelling seasonality and trend with ARIMA models.

Stationary processes in the frequency domain: the spectral density function, the periodogram, spectral analysis.

Concept of state-space models: dynamic linear models and the Kalman filter.

Textbooks

1. Jonathan DC and Kung-Sik C (2008). Time Series Analysis - With Applications in R. Springer.
2. Spyros M, Steven W and Rob H (1997). Forecasting – Methods and Applications, *3rd edition*. Wiley.

Reference Books

1. Chatfield C (2003). The Analysis of Time Series, *6th edition*. Chapman & Hall.
2. Shumway RH and Stoffer DS (2011). Time Series Analysis and Its Applications: With R Examples. Springer.
3. Brockwell PJ and Davis RA (2002). Introduction to Time Series and Forecasting. *3rd edition*. Springer.

AST408: Generalized Linear Models **Credit 3**

Course Description

This course is about introducing statistical methods including the general linear model for quantitative responses (including multiple regression, analysis of variance, and analysis of covariance), binomial regression models for binary data (including logistic regression and probit models), and models for count data (including Poisson regression and negative binomial models). All of these techniques are covered as special cases of the Generalized Linear Model, which provides a central unifying statistical framework for the entire course.

Course Objectives

After completing the course students will be familiar with a unified procedure of several modeling techniques. They will be familiar with the exponential family of distributions, the class of generalized linear models (GLM) as regression models with responses from the exponential family of distributions, the concepts of link functions for modeling the correspondence between the expected value of the responses and covariates and of variance functions for specifying the correspondence between the expected values and variances of the responses, analyzing data from important special cases of GLMs, in particular logistic regression and Poisson regression, and extensions of the GLM framework using quasi-likelihood based on specified link and variance functions.

Learning Outcomes

After completion of the course, students are expected to

CLO1 gain a comprehensive understanding of the exponential family of distributions and their relevance to generalized linear models.

CLO2 describe and apply the generalized linear models for both quantitative and qualitative response variables within the GLM framework.

CLO3 develop skills to analyze and interpret data using various special cases of GLMs, with a particular focus on logistic regression and Poisson regression.

CLO4 extend the GLM framework using quasi-likelihood methods based on specified link and variance functions, allowing for flexible modeling of complex data structures.

CLO5 develop the ability to conduct comprehensive data analyses, including fitting, diagnosing, and validating various GLM models, to address real-world data challenges effectively.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	3	3	2		2		
CLO2	3	2	3	3		2		1
CLO3	2		1	2	1	2		1
CLO4	3	3	2	1				
CLO5	3	2	2	1	1	1		1

Contents

Generalized linear models: exponential family of distributions; estimation: method of maximum likelihood, method of least squares, estimation of generalized linear models; inference: sampling distribution for scores, sampling distribution for maximum likelihood estimators, confidence intervals for model parameters, adequacy of a model, sampling distribution for log-likelihood statistic, log-likelihood ratio statistic (deviance), assessing goodness of fit, hypothesis testing; multiple regression: maximum likelihood estimation, log-likelihood ratio statistic.

Models for binary responses: probability distributions, generalized linear models, dose response models, general logistic regression, maximum likelihood estimation and log-likelihood ratio statistic, other criteria for goodness of fit, least square methods; multinomial distributions; nominal logistic regression models; ordinal logistic regression models.

Models for count data, Poisson regression and log-linear models: probability distributions, maximum likelihood estimation, hypothesis testing and goodness of fit.

Textbooks

1. Dobson AJ and Barnett AG (2008). An Introduction to Generalized Linear Models, *3rd edition*. Chapman & Hall.

Reference Books

1. McCullagh P and Nelder JA (1989). Generalized Linear Models, *2nd edition*. CRC Press.
2. Agresti A (2007). An Introduction to Categorical Data Analysis. *2nd edition*. Wiley.

3. Generalized Linear Models and Extensions, *4th edition*.

AST409: Official Statistics **Credit 3**

Course Description

Official statistics is a branch of statistics that deals with the collection, analysis, interpretation, and dissemination of data produced by government agencies and other official bodies. Its purpose is to provide accurate and reliable information on various aspects of society, such as the economy, population, health, education, and crime. Official statistics are used for various purposes, including policy-making, research, business-decision making etc.

Course Objectives

The objectives of the course are to make students familiar with different sources of official statistics and their institutional, legal, and ethical aspects, and to prepare students with strong methodological basis for measuring, processing, producing, and using official statistics for national policies and planning.

Course Learning Outcomes (CLOs)

After completing this course, students are expected to

- CLO1 understand the fundamental concept of official statistics and their importance in practice
- CLO2 understand the methodological aspect of measuring, collecting and processing official statistics
- CLO3 be familiar with the legale and ethical aspects of using official statistics
- CLO4 know in details about the different types and sources of official statistics and their use in practice
- CLO5 develop skills to demonstrate the use official statistics for nataional policies and planning and explain it with different stakeholders

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3		1	2	2	2	2	2
CLO2	2	1	3	2	3	2	1	1
CLO3	1		1	2	2	2	1	3
CLO4	2		2	2	2	3	2	2
CLO5	2		1	2	3	2	2	2

Contents

Domains of official statistics, concepts, definitions, and statistical standards; fundamental principles of official statistics; code of practice of official statistics in Bangladesh; Legal and

ethical framework for official statistics; Sector-specific data quality assessment framework. Role of official statistics in monitoring the progress of global and national goals and policies. Data availability of the indicators of global and national goals and policies.

Use of administrative data in official statistics: sources of administrative data, collection, process, and use of administrative data in official statistics.

Use of survey data in official statistics; sources of surveys in Bangladesh; analysis of survey data for official statistics; use of survey weighting for official statistics.

Demographic Statistics: Population growth, techniques to measure it, doubling time concept in demography; population estimates and projections; different techniques of population projection- component method, arithmetic/linear method, geometric method, exponential method, matrix method, Lee-Carter model, Bayesian model etc., for population projections.

Health and well-being statistics: maternal and child health statistics, health services, general and chronic health statistics, health workforce.

Macro-Economy Statistics: GDP, GDP per capita, GDP growth, Inflation rate, Remittance etc. Business Economy Statistics: Short-and long-term business statistics; Science and technology; Environment and energy; Transport and travel; Labour Force Statistics: Income and social conditions; Employment, unemployment, and earnings.

Agricultural statistics: Food production and consumption (crop, forestry, and fisheries), Food Balance Sheet, Food insecurity experience sclse (FIES), land use statistics, and climate and environmental statistics.

Education statistics: Enrollment, dropout, completion, transition rates at different levels of education, teacher student ratio.

National income: concepts and methods of measurement; social accounting matrix. Distribution of income and wealth: Pareto and Lognormal distribution.

Poverty: concept, steps in measurement. Defining poverty lines: subjective and objective poverty lines. Measures of poverty: headcount, gap, squared gap (Foster-Greer-Thorbecke), Sen, Sen-Shorrocks-Thon, Watts poverty indexes. Robustness of indexes: sampling error, measurement error. Poverty mapping, poverty comparisons. Vulnerability to poverty: measuring and quantifying vulnerability. The effect of taxation on inequality and poverty. Multidimensional poverty index.

Inequality: concept, measurement of positive and normative measures; Lorenz curve; Gini coefficient; Atkinson's index, Theil's index, Human development index. Inequality comparisons.

Psychometry: concept, measurement; intelligent and achievement tests; z-score and t-score; intelligent quotient.

Official statistics for all SDG indicators; Data visualization and GIS mapping; time series of official statistics.

Textbooks

1. Linneman TJ (2014). Social Statistics: Managing Data, Conducting Analyses, Presenting Results. Routledge.
2. Stillwell J and Clarke M (2011). Population Dynamics and Projection Methods, *4th edition*, Springer.

3. Haughton J and Khandker SR (2009). Handbook on Poverty and Inequality. Washington, DC: World Bank.

Reference Books

1. Rahman PkMM (2016). Fundamentals of Social Statistics and Social Development. Osder.
2. Chadha NK (2009). Applied Psychometry. Sage.
3. NIPORT and ICF (2023). Bangladesh Demographic and Health Survey, 2022. Dhaka, Bangladesh and Rockville, Maryland, USA.

AST430: Statistical Computing IV **Credit 2**

Course Description

This course provides hands-on experience with different inferential procedures and multivariate techniques, focusing on their application to real-life problems. Students will learn to use statistical software (R package) to analyze data using principal component analysis, factor analysis, discriminant analysis, and canonical correlation analysis. The course aims to equip students with the skills necessary to interpret and present their findings clearly.

Course Objectives

The objectives of the course are to provide students with hands-on experience in applying different inferential procedures and multivariate techniques using statistical software (R package).Teach students how to extract and present the most important features from complex data. Finally enable students to interpret analysis results and prepare scientific publications.

Learning Outcomes

Upon completion of the course, students will be able to

CLO1 understand and explain the theoretical foundations and concepts of different multivariate analysis techniques, including principal component analysis, factor analysis, discriminant analysis, and canonical correlation analysis.

CLO2 analyze real-life data using multivariate techniques and inferential procedures, drawing connections among various ideas and concepts.

CLO3 evaluate the results obtained from multivariate analysis, justify decisions, and write scientific publications based on the analysis.

CLO4 apply statistical software (R package) to perform multivariate analysis and solve real-life problems effectively.

CLO5 develop new or original work by implementing advanced multivariate techniques and presenting significant findings clearly to stakeholders and researchers from other disciplines.

CLOs	Program Learning Outcomes (PLOs)							
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	3	2	1	1	2	1
CLO2	3	2	3	3	2	2	2	1
CLO3	2	2	3	3	2	2	2	1
CLO4	2	3	3	3	3	2	2	1
CLO5	2	2	3	3	2	2	2	1

Contents

Problems related to Inference and Multivariate analysis: principal component analysis, factor analysis, discriminant analysis, canonical correlation analysis, linear programming problem; Problems related to Machine Learning techniques: classification methods, resampling methods-different cross-validation and bootstrap, linear model selection procedures, regularized methods-lasso, elastic net, group lasso, and bridge, dimension reduction techniques, linear programing probleems.

AST431: Statistical Computing V **Credit 2**

Course Description

This statistical computing course focuses on real-life applications of the topics covered in the courses: time series analysis and econometrics methods. The course will equip students with the tools and techniques to analyze and interpret time series, and economic and financial data, make forecasts, estimate economic relationships, and contribute to economic research and policy analysis by integrating statistical software R, Stata, or Python.

Course Objectives

The course's main objective is to familiarize the students with statistical software and packages (R, Stata, or Python) to apply time series and econometric methods for analyzing data and solving real-life problems. The aim is to prepare students with the skills and knowledge necessary to pursue careers in economics, finance, data analytics, or research institutions where time series analysis and econometrics are essential.

Learning Outcomes

Upon completion of the course, students are expected to be able to

CLO1 comprehensive understanding of time series and econometric models, and the skill to interpret and evaluate results and make informed decisions based on their analyses.

CLO2 understand how to fit various time series and econometric model to real-life data and identify the best models.

CLO3 proficiently use of statistical software such as R, stata to perform different kinds of econometric modeling and visualization.

CLO4 proficiently use of statistical software such as R, stata to perform different kinds of time series analysis and visualization.

CLO5 discuss the use of statistical models related to time series and econometric data frame and solutions to the problems faced by the different stakeholders.

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	2	2	1				
CLO2	2	2	3	2				
CLO3	1	1	3	1	2			
CLO4	1		1	2	3			1
CLO5	1		1	2	1	3	3	1

Contents

Exploratory analysis of time series data, fitting time series models such as ARMA, ARIMA, etc. Fitting econometric models.

AST432: Statistical Computing VI **Credit 2**

Course Description

This course deals with the implementation of GLMs and survival models to real-life data from diverse areas such as medicine, biological science, engineering, and social science. In addition, this course demonstrates the use of statistical software and packages (e.g., R or Stata) to compute the models using real-life data.

Course Objectives

The objectives of the course are to teach students the application of a broader range of statistical models to analyze real-life data from diverse areas. Secondly, to equip students with skills to use statistical software and packages (e.g., R or Stata) to compute the models and discuss the models and results with other researchers and stakeholders.

Learning Outcomes

Upon completion of the course, students will be able to

CLO1 develop a solid understanding of the implementation of GLMS and survival models for their use in practice

CLO2 understand how to formulate and fit a model for a dataset and interpret the estimates of the model parameters

CLO3 develop computational skills to use statistical software for fitting the models to a dataset

CLO4 demonstrate skills to use these models to analyze data from diverse areas and discuss the findings with researchers from other disciplines

CLO5 discuss the use of statistical models for data-driven solutions to the problems faced by the industries and other stakeholders

	Program Learning Outcomes (PLOs)							
CLOs	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8
CLO1	3	1	2	1				
CLO2	2	1	3	2				
CLO3	1	1	2	2	3			
CLO4	1		2	3	2	1		1
CLO5	1		1	1	1	3	3	1

Contents

Exploratory analysis of categorical data, fitting glm: logistic regression models, binary logistic, multinomial logistic, ordinal logistic, log-linear models; Fitting survival models: AFT, Cox PH models, competing risks models, and frailty models.

AST440: Oral IV **Credit 2**

Each student must be examined by a committee of selected members at the end of the academic year.

AST450: B.S. Capstone Project **Credit 3**

Each student will be required to prepare a project report and present the report in a seminar. The project report is expected to reflect a student's ability to apply his or her knowledge of statistical methods in a nobel research work, which could be applied or theoretical. For the project work, each student will be assigned to a supervisor (a faculty member of ISRT) at the beginning of the academic year. The project report submission and presentation should be made before the final examination. The 50% weight of the course will be allotted to project work, 10% for supervisor, and the remaining 40% for seminar presentation. The internal members of the examination committee will evaluate the performance in the seminars, and the report will be evaluated by two internal examiners nominated by the examination committee. A supervisor cannot evaluate the project report s/he has supervised.
