



UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Scienze Economiche, Aziendali e Statistiche		
ACADEMIC YEAR	2019/2020		
BACHELOR'S DEGREE (BSC)	STATISTICS FOR DATA ANALYSIS		
INTEGRATED COURSE	SOCIAL STATISTICS I - INTEGRATED COURSE		
CODE	18159		
MODULES	Yes		
NUMBER OF MODULES	2		
SCIENTIFIC SECTOR(S)	SECS-S/05		
HEAD PROFESSOR(S)	BOSCAINO GIOVANNI	Ricercatore	Univ. di PALERMO
OTHER PROFESSOR(S)	BOSCAINO GIOVANNI	Ricercatore	Univ. di PALERMO
CREDITS	12		
PROPAEDEUTICAL SUBJECTS	01736 - PROBABILITY THEORY 20581 - EXPLORATORY STATISTICS 16127 - MATHEMATICS		
MUTUALIZATION			
YEAR	2		
TERM (SEMESTER)	2° semester		
ATTENDANCE	Not mandatory		
EVALUATION	Out of 30		
TEACHER OFFICE HOURS	BOSCAINO GIOVANNI Monday 12:00 14:00 Dipartimento SEAS, edificio 13, Il piano lato ascensore. Per coordinare meglio gli impegni degli studenti e quelli imprevisti del docente, si invitano gli studenti a contattare il docente via email in modo da concordare il giorno e l'orario dell'appuntamento (anche differente da quello qui riportato).		

PREREQUISITES	The course needs the student knows descriptive statistics, inferential statistics, probability, and basics of mathematics and informatics.
LEARNING OUTCOMES	KNOWLEDGE AND UNDERSTANDING Acquisition of the "culture" and the specific language of the discipline in order to: 1) adapt language to different contexts; 2) identify relationship between two or three categorical variables; 3) use SAS in an elementary way. Students are required knowledge and understanding of specific issues of the discipline and of the mathematical language used of the formalisation of the same. Students easily achieve this by adapting (or newly acquired) the previous knowledge through participation in lectures and/or consultation of the reference books. APPLYING KNOWLEDGE AND UNDERSTANDING The ability to apply the learned lessons should allow to face abstract and real situations. Students will 1) identify the suitable statistical tool, test, and measure with respect to the variable measurement scale, in order to summarise in a simple way the relationship among variables; 2) manage data using SAS software; 3) be able to explain and comment results; 4) friendly use SAS as tools for generating tables and plots. Skills are acquired through a work of individual and collegial exercise class coordinated by the Teacher. MAKING JUDGEMENTS Judgement and reflection must allow independent choices about the settings to follow even without direct guidance. Student will be able to analyse results also in a critical point of view, highlighting strength and weakness of adopted methodology and results. This will be used in each step of the learning process by also following a critical approach to the provided stimuli. COMMUNICATION SKILLS Students will get specific languages and rigorous characteristic of each formalised aspect of the discipline. The aim is to give the student the ability of adapting his/her speech (also in writing) to the spokesperson, usually a not-statistician. The habit of speech and the exchange of information should take place by forcing the use of correct forms of communication, during exercise class. LEARNING SKILLS Students are constantly encouraged to acquire a critical view on issues such as to allow the passage from theoretical preparation to empirical application of the same in a more autonomous as possible. Students will be able to understand basic scientific literature; to increase their knowledge attending higher-level courses and/or on their own; to adapt their knowledge to different contexts and interlocutors. The dialectic confrontation with classmates and with the teacher will be used for the expansion of this skill.
ASSESSMENT METHODS	CATEGORICAL DATA MODULE Student's assessment is based on oral exam and written test. CATEGORICAL DATA MODULE TEST The written test aims to check student's knowledge, abilities (with respect to excel too) and his/her talent in reporting them in a writing along with an adequate statistical language. The test lasts 1 hour. It considers at most 3 open questions (at most 3 sub-questions for each question), both practical and theoretical. Questions will be formulated in order to be easy understandable, clear, with different difficulty, and with results easily comparable among students. The pass-mark (expressed by a mark of 18 on a 0-30 scale) will be reached when the student will show a basic use of the core concepts/terms of the module, and i) the identification of the suitable methodology even in case of mere calculus errors, for the practical question; and ii) a basic (not fully argued) answer statistically coherent to the theoretical question. CATEGORICAL DATA MODULE ORAL EXAM Once the student passes the written test, he/she will access to the oral exam. Oral exam consists of at least 2 questions aimed to better graduate the assessment of the student's knowledge and abilities and his/her talent in reporting them by a statistical speech (it will complete the outcome of the test). Questions could regards a practical exercise too. Pass-mark will be reached when the student will show a basic use of the core concepts/terms of the module, and i) the application of the suitable methodology even in case of mere calculus errors, for the practical question; and ii) a basic (not fully argued) statistically coherent answer to the theoretical question. Better

	he/she will perform, higher will be the assessment. CATEGORICAL DATA MODULE FINAL MARK Final mark will take into account the outcomes of both written test and oral exam, in particular: i) mastery of course topics; ii) ability in applying knowledge; iii) correct use of statistical language. The examiners could also take into account some student context characteristics (disability, or level of participation during lessons). Test and oral exams are valued on a 0-30 scale, with 18 as pass-mark: the whole mark for the module is computed as the simple mean of the test and oral exam marks. SAS LABORATORY MODULE The verification of learning consists of a written and oral test. SAS LABORATORY WRITTEN TEST The written test aims to detect the knowledge, skills possessed by the student (including preliminary education). The test will consist of a set of multiple-choice questions and programming questions. The threshold of sufficiency (equal to a score of 18 on a scale of 18-30), necessary to pass the written test of the module, is reached if the student shows knowledge and ability to use at least the basic procedures of SAS. The more the student demonstrates awareness of the SAS procedures and their proper use, the higher will be the evaluation. SAS LABORATORY ORAL TEST The oral test consists of a commentary on the written test, aimed at deepening the student's knowledge with questions aimed at better grading the evaluation of knowledge, skills, abilities and transversality with the topics of the courses taken previously, possessed by the student, as well as his ability to convey them with appropriate statistical language, in addition to what already emerged from the written test. The test consists in two parts: 1) a questionnaire with 10 multiple-choice questions (1 point for each correct answer); 2) the writing of a SAS program based on 3 tasks (one of them regards macro-variable and importing a dataset) that can give 20 points maximum. The more the student has given evidence of his argumentative and expository skills, as well as his statistical language skills and connections with other disciplines, the more positive the assessment will be. METHOD OF FINAL EVALUATION OF THE SAS LABORATORY MODULE The final assessment of the examination will take into consideration three aspects: (i) the mastery of the subjects; (ii) the ability to apply the knowledge; and (iii) the property of language, assessed in the written and oral test as a whole. The teacher will also have the opportunity to take into account the contextual factors of the examination (such as active participation during lessons and exercises, or the presence of any disability) in determining the outcome of the test. The resultant grade for the module is the simple arithmetic mean of the two tests. Final Passing Mark Final passing mark for the whole course is in [18; 30]: it comes from the simple mean of the final marks of the two modules. Some notes: Student can take the exam of just one single module at time. The passing assessment will be in charge till the end of the academic year during the exam has been performed. After that date, if he/she has passed the exam of the other module will have to redo the exam of the previous module. If the student wants to take just one module exam, test lasts one hour and a half. During the written test cellular phones, smartphones, and own notebook/tablet will not allowed, such as internet access and formulary, under penalty of invalidation of the exam. Student can withdraw from the exam (oral and written) at any time. If the student does not pass the exam, he/she can do it again at next scheduled exam. English questions will be posed when both modules are passed.
TEACHING METHODS	The course will be held with lectures and exercises on the blackboard and PC, actively engaging students to the resolution of the questions. SAS for real data analysis. The text of the SAS practice tests will be in English, to stimulate the student to acquire SAS terminology.

MODULE CATEGORICAL DATA

Prof. GIOVANNI BOSCAINO

SUGGESTED BIBLIOGRAPHY

Dispensa di dati categoriali fornita durante il corso

B.S. Phillips (1972) Metodologia della ricerca sociale. Societa' editrice Il Mulino. Cap. X pagg. 271-357

Kendall and Gibbons (1990) Rank correlation and methods. Edward Arnold. Cap. 4 e 6

Agresti (2007) An introduction to categorical data analysis. John Wiley. Cap. 1, 2, 4 (4.4 e 4.5 esclusi)

AMBIT	50250-Statistico, statistico applicato, demografico
--------------	---

INDIVIDUAL STUDY (Hrs)	98
-------------------------------	----

COURSE ACTIVITY (Hrs)	52
------------------------------	----

EDUCATIONAL OBJECTIVES OF THE MODULE

The aims of the module regard the analysis of categorical data, used mostly in a socio-economic-biologic field. During the course, many techniques, statistics, and tests will be taught taking into account the data drawing scheme too. the aim is to teach the student different tools for different situations. Finally, students will be more friendly with Excel worksheet.

SYLLABUS

Hrs	Frontal teaching
1	Introduction: Programme, aims, exams procedures. Basics of measurements.
3	rank correlation: kendall's tau
2	Rank correlation for more than 2 rankings: Kendall's W
2	Inference for rank correlation methods
2	Two ways tables: relationship among rows and columns. Symmetric and asymmetric relationship: data drawn scheme.
5	2x2 tables from 2 binomial independent samples: proportion comparison via Relative Risk, Odds, Odds Ratio and inference. independence in a 2x2 contingency table: Pearson χ^2 and residuals
4	lxJ contingency table independency test: M^2 for ordinal variables. Fisher exact test.
6	3 ways tables: Simpson's paradox, marginal and partial association, odds, and odds ratio. Conditional association and independence. Cochran-Mantel-Haenszel test and statistics.
7	Logit model, with metric and non-metric covariates. Estimation, inference and meaning of logit regression, with respect to a 2x2 table
Hrs	Practice
5	Exercises on cograduation with and without ties, on 2 rankings and on more than 2 rankings. Proportions comparison for two binomial independent samples.
10	Pearson χ^2 ; exercises on mutual, conditional, and marginal independency and on homogeneous association.
5	Logit models: definition, estimations, inference with respect to a 2x2 table.

MODULE
"SAS" LABORATORY

Prof. GIOVANNI BOSCAINO

SUGGESTED BIBLIOGRAPHY

Guida online di SAS: <http://support.sas.com/documentation/94/index.html>
SAS Certification Prep Guide: Base Programming for SAS 9 - SAS PRESS
SAS Certified Specialist Prep Guide - Base Programming Using SAS 9.4 - SAS Documentation
Materiale fornito dal docente

AMBIT	50250-Statistico, statistico applicato, demografico
INDIVIDUAL STUDY (Hrs)	98
COURSE ACTIVITY (Hrs)	52

EDUCATIONAL OBJECTIVES OF THE MODULE

The main objective of the course is to introduce students to basic SAS programming and the conscious use of SAS Studio software. At the end of the course, the student will be able to apply the basic procedures of the software for data management and manipulation and elementary statistical analysis.

SYLLABUS

Hrs	Frontal teaching
2	Introduction
10	SAS basics: tools, instruments, data managing, variates managing
10	Basic SAS procedures and their options
10	How to write a report

Hrs	Practice
6	SAS basics: tools, instruments, data managing, variates managing
6	Basic SAS procedures and their options
8	writing a report