MSc Human Nutrition Sciences (LM-61)

Integrated course: **Statistics and Processing**
SSD: **MED/01** e **ING/INF-05**
CFU: **6**
Responsible teacher: **Montanari Paolo**   e-mail: paolo.montanari@unicamillus.org
(https://www.unicamillus.org/it/personnel/montanari-paolo)

Course: **Medical statistics and epidemiology**
SSD: **MED/01**
CFU: **3**
Teacher: **Montanari Paolo**   e-mail: paolo.montanari@unicamillus.org
(https://www.unicamillus.org/it/personnel/montanari-paolo)

Course: **Data and Information management**
SSD: I**NG/INF-05**
CFU: **3**
Teacher: **Montanari Paolo**   e-mail: paolo.montanari@unicamillus.org
(https://www.unicamillus.org/it/personnel/montanari-paolo)

PREREQUISITES
Elementary concepts of mathematics

EDUCATIONAL OBJECTIVES
Provide students with the knowledge and skills necessary to identify the most suitable statistical tools for the study they want to conduct. Make students autonomous in conducting the main statistical analyses and the main statistical tests thanks to the use of the IT skills acquired during the course. Provide students with a broad knowledge of computer terminology, the main hardware and software components of computers, their functioning and their fields of application. Make students autonomous in the use of spreadsheet management software.

EXPECTED LEARNING OUTCOMES
Knowledge and understanding
At the end of the course the student will have to:
1. know the methods and tools of descriptive statistics and inferential statistics;

2. know the various frequency distributions and have understood the main differences between the distributions and know how to exploit the properties;

3. understand what a statistical test is for and know the main statistical tests;

4. know and understand the fundamental methodological aspects of epidemiology;

5. know the tools of epidemiology and their fields of application;

6. have an idea of what a computer is and what its main uses are;

7. have knowledge of the main hardware components, storage devices and input and output devices;

8. understand the difference between system software and application software;

9. know what the main application software are;

10. understand what a database is and what it is used for and what its fundamental parts are;

11. understand what an Information System is;

12. have a basic understanding of cybersecurity.

Ability to apply knowledge and understanding

At the end of the course, the student will have to:

1. know how to use the methods and tools of descriptive statistics and inferential statistics;

2. know how to exploit the properties of frequency distributions;

3. knowing how to use the various statistical tests and knowing how to identify the most appropriate one for the study you want to conduct;

4. know how to identify the most appropriate epidemiological indicators depending on the type of analysis;

5. know how to use the techniques and tools of basic epidemiology in different types of scientific studies;

6. be able to read and use epidemiological studies and evaluate their results based on the interpretation of available information;

7. know how to identify the main hardware components present within a generic computer, the storage devices and the input and output devices;

8. know how to identify the field of use of the main application software;

9. have a fair amount of dexterity with the fundamental operations that can be performed with a program for managing spreadsheets.

Course program "Medical statistics and epidemiology":

Introduction to statistics: descriptive statistics; inferential statistics; terminology; qualitative nominal and ordinal characters; discrete and continuous quantitative characters. (video lesson 1)

Data representation: representation by statistical unit-modality; mode-frequency representation; absolute, relative, cumulative absolute, cumulative relative frequencies; relative frequency distribution; graphical representation of data; group data into classes. (video lessons 2-3)

Indicators for the description of the distributions: central trend indices: arithmetic average, weighted average, average for grouped data, geometric average, mode; position indicators: median, median for pooled data, quartiles and percentiles; variability indicators: range, deviance, variance, standard deviation,

standard deviation for pooled data, coefficient of variation, interquartile deviation; 5-number summary and boxplot. (video lessons 4-7)

Distributions: distributions of observations; symmetric and asymmetric distributions; shape indices of a distribution: asymmetry (skewness) and kurtosis (kurtosis); Gaussian distributions $N(\mu, \sigma)$ and standard Gaussian $N(0,1)$; standardization; find the proportion given an interval and find the interval given a proportion; Student's T distributions at different degrees of freedom. (video lessons 8 - 11)

Samples and inference: sample mean, variance and standard deviation; inference and errors in the inference process; distribution of the sample mean and central limit theorem; standard error of the sample mean; confidence levels; confidence intervals for the mean; sampling techniques: fraction sampling, simple random sampling, systematic sampling, stratified sampling, quota sampling and cluster sampling; generation of random numbers with uniform distribution or with Gaussian distribution of mean and standard deviation fixed by MS Excel functions. (video lessons 12-14)

Correlation, interpolation and regression: scatter plot; covariance; linear correlation coefficient; interpolating curves, residuals and least squares curves; line of least squares; coefficient of determination. (lessons 15-17)

Statistical tests for hypothesis testing: definition of probability and its interpretation as a limit of relative frequency; statistical tests; hypothesis H0 and H1; p-value and level of significance; 1st and 2nd type errors; power of the test; decision rule; operational sequence for conducting a statistical test; parametric and non-parametric tests; summary scheme of the types of tests; z-test and t-test for checking whether an observation belongs to a population. (video lessons 18-19)

Test for the verification of the association between two characters: test for the verification of the association between quantitative characters based on the linear correlation coefficient of Bravais-Pearson; test for the verification of association between characters on an ordinal scale based on the Spearman rank correlation coefficient; examples using MS Excel. (video lessons 20-21)

Chi-square test for frequency analysis: observed frequencies and expected frequencies; verification of frequency homogeneity, verification of goodness of adaptation of the empirical distribution to the expected trend of the theoretical distribution, verification of the association between two characters; correction of Yates in the case of only one degree of freedom; Fisher's exact test; McNemar test for paired data (before and after treatment); I use Excel functions. (video lessons 22-26)

Test for the comparison of medians: transformation of observations into ranks; test for the comparison of the medians of unpaired observations (U of Mann - Whitney, K of Kruskal-Wallis); test for the comparison of the medians of paired observations (Wilcoxon's T); examples using MS Excel. (video lessons 27-29)

Test for comparison of the averages: t-test for comparison of averages of unpaired and paired observations; examples using MS Excel. (video lesson 30)

Elements of epidemiology: definitions, objectives and study models; recall of mathematical tools: ratios, proportions, rates, odds; disease frequency measures: prevalence, cumulative incidence and incidence rate, odds; static cohort and dynamic cohort; measures of association: relative risk (RR) and odds ratio (OR), interpretative scale, confidence intervals; prospective studies (or cohort studies or longitudinal studies): epidemiological studies and clinical trials, single blind and double blind, hazard ratio; retrospective studies (or case-control studies). (video lessons 31-35)

Complements of epidemiology: confounding; standardization; screening tests and diagnostic tests; sensitivity, specificity, predictive value of positive test result (PPV), predictive value of negative test result (VPN); relationship between prevalence, PPV and VPN. (video lesson 36)

Course program "Data and information management":

Introduction to the world of computers: what is a computer, how computer works; terminology; main operations; a look at the computer history; The main types of computers; an introduction to hardware, main components: input, processing, output, memory and communication devices; an introduction to software: system software and application software; computer network and the Internet; computer and society. (video lessons 1-3)

The language of computers: data and program representation; representing numerical data: the binary numbering system; coding systems for Text-Based Data and other types of data; Representing Software Programs: Machine Language. (video lesson 4)

Hardware: inside the system unit: motherboard, CPU, GPU, memory, bus, expansion cards, ...; peripheral devices; the system clock and the machine cycle; strategies to improve the performance of a computer. (video lessons 5-7)

Storage systems: storage systems characteristics; hard drives; optical discs and drives; flash memory storage systems; network and cloud storage systems; smart card; holographic storage; storage systems for large computer systems. (video 8-10)

Input and Output devices: pointing and touch devices; scanners, readers, digital cameras; audio input systems; display devices; printers; audio output devices. (video lessons 11-13)

System Software (Operating Systems and Utility Programs): System Software vs. Application Software; functions of an Operating System; differences among Operating Systems; Operating Systems for personal computer and servers; Operating Systems for mobile devices and larger computer; utility programs: types and functions; the future of Operating Systems. (video lessons 14-16)

Application Software: software ownership rights; desktop and mobile software, installed and cloud software; main types of application software: word processing, spreadsheet, database management systems, graphics and multimedia software, other types of application software. (video lessons 17-20)

Database: introduction and definitions; entities and relationships; data definition; data dictionary; data integrity, data security, data privacy; data organization; types of DBMS; database models; relational model; tables, forms, queries, reports. (video lessons 21-23)

Artificial Intelligence Systems: introduction and definitions; intelligent agents; expert systems; robotics. (video lesson 24)

Information Systems and system development: responsibility for system development; outsourcing; the system development life cycle; approaches to system development. (video lessons 25-26)

IT security: definitions; unauthorized access and unauthorized use; protecting against unauthorized access and unauthorized use; computer sabotage; protecting against computer sabotage; access systems based on the use of biometric data; firewall; encryption; private key cryptography; public key cryptography; virtual private networks (VPN); online theft, online fraud and other dot cons; protecting against online theft, online fraud and other dot cons. (video lessons 27-30)

Exercises with spreadsheets: definitions and tools; basic operations; formulas; relative and absolute cell references; functions; statistical functions; date functions; text functions; nested functions; formatting; graphics and their customization; settings for printing; data transposition; cell comments. (video lessons 31-36)

Teaching mode:

Distance learning. For each course, there are 36 video lessons of approximately 15 minutes each, which the student can access from the course web portal and follow in asynchronous mode. At the end of each video lesson there are multiple choice learning verification quizzes. In addition to the video lessons, 3 hours of interactive teaching in synchronous mode are provided for each course.

Learning assessment methods:

For each course, the exam consists of a written test in the form of a questionnaire of 31 multiple choice questions (5 options of which only one is correct). Each correct answer: 1 point; each incorrect or omitted answer: 0 points.

Recommended texts and bibliography

Fowler J. , Jarvis P., Chevannes M. *Statistica per le professioni sanitarie* Edises;

Italo D'Ascanio *Statistica sanitaria pratica* Società Editrice Universo;

Paolo Chiari, Daniela Mosci, Enrico Naldi, *Evidence-Based Clinical Practice. La pratica clinico-assistenziale basata su prove di efficacia* 2/ed, McGraw-Hill

Deborah Morley and Charles S. Parker, *Understanding Computers: Today and Tomorrow (16th edition)* - Cengage Learning

Grassi R., Pinto G., Serra N. *Sistemi per l'elaborazione dell'informazione* Edises