



UNIVERSITÀ DEGLI STUDI DI PALERMO

DEPARTMENT	Scienze Economiche, Aziendali e Statistiche		
ACADEMIC YEAR	2016/2017		
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 06647 - STATISTICS 1 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.
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) perform a basic survey sampling and build a questionnaire. 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 school 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 situations and examples drawn from everyday life and consolidated use of discipline. Students will 1) identify the suitable statistical tool, test, measure with respect to the variable measurement scale, in order to summarise in a simple way the relationship among variables; 2) plan the suitable survey with respect to the objectives; 3) be able to explain and comment results; 4) friendly use Excel sheets as tools for managing data in order to reach their statistical goal. 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	Student's assessment is done through an oral exam. The oral exam is subject to passing the written test (its outcome will be considered for the final mark). The exam will be chaired by the chair of the course. The course consists in two modules: first module will be developed during the first part of the semester, and then followed by the second module. Therefore, the midterm exam will be established at the end of the first module. TEST The written test (midterm one too) aims to check student's knowledge, abilities (with respect to excel too) and his/her talent in reporting them in a text along with an adequate statistical language. The test lasts 3 hours. It considers the two course modules: 2 open questions for each module (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. Midterm test will regard just first module and its duration is 1 hour and a half: 2 open questions (at most 3 sub-questions each), 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.

	ORAL EXAM Once the student passes the written test, he/she will access to the oral exam. Oral exam regards both course modules and it will examine in depth the test and further student's knowledge. Oral exam consists of at least 2 questions aimed to better graduate the assessment of the student's knowledge and abilities (also to made links to others subjects) and his/her talent in reporting them by a statistical speech (it will complete the outcome of the test). Questions could regards a practical example. 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. FINAL MARK FOR THE MODULE Final mark will consider of the outcomes of both written test and oral exam and English language performance, taking into account: 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. FINAL MARK FOR THE COURSE 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. With respect to Course Degree Rules, two more questions will regards the student's knowledge of English Language. Study reference books lists of the two course modules consider also two brief English texts: the student will translate in Italian language a few sentences (selected by the examiner) of one of them and answer in English to a question about one of them. The aim is keeping the student trained in the use of English. The outcome will give the student 1 mark bonus to the final mark of the course. 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, under penalty of invalidation of the exam. Students can use their own formulary just for the Survey Samopling module. 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 on PC, actively engaging students to the resolution of the questions.

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. 1 pagg. 1-24, cap 2 pagg 25-28, cap 3 pagg 40-50, cap 6 pagg 117-120

B.S.Everitt (1994) The analysis of contingency table. Chapman & Hall. Pagg. 1-17, 60-67

Agresti (1996) An introduction to categorical data analysis. John Wiley. Pagg. 1-30, 39-41, 53-64, cap. 4 e 5 fino a pag. 118

Colett (1994) Modelling binary data. Chapman and Hall. pagg. 1-74

Testo per la prova in inglese: Kendall and Gibbons (1990) Rank correlation and methods. Edward Arnold. Cap. 1 da 1.1. a 1.6 (incluso)

AMBIT	50250-Statistico, statistico applicato, demografico
INDIVIDUAL STUDY (Hrs)	94
COURSE ACTIVITY (Hrs)	56

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.

SYLLABUS

Hrs	Frontal teaching
1	Introduction: Programme, aims, exams procedures. Basics of measurements.
5	Rank correlation methods: how to order? Kendall Tau, Tau as agreement and accuracy measure. Spearman rho vs Tau. Generalized correlation coefficient. Tau with tied rankings. Rank correlation for >2 rankings: Kendall W.
4	Two ways tables: relationship among rows and columns. Symmetric and asymmetric relationship: data drawn scheme.
4	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 residual analysis.
4	IxJ 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-Haentzel test and statistics.
6	Logit model, with metric and non-metric covariates.
2	Estimation, inference and meaning of logit regression, with respect to a 2x2 table. goodness of fit statistics.
Hrs	Practice
7	Exercises on cograduation with and without ties, on 2 rankings and on more than 2 rankings. Proportions comparison for two binomial independent samples.
6	Pearson χ^2 and residuals analysis; exercises on mutual, conditional, and marginal independency and on homogeneous association.
3	Logit models: definition, estimations, inference with respect to a 2x2 table.
8	introduction to SAS software: logic, main concepts, data management. Main procedures for categorical data.

MODULE SAMPLE SURVEYS AND OPINION POLLS

Prof. GIOVANNI BOSCAINO

SUGGESTED BIBLIOGRAPHY

Cicchitelli, Herzel, Montanari (1992) Il campionamento statistico. Il Mulino.
Cochran (1977) Sampling Techniques. Wiley & Sons.
Dispense del corso
Module teacher notes.

Testo per la prova in inglese: Cochran (1977) Sampling Techniques. Wiley & Sons. Cap.1 - par. 1 e par. 2

AMBIT	50250-Statistico, statistico applicato, demografico
INDIVIDUAL STUDY (Hrs)	94
COURSE ACTIVITY (Hrs)	56

EDUCATIONAL OBJECTIVES OF THE MODULE

The main aim of this module regards basic statistical methods for survey sampling (questionnaire, dimension, and schemes), estimation of an unknown parameter of a population (unknown) and its standard error. At the end of the module, students will be able to perform a simple survey and collect data.

SYLLABUS

Hrs	Frontal teaching
1	Introduction: aims, exams procedures.
1	Introduction to survey sampling: history and sampling from finite population.
6	Simple survey sampling: estimation of the mean, proportion, and total. Setting optimum sampling dimension.
15	Stratified survey sampling: estimation of the mean, proportion, and total. Allocation: proportional, equal, and optimum.
4	Comparisons between simple and stratified sampling (for the three different allocations).
5	Collecting data: questionnaire building and administration.
Hrs	Practice
6	Simple survey sampling: estimation of the mean, proportion, and total. Setting optimum sampling dimension.
14	Stratified survey sampling: estimation of the mean, proportion, and total. Allocation: proportional, equal, and optimum.
4	Questionnaire building with examples.