UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistics - 2

2021-1-E3301M192-T2

Learning objectives

Economic disciplines study a variety of phenomena often showing different features. This course provides the main statistical methods to collect, represent, synthetize and analyze data for such phenomena.
Students will learn how to select and apply the suitable statistical method to describe single phenomena and/or to interpret their relations.

Contents

The course provides the main tools for synthesizing the main features of statistical data and for analyzing the relations between them.

The meaning of Statistics

Statistics as a science

Applications of Statistics

The branches of Statistics

Summarizing univariate data

Data collection

Ratios of statistical data

Frequency distributions and graphical displays

Central tendency measures

Variability measures

Concentration measures

Skewness measures

Mathematical models for frequency distributions

Summarizing multivariate data

Main interpolation methods

The least squares method

The least square line and its properties

Bivariate and partial frequency distributions

Independence and association measures

The regression function and the regression line

Concordance and correlation measures

The least square plane

Multiple correlation in the case of more than one independent variable

Detailed program

1. **Introduction to Statistics**
2. **Statistics and its partitions**

· Descriptive Statistics

· Inferential Statistics

3. Foundations of descriptive Statistics

· Statistical data definition

- How to observe and collect survey and population data
- Data collection, examination and selection
- Preparation of data and statistical tables
- Statistical-mathematical data processing.

4. Statistical Ratios

· Their definition, aims and use

· Ratios of statistical data: composition ratio, density ratio, derivation ratio, coexistence ratio, Index Numbers

5. Univariate descriptive Statistics

· Absolute, relative, cumulated, retro-cumulated and specific frequencies

· Graphical tools for representing frequency distributions.

· **The means**: the mode, the median and quantiles (quartiles, deciles, centiles, ...) the arithmetic mean, the harmonic mean, the quadratic mean and the geometric mean. Chisini's invariance principle for the means. Annual average variation rate and mean index number.

· **Variability**: general concept and classification of the indices to measure it.

- Absolute indices of variability: intervals of variation, mean deviation from a mean value, mean difference.
- Relative indices of variability

· **Concentration (inequality):** general concept and fields of applications

- Lorenz diagram and its properties.
- Gini concentration ratio as a ratio among areas and with reference to the mean difference
- Requirements of inequality measures

· **Asymmetry:** general concept, within the study of the shape of a distribution

- The definition of symmetry for a frequency distribution
- Indices for measuring the direction of asymmetry

· **Models for the analytic representation** of frequency distribution for quantitative continuous variables

- general requirements
- The normal curve: its analytical formulation, properties and usage.

- The standard normal curve and the usage of its statistical tables.
- Criteria for data normality
- Some notes on the Log-normal and the Pareto curve.

6. Interpolation

· general definition and usage

· Interpolation of a given set of points

· Interpolation among a given set of points

- Choice of the interpolating curve and of the fitting criterium.
- The Least square fitting method.
- The least square line: parameters determination and properties
- Analysis of the residuals and of the goodness of fit.

7. Bivariate descriptive statistics

· Bivariate frequency distributions

· Dependence in distribution (association).

- Distributive independence and maximum dependence
- The contingencies.
- Measures of association and interpretation of the dependence in distribution

· Mean dependence and the case of mean independence

- Pearson correlation ratio as a measure of the degree of mean dependence

· The regression piece-wise line. The regression line and the least square line for a bivariate frequency distribution, with the assessment of its goodness of fit.

· The positive and negative correlation among variables.

- The covariance and its properties.
- The Bravais-Pearson linear correlation coefficient: definition, properties, usage and interpretation.

8. Regression and correlation for three variables

· Introduction

· Some models.

· The least square method.

· The least square plane.

· The least square parameters determination, using the properties of the arithmetic mean to yield the normal system.

· Properties of the residuals and of the least square plane.

· Total variance, residual and explained variance.

· Goodness of fit of the interpolating plane.

· Multiple correlation coefficient.

· Improvement of the goodness of fit, when passing from the least square line to the least square plane.

· Total and partial regression coefficients.

· Partial correlation coefficients.

Prerequisites

This course requires the use of common concepts of mathematics, from the secondary school.

Teaching methods

Traditional method (8 credits): 45,5 hours of theoretical lectures and 18 hours of practical lectures

Assessment methods

The exam is written and oral. The written test consists of three open questions about theory and three numerical exercises. The theoretical questions tests students' knowledge and understanding of the main concepts of the subject. The exercises measures students' ability in the application of such concepts to solve simple practical problems. Students with a mark greater or equal to 18/30 in the written test are admitted to the oral exam. The oral exam is a discussion on the written test and on subjects from the program. The ability to comment the practical problems and to express the concepts with an appropriate language will also be considered in the global evaluation.

Students can choose to split the written test into two written parts, by taking the first written part at the end of April/the beginning of May and the second one during the regular exam session of June. Both parts contain two open questions about theory and two numerical exercises (as explained before). Each part contributes to 50% of the final evaluation. After being successful in the two written parts, a final interview could be required for completing the exam.

Textbooks and Reading Materials

M. Zenga “Lezioni di statistica descrittiva”, Ed. Giappichelli, 2007

M. Zenga “Esercizi di statistica”, Ed. Giappichelli, 1993

M. Zenga “Richiami di matematica”, Ed. Giappichelli, 1992

M. Zenga “Metodi statistici per l’Economia e l’Impresa”, Ed. Giappichelli, 1994

D. Piccolo, “Statistica per le decisioni”, Ed. Il Mulino, 2004

G. Leti “Statistica descrittiva”, Ed. Il Mulino, 1983

Semester

Second semester

Teaching language

Italian