E.G.S.Pillay Engineering College

An Autonomous Institution Affiliated to Anna University,
Chennai | Approved by AICTE, New DelhiAccredited by NAAC
with A++ Grade | Accredited by NBA T1 (B.E. - CSE, B.E. - ECE

Academic Year:2025-26 |Odd Semester|3rd SemesterB.TECH Programs |Regulation: 2023

Department of Artificial Intelligence and Data science

COURSEPLAN

COURSE SUMMARY

Course Code &Name:	2301MA301-Probability and Statistics
Semester:	III(03)
Credit Hours :	45(T)+ 15(TU)= 60Sessions|(LTPC:3104)
Course Duration:	July 2024toNov2024
Course Instructor/Coordinator :	Mrs.D.Anandhi (anandhi@egspec.org)
Domain Coordinator:	Dr.R.Deepa (deepa.ar@egspec.org)
Course Type:	Theory
Core/Elective:	Core
Prerequisites:	2301MA102 - Engineering Mathematics-I 2301MA202- Engineering Mathematics- II
Course Description :	The basic analytical concepts of probability theory, statistical design of experiments and data analysis Probability Distribution is a key concept in machine learning, data engineering, and artificial intelligence. It enables businesses to make informed decisions based on the data available, and is used to create predictive models that can make predictions about future events. Now that data science receives a lot of attention, the three disciplines of data analysis, databases, and sciences are discussed with respect to the roles they play. In several discussions, I observed misunderstandings of Artificial Intelligence. Hence, it might be the right time to give a personal view of AI and the part of machine learning therein.

	Since the relation between machine learning and statistics is so close that sometimes the boundaries are blurred, explicit pointers to statistical research are made.

Course Objectives:	1. To introduce the basic concepts of probability and random variables 2. To acquaint the knowledge of testing of hypothesis for small and large samples this plays an important role in real life problems 3. To impart knowledge of handling random vectors which represent random variables in multi-dimensional
Course Outcomes(Statements)CO1:	Use the fundamental concepts of probability and have knowledge of standard distributions which can describe real life phenomenon.(K3)
CO2:	Apply the basic concept of one and two dimensional random variables in engineering applications.(K3)
CO3:	Simulate the concept of testing of hypothesis for small and large samples in real life problems.(K3).
CO4:	Apply the basic concepts of classifications of design of experiments in the field of statistical quality control (K3).
CO5:	Develop exposure to the principal component analysis of random vectors and time series. (K3)
Mapping POs/PSOs	

CO #	PO 1	PO2	PO3	PO4	PO5	PO6	PO 7	PO 8	PO9	PO 10	PO 11		PSO1	PSO2	PSO3
1(L2)	3	2	1	-	-	-	-	-	-	-	1		-	-	-
2(L3)	3	2	1	-	-	-	-	-	-	-	1		-	-	-
3(L3)	3	2	1	-	-	-	-	-	-	-	1		-	-	-
4(L3)	3	2	1	-	-	-	-	-	-	-	1		-	-	-
5(L4)	3	2	1	-	-	-	-	-	-	-	1		-	-	-
·	3	2	1	-	-	-	-	-	-	-	1		-	-	-

Office Hours :	03Theory sessions |02Tutorialsessions (Ref TT)

Course Assessment &Grading Policy: (Theory Component)	CAT-I(50)	CAT-II(50)	Activity-I	Activity-II)	ESE(100)
			( )		

Over all Impact of the Course Outcomes i	Slight (low) : - POs (3,12)| PSOs (1,2)Moderate(Medium) :POs (2) Substantial(High) :POs(1)

n Attainment POs/ PSOs:						
Instruction Methods:	1. Problem solving 2. Discussion 3. Critical Thinking 4. Mind Mapping 5. Flipped Classroom					
(Exposure to) Modern Tools usage:	1. Power point Presentation 2. Statistical package					
Scope for Problem based learning :	Problem based learning;Specific toCO1toCO5,students will have to take up team work					
Scope for Innovative Teaching:	1. Flipped Classroom 2. Inquiry based learning					
Course Attainment(Benchmark| Target):	CO1	CO2	CO3	CO4	CO5	
Set Benchmark:	60	60	60	60	60	
Set Target:	65	65	65	65	65	

SESSION- WISE PLAN(CO1)

CO1: Use the fundamental concepts of probability and have knowledge of standard distributions which can describe real life phenomenon.(K3)	Bloom's level# :Bl3 Apply
Contents/Topics:Module 1:Random Variables and distributions Introduction:Course over view,Expectation sand learning outcomes Random variable - Distribution function - properties Probability mass function - Probability density function - moments Standard Distributions - Binomial, Poisson and Normal distributions.	

Mapping POs /PSOs

P O 1	PO2	P	P O	P	P O 6	P O 7	P O 8	P	PO 1 0	PO 1 1	PS O 1	PS	P S O 3
3	2	1	-	-	-	-	-	-	-	1	-	-	-

Contact Hours: (9+3)/60 (20% contributing to The course)	Benchmark/Target:60/65

Session	Cumulative Session No	Course Content/ Topics	Content Delivery Mechanism	Pedagogy
1	1.	Course overview, learning outcomes	Interactive lecture	Q&A
2	2.	Random variable	Interactive lecture,examples	Active learning, Inquiry-Based learning
3	3.	Distribution function andProperties	Interactive lecture,examples	Active learning
4	4.	Solving problems(Tutorial	Discussion	Practice
5	5.	Probability mass function	Interactive lecture,examples	Active learning
6	6.	Probability density function	Interactive lecture,examples	Active learning
7	7.	Moments	Interactive lecture,examples	Active learning
8	8.	Solving problems(Tutorial)	Discussion	Practice
9	9.	Binomial distributions	Interactive lecture,examples	Active learning
10	10.	Poisson distributions	Interactive lecture,examples	Active learning
11	11.	Normal distributions	Interactive lecture,examples	Active learning
12	12.	Solving problems(Tutorial)	Discussion	Practice

Assessment Strategy:	CAT-I(50)	CAT-II(50)	Activity-I(5)	Activity-II(5)	ESE(100)

Marks :	16	-	-	-	-
Other Assessment Strategy/ Tools:		Formative Assessment(Assignment)			
MeasuringCO1attainmentthrough:		CIE=	SEE =	CES=	T =

SESSION-WISE PLAN(CO2)

CO2:Apply the basic concept of one and two dimensional random variables in engineering applications.(K3)	Bloom's level# :Bl3 Applying

Contents/Topics:Module2:TWO DIMENSIONAL RANDOM VARIABLES

Joint distributions - Marginal and conditional distributions Expected values of functions of two variables Correlation and regression (for discrete data only) - Central limit theorem (Statement)

Mapping POs/ PSOs

P O 1	PO 2	P	P O	P	P	P O 7	P O 8	P	PO 10	PO 11	PS O 1	PS	P S O 3
3	2	1	-	-	-	-	-	-	-	1	-	-	-

Contact Hours:(9+3)/60(20% contributing to the course)	Benchmark/Target:60/65

Session	Cumulative Session No	Course Content/ Topics	Content Delivery Mechanism	Pedagogy
1	13.	Joint distributions	lecture, Interactive discussion	Active learning,Group discussions
2	14.	Marginal distributions	lecture,Interacti ve discussion, Real time Scenarios	Active learning,Group discussions
3	15.	conditional distributions	Interactive lecture,Real time Scenarios	Active learning,Group discussions

		0		0	0		0	
5	17.	Expected values of functions of two variables	Interactive lecture,Real time Scenarios	Active learning,Group discussions				
6	18.	Expected values of functions of two variables	Interactive lecture,Real time scenarios	Active learning,Group discussions				
7	19.	Solving problems(Tutorial)	Discussion	Practice				
8	20.	Correlation (for discrete data only)	lecture, Interactive examples	Active learning,Group discussions				
9	21	regression (for discrete data only)	lecture, Interactive examples	Active learning, Group discussions				
10	22	Correlation and regression problems	lecture, Interactive examples	Active learning,Group discussions				
11	23	Central limit theorem	Interactive lecture, examples	Active learning,Q&A				
12	24	Solving problems(Tutorial)	Discussion	Practice				
Assessment Strategy:		CAT-I(50)	CAT-II(50)	Activity-I(5)	Activity-II(5)	ESE(100)		
Marks :		18	-	5	-	20		
Other Assessment Strategy/ Tools:		Formative Assessment(Assignment)						
MeasuringCO2attainment through:		CIE=	SEE =	CES=	T =			

SESSION- WISEPLAN(CO3)

CO3: Simulate the concept of testing of hypothesis for small and large samples in real life problems.	Bloom's level# :Bl3 Applying

Contents/Topics:Module3:TESTING OF HYPOTHESIS

Sampling distributions - Tests for single mean, proportion and difference of means (large and small samples)Tests for single variance and equality of variances - Chi square test for goodness of fit Independence of attributes.

Mapping POs/PSOs

PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	P	PO11		PSO1	PSO2	PSO3
3	2	1	-	-	-	-	-	-	-	1		-	-	-

Contact Hours: (9+3)/60 (20%contributing to the course)	Benchmark/Target:60/65

Session	Cumulative Session No	Course Content/ Topics	Content Delivery Mechanism	Pedagogy
1	25.	Sampling distributions	Lecture, Interactive discussion	Active learning,Group discussions
2	26.	Tests for single mean	Lecture, Interactive discussion	Active learning,Q&A
3	27.	Tests for single proportion	Lecture, Interactive discussion	Active learning, Group discussions
4	28.	Test for difference of means	Lecture, Interactive discussion	Active learning,Group discussions
5	29.	Solving problems(Tutorial)	Discussion	Practice

6	30.	Tests for single variance	lecture, Interactive examples	Active learning,Q&A
7	31.	Test for equality of variances	lecture, Interactive examples	Active learning,Q&A
8	32.	Solving problems(Tutorial)	Discussion	Practice
9	33.	Chi square test for goodness of fit	lecture, Interactive examples	Active learning,Q&A
10	34.	Independence of attributes.	lecture, Interactive examples	Active learning,Q&A
11	35.	More problems on chi square test	lecture, Interactive examples	Active learning,Q&A
		Solving problems(Tutorial)	Discussion	Practice

Assessment Strategy:	CAT-I(50)	CAT-II(50)	Activity-I(5)	Activity-II(5)	ESE(100)
Marks :	16	16	-	-	20
Other Assessment Strategy/ Tools:		Formative Assessment(Quiz)			
MeasuringCO3attainmentthrough:		CIE=	SEE =	CES=	T =

SESSION-WISE PLAN(CO4)

CO4: Apply the basic concepts of classifications of design of experiments in the field of statistical quality control.	Bloom's level# :Bl3 Applying
Contents/Topics:Module4:DESIGN OF EXPERIMENTS One way and two way classifications Completely randomized design - Randomized block design latin square design - 2^2 factorial designs.	
Mapping POs/PSOs	

P O 1	PO2	P											
3	2	1	-	-	-	-	-	-	-	1	-	-	-
Contact Hours:(9+3)/60(20%contributing to the course)									Benchmark/Target:60/65				

Session	Cumulative Session No	Course Content/ Topics	Content Delivery Mechanism	Pedagogy
1	37.	One way classifications	Interactive lecture	Q&A
2	38.	One way classifications	Discussion	Active learning,Collabo r ative problem-solving
3	39.	two way classifications	Interactive lecture	Active learning, Collaborativ e problem-solving
4	40.	two way classifications	Interactive lecture	Active learning, Collaborativ e problem-solving
5	41.	Solving problems(Tutorial)	Discussion	Practice
6	42.	Completely randomized design	lecture, Interactive discussion	Active learning ,Q &A, Group discussions
7	43.	Completely randomized design	Interactive lecture	Active learning, Q&A, Group discussions
8	44.	Randomized block design	Interactive lecture	Active learning, Q&A, Group discussion s
9	45.	Solving problems(Tutorial)	Discussion	Practice

10	46.	Latin square design	Interactive lecture	Active learning, Q&A, Group discussions
11	47.	$2^{'}$ factorial designs.	Interactive lecture	Active learning, Q&A, Group discussions
12	48.	Solving problems(Tutorial)	Discussion	Practice

Assessment Strategy:	CAT-I(50)	CAT-II(50)	Activity-I(5)	Activity-II(5)	ESE(100)
Marks :	-	16	-	-	20

Other Assessment Strategy/ Tools:	Formative Assessment(Quiz)

Measuring CO4attainment through:	CIE=	SEE =	CES=	T =

SESSION-WISE PLAN (CO5)

CO5: Develop exposure to the principal component analysis of random vectors and Time Series.	Bloom's level# :Bl3 Apply

Contents/Topics:Module5:MULTIVARIATE ANALYSIS AND TIME SERIES

Random vectors and matrices - Mean vectors and covariance matrices -Principal components Population principal components - Principal components from standardized variables.Time series - components - Trend-Determination of trend by moving averages - least square method Seasonal Variations-Ratio to moving average method.

Mapping POs/PSOs

P O 1	PO 2	P	P O	P	P O 6	P O 7	P	P	PO 10	PO 11	P S O 1	PS	P S O 3
3	2	1	-	-	-	-	-	-	-	1	-	-	-

			0	0	0	0	0

Session	Cumulative Session No	Course Content/ Topics	Content Delivery Mechanism	Pedagogy
1	49.	Random vectors and matrices	Interactive lecture, examples	Active learning, Discussion
2	50.	Mean vectors and covariance matrices	Interactive lecture, examples	Active learning, Discussion
3	51.	Principal components	Interactive lecture, examples	Active learning, Discussion
4	52.	Solving problems(Tutorial)	Discussion	Practice
5	53.	Population principal components	Interactive lecture, examples	Active learning, Q&A
6	54.	Principal components from standardized variables	Interactive lecture, examples	Active learning, Q&A
7	55.	Time series - components	Interactive lecture, examples	Active learning ,Q &A
8	56.	Trend- Determination of trend by moving averages	Interactive lecture, Real time Scenarios	Active learning, Q&A, Group

				discussions
9	57.	Solving problems(Tutorial)	Discussion	Practice
10	58.	least square method	Interactive lecture, examples	Active learning ,Discussion
11	59.	Seasonal Variations	Interactive lecture, examples	Active learning ,Discussion

12	60.	Solving problems(Tutorial)	Discussion	Practice

Assessment Strategy:	CAT-I(50)	CAT-II(50)	Activity-I(5)	Activity-II(5)	ESE(100)
Marks :	-	18		5	20

Other Assessment Strategy/ Tools:	Formative Assessment(Assignment),Seminar Presentation

MeasuringCO5attainmentthrough:	CIE=	SEE =	CES=	T =

INTERNAL ASSESSMENT PLAN(Theory Part):

CO#/ Assessment		Marks in CAT-I			Marks in CAT-II			Marks in ACTIVITY-I			Marks in ACTIVITY- II			TOTAL MARKS
Bloom's level		2	3	4	2	3	4	3	4	5	3	4	5	
CO1	Part-A	4	-	-	-	-	-	-	-	-	-	-	-	16
	Part-B		12	-	-	-	-	-	-	-	-	-	-	
	TOTAL	4	12	-	-	-	-	-	-	-	-	-	-	
CO2	Part-A	06	-	-	-	-	-	-	-	-	-	-	-	18
	Part-B	-	12	-	-	-	-	05	-	-	-	-	-	
	TOTAL	06	12	-	-	-	-	05	-	-	-	-	-	
CO3	Part-A	06	02	-	-	-	-	-	-	-	-	-	-	32
	Part-B	-	24	-	-	-	-	-	-	-	-	-	-	

	TOTAL	06	26	-	-	-	-	-	-	-	-	-	-	
CO4	Part-A	-	-	-	04	-	-	-	-	-	-	-	-	16
	Part-B	-	-	-	-	12	-	-	-	-	-	-	-	
	TOTAL	-	-	-	04	12	-	-	-	-	-	-	-	
	Part-A	-	-	-	04	02	-	-	-	-	-	-	-	
CO5	Part-B	-	-	-	-	12, 1	-	-	-	-	-	-	-	18
	TOTAL	-	-	-	04	14	-	-	-	-	-	-	-	
	Part-B	-	-	-	-	-	-	-	-	-	05	-	-	
	TOTAL	-	-	-	-	-	-	-	-	-	05	-	-	
TOTAL (BL)		24	26	-	10	40	-	05	-	-	05	-	-	110
TOTAL		50			50			05			05			

Overall Quality of Question Paper as per Bloom's Taxonomy levels						
Bloom's level	1	2	3	4	5	6
Marks	-	34	66	10	-	-
% of Questions	-	30.90	60.00	09.09	-	-
		%	%	%		

PROPOSED ACTIVITY-I(CO2)

Apply the basic concepts of one and two dimensional

random variables in engineering applications.

Objective:

This assignment aims to apply discrete and continuous random variables for solving engineering and real life problems.

Students will be assessed on their ability to:

Acquire skill in handling situation with more than one random variable with time function.

Assignment questions:

1. If P(x, y)=k(2x+3y), x=0,1,2 and y=1,2,3. Simplify (i) P(x=0/y=2). (ii)Find P(x +y) and P(x +y)>3. Also find P(x≤1),P(y≤2), P(x≤1, y≤2)
2. If the joint p.d.f of x and y is given by f(x ,y)= $\cap\ y < 3$ Analyze (i) P(X<1 $\cap\ y < 3$).(ii) P(X<1 | Y<3). (iii) P(x +y<3)
3. A statistical survey on the Heights of fathers and sons are taken and the observed data's are given below.

Height of father	65	66	67	67	68	69	70	72
Height of Son	67	68	65	68	72	72	69	71

Correlate the heights by designing two dimensional random variable and discuss the nature of correlation coefficient.

4. If the joint density function of the random variable X and Y f(x , y)= 2-x-y, 0≤x≤1, 0≤y≤1. Examine the Correlation.
5. The joint density function of X and Y is $f(x,y) = \begin{cases} e^{-(x+y)}; 0 \le x \le \infty \\ 0, \text{elsewhere} \end{cases}$ Are X and Y are independent? Find $P(X < 1); P(X + Y < 1)$

This assignment provides an opportunity for students to apply their knowledge of

probability and Distribution concepts to analyze real-world problems and develop critical thinking skills.

Students' can gain valuable skills and knowledge for various fields.

Scheme of Evaluation

S.No	Content	Mark Allocation
1.	Solution for All Problems	8
2.	Presentation	1
3.	On-time Submission	1
Total		10

PROPOSED ACTIVITY-II (CO5)

Develop exposure to the principal component analysis of random vectors and Time Series.

Objective:

The objective of time series analysis is to understand the underlying structure and function of the data points collected over time and to make informed decisions based on this analysis.

Students will be evaluated on their capacity to:

- Test the hypothesis of single and difference mean
- Determine differences between research results from three or more unrelated samples or groups.

Assignment questions:

1 . i)Calculate the covariance matrix for the two random variable X and Y their joint probability function is represented by the entries in the following table

X/Y	X=0	X=1
Y=-1	0.24	0.06
Y=0	0.16	0.14
Y=1	0.40	0.00

(ii)Compute the principle components to the matrix A=

2.

(i) Compute the quadratic trend of the form $y=a+bx+cx^2$ for the data given below:

Years	1985	1986	1987	1988	1989	1990	1991
Production	10	11	12	9	10	13	11

(ii) Determine the equation of the straight line which best fits the following data:

Years	1984	1985	1986	1987	1988
Saleslin Rs.1000)	35	56	79	80	40

3.(i) Calculate the 3-yearly moving average of the data given below:

Years	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	
Sales (millions in Rs)	3	4	8	6	7	11	9	10	14	12	

Draw a graph to represent the moving averages ,also predict the sale for 1993.

(ii) Calculate the seasonal variation by the ratio to trend method from the data given below

Year	I Quarter	II Quarter	III Quarter	IV Quarter
1994	60	80	72	68
1995	68	104	100	88
1996	80	116	108	96
1997	108	152	136	124
1998	160	184	172	164

.

4.Assuming 5-yearly moving average calculate trend values from the data given below also plot the

Ye	19	19	19	19	19	19	19	19	19	19	19	19	19

ars	71	72	73	74	75	76	77	78	79	80	81	82	83
Production	105	107	109	112	114	116	118	121	123	124	125	127	129

5. (i)Compute the average seasonal movement for the following series

Year	Quarterly production			
	1	2	3	4
1974	3.5	3.9	3.4	3.6
1975	3.5	4.1	3.7	4.0
1976	3.5	3.9	3.7	4.2
1977	4.0	4.6	3.8	4.5
1978	4.1	4.4	4.2	4.5

(ii)Given below are the figures of production of a sugar factory

Year	1974	1975	1976	1977	1978	1979	1980
Production	77	88	94	85	91	98	90

Fit a straight by least a squares method and tabulate the trend values.

Scheme of Evaluation

S.No	Content	Mark Allocation
1.	Solution for All Problems	8
2.	Presentation	1
3.	On-time Submission	1
	Total	10

Course Co-ordinatorHOD