



# UNIVERSITÀ DEGLI STUDI DI PALERMO

<b>DEPARTMENT</b>	Scienze Economiche, Aziendali e Statistiche
<b>ACADEMIC YEAR</b>	2023/2024
<b>BACHELOR'S DEGREE (BSC)</b>	STATISTICS FOR DATA ANALYSIS
<b>SUBJECT</b>	LINEAR STATISTICAL MODELS
<b>TYPE OF EDUCATIONAL ACTIVITY</b>	A
<b>AMBIT</b>	50244-Statistico - probabilistico
<b>CODE</b>	23136
<b>SCIENTIFIC SECTOR(S)</b>	SECS-S/01
<b>HEAD PROFESSOR(S)</b>	CHIODI MARCELLO      Professore Ordinario      Univ. di PALERMO
<b>OTHER PROFESSOR(S)</b>	
<b>CREDITS</b>	9
<b>INDIVIDUAL STUDY (Hrs)</b>	141
<b>COURSE ACTIVITY (Hrs)</b>	84
<b>PROPAEDEUTICAL SUBJECTS</b>	19596 - STATISTICAL INFERENCE
<b>MUTUALIZATION</b>	
<b>YEAR</b>	2
<b>TERM (SEMESTER)</b>	2° semester
<b>ATTENDANCE</b>	Not mandatory
<b>EVALUATION</b>	Out of 30
<b>TEACHER OFFICE HOURS</b>	<b>CHIODI MARCELLO</b> Tuesday    15:00    17:00    stanza del docente (edificio 13); eccezionalmente su teams Friday      12:00    13:00    stanza del docente (edificio 13); eccezionalmente su teams

DOCENTE: Prof. MARCELLO CHIODI

<b>PREREQUISITES</b>	statistical inference, calculus, linear algebra and probability calculus
<b>LEARNING OUTCOMES</b>	<p>Conoscenza e capacita' di comprensione            Distinzione dei diversi tipi di correlazione (semplice parziale multipla);            Conoscenza del linguaggio e della terminologia dei modelli di dipendenza.            Comprensione della unicità di approccio all'analisi della dipendenza da variabili quantitative e qualitative            Comprensione dell'analisi della verosimiglianza per la costruzione di test e stimatori e dei diversi tipi di scomposizione della devianza e del concetto di varianza "spiegata"            Conoscenza delle tecniche di selezione dei modelli e di analisi dei residui</p> <p>Capacita' di applicare conoscenza e comprensione            Essere in grado di applicare la metodologia dei modelli di dipendenza a diversi casi di dati reali, provenienti da contesti diversi e con modalita' diverse di acquisizione dei dati            Capacita' di distinzione fra diverse tipologie di matrici del disegno in funzione della natura dei dati e delle variabili.            Capacita' di impiego del software specialistico (R) per l'analisi dei modelli lineari e non</p> <p>Autonomia di giudizio            Essere in grado di produrre risultati in relazione al fenomeno.            Giudicare l'utilita' dei modelli impiegati per analizzare relazioni tra variabili e analizzare criticamente i risultati utilizzando tecniche di analisi dei residui.            Capacita' di effettuare un processo di selezione di un modello statistico di dipendenza, integrando con le informazioni provenienti dall'analisi dei residui            Capacita' di giudicare il software da utilizzare per l'implementazione di particolari tecniche</p> <p>Abilita' comunicative            Essere in grado di redigere un rapporto tecnico;            scegliere le rappresentazioni grafiche piu' opportune in relazione ai problemi;            esporre verbalmente il contenuto del rapporto sia in forma sintetica che in forma estesa, in funzione delle conoscenze statistiche del destinatario            Capacita' d'apprendimento            Essere in grado di consultare la letteratura scientifica sull' argomento; capacita' di apprendere le estensioni dei modelli studiati a lezione; capacita' di apprendimento di software statistico specialistico anche diverso da quello impiegato in aula.</p>
<b>ASSESSMENT METHODS</b>	written and oral exam
<b>EDUCATIONAL OBJECTIVES</b>	<p>Educational Objectives            The main objective is to give the fundamental instruments to build and to analyze statistical models founded on linear model and on general linear models also using the software R</p>
<b>TEACHING METHODS</b>	<p>lessons, practical exercises and real data analysis.            Lectures, exercises and real data analysis with open source software. Most of the topics are introduced directly from real data. During the course, the teacher will share with the students a short article, a book chapter, or a part of it in English, which will be the subject of debate with the students.</p>
<b>SUGGESTED BIBLIOGRAPHY</b>	<p>Draper-Smith, Regression Analysis, John Wiley            Per maggiori approfondimenti verranno indicate le fonti bibliografiche piu' utili durante il corso.</p> <p>Materiale didattico            Il materiale didattico (appunti, software, casi di studio, dati e codice R) sara' a disposizione degli studenti sul web</p> <p>Verra' anche usato un software didattico ossia il package R scritto dal docente MLANP            che nella versione più aggiornata insieme a vari notebook didattici è disponibile sul sito del docente <a href="http://www.marcellochiodi.com">www.marcellochiodi.com</a> o sulle pagine istituzionali</p>

## SYLLABUS

Hrs	Frontal teaching
2	Introduction and introductive real problems
6	multiple statistical variables, data matrices. correlation among many variables. partial correlation, principal component analysis
4	bivariate normal distribution. Multivariate normal distribution (outlines), regression functions and conditional distributions
4	regression function: linear model. systematic and random components; design matrix X; repeated observations, factorial designs, polynomial regression, dummy variables, interactions

## SYLLABUS

Hrs	Frontal teaching
8	basic assumptions in linear model; likelihood; LR test; ML estimators; Ordinary least squares; sampling distributions of estimators; deviance components. Hypothesis test. General F test. Confidence interval for prediction errors
6	Multiple regression. test of specific hypothesis. test for regressor deletion, Collinearity. variable selection. AIC prediction errors and cross validation
4	residual analysis
6	linear models with categorical explicative variables: one way and two way analysis of variance; divergence from linearity; covariance analysis. residual analysis
8	GLM: generalized linear models. link functions and exponential family. logistic regression, poisson regression and gamma regression. parameter estimation and model selection
Hrs	Practice
8	multiple statistical variables, data matrices. correlation among many variables. partial correlation, principal component analysis
2	bivariate normal distribution. Multivariate normal distribution (outlines), regression functions and conditional distributions
20	Exercises on linear models with R: estimation, test and variable selection
6	logistic regression, poisson regression and gamma regression. parameter estimation and model selection with R