

Elements of Statistics

Lecturer: Anatoliy A. Peresetskiy

Class teachers: Olga S. Vilenskaya, Vladislav V. Smirnov, Kirill A. Melkumiants

Course description

Elements of statistics is a two-semester course for second year ICEF students. This is a course for students specializing in economics. The course is taught in Russian and English.

Basic ideas of statistics, such as descriptive statistics, population and sample, parameters estimation, statistical hypotheses checking etc, are studied in the course, as well as elements of probability theory which are necessary for understanding the course.

Teaching objectives

The main objective of the course is to give a sound and self-contained (in the sense that the necessary probability theory is included) description of classical or mainstream statistical theory and its applications.

The students should learn to carry out a simple analysis of data (to find mean, median, standard deviation and other descriptive statistics), to present the data graphically (histograms, stemplots). They should understand the differences between population and sample, and theoretical and sample characteristics. Since it is not worth while to teach Statistics without elements of probability theory, studying its basic notions and results is a part of the course. Students should understand what probability space, random event, probability of an event are. They should know how to calculate probabilities of complex events, solve elementary combinatorial problems, use the full-probability and Bayes formulas. The students should have a clear understanding of what a random variable and its distribution are.

The students should learn to formulate and solve basic problems of statistics, such as parameters estimation, statistical hypotheses checking, correlation analysis, analysis of variance. One of the course aims is to prepare students for further studying of Econometrics on the basis of studying simple and multiple regression models.

The course is not mathematically rigorous. Proofs, and even exact statements of results, are often not given. The problems are an essential part of the course. A serious effort has been made in the problems to illustrate the variety of ways in which the theory may be applied.

By the end of the course the students should have gained an understanding of the underlying theory and an ability to carry out relevant calculations and apply standard methods in practice

In the end of the year students should pass an UL exam “Elements of Statistics”.

Teaching methods

The following methods and forms of study are used in the course:

- lectures
- classes
- self study
- weekly written home assignments (to be graded)
- teachers’ consultations

Grade determination

Approximate formula for the final grade in HSE program in statistics is given below:

- Home assignments (total HA sum)/(total number of HA+Quizes = 19+2=21) — 0.10
- Midterm 1 — 10%
- Exam (winter) — 15%
- Midterm 2 — 10%
- Final exam — 30%
- Max(London Stat-1, London Stat-2) — 25%
- Extra additions to score
 - Activity in class (for each point) — 0.8%*
 - Misprints (for each misprint) — 0.8%*

**to be revised*

If a student get less than 25 point at the final exam it means he FAILED, and have to pass a Make-up, doesn’t matter what the final score is.

Main reading

1. P.Newbold. Statistics for Business and Economics. Prentice–Hall, 4th edt., 1995; (N) 5th edition, 2003.

Additional reading

1. С.А. Айвазян, В.С. Мхитарян. Теория вероятностей и прикладная статистика. ЮНИТИ. Москва, 2001.
2. Гмурман В. Е. Теория вероятностей и математическая статистика. М., “Высшая школа”, 1998.
3. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике. М., “Высшая школа”, 1998.
4. Шведов А.С. Теория вероятности и математическая статистика. ВШЭ, 1995.
5. Hogg R.V. and Tanis E.A., Probability and Statistical Inference, Prentice Hall, 1993.
6. Johnston A.R. and Bhattacharyya G.K., Statistics. Principles and Methods. 3rd edition, Wiley, 1996
7. Elements of Statistics, Study Guide. University of London, 2006.
8. Statistics 2, Study Guide. University of London, 2006.

Internet resources and databases

1. http://www.londonexternal.ac.uk/current_students/programme_resource_lse/exam_archive/exam_papers_2006.shtml
2. <http://www2.gsu.edu/%7Edscbms/ibs/qcontent.html>
3. <http://www.math.uah.edu/stat/>
4. <http://www.math.yorku.ca/SCS/Gallery/>
5. http://www.bbn-school.org/us/math/ap_stats/applets/applets.html
6. <http://wise.cgu.edu/links/applets.asp>

Course outline

1. Primary data analysis

- Graphical presentation of one variable data. Dotplot. Stemplot. Histogram.
- Data irregularities. Outliers. Clusters. Histogram shape.
- Descriptive statistics. Measures of central tendency. Arithmetic mean, median, mode, geometric mean. Measures of variation. Range, sample standard deviation, interquartile range, mean absolute deviation, mean relative deviation.
- Descriptive statistics and linear transformation of data.
- Measures of the relative standing. Quartiles. Percentiles. Z-score.
- Grouped data.

(N Ch.1, 2)

2. Elements of Probability theory

- Probability. Outcome space. Events algebra. Independent events. Mutually exclusive events. Conditional probability. Full probability formula. Bayes' formula.
- Discrete Random Variable. Probability density function (p.d.f.). Mathematical Expectation. Variance. Standard deviation. Binomial distribution.
- Continuous random variables. Distribution function. Probability density function. Uniform distribution. Normal distribution.
- Law of large numbers. Normal approximation to binomial distribution. Linear transformation of a random variable.
- Two random variables. Covariation. Correlation. Uncorrelated and independent random variables. Mean and Variance of a linear combination of two random variables.

(N Ch.3, 4, 5)

3. Elements of Mathematical Statistics

- Random Samples. Duality of Interpretation. Estimation of population parameters. Sample mean and sample variance. Mean and variance of the sample mean. Estimation of proportions.
- Point Estimation. Properties of the estimators. Unbiasedness, efficiency, consistency, estimators for mean and variance.
- Interval estimation. Confidence intervals. Estimation of the mean. Normal approximation for large samples, small samples (Student distribution). Difference of two means. Proportions.
- Hypothesis testing. Hypothesis testing with confidence intervals. Hypothesis testing with test-statistics. Two-sided and one-sided p -values.
- χ^2 goodness-of-fit test. Contingency tables.

(N Ch.6–9, 11)

4. Models of simple regression

- $X - Y$ plot. Line fitting. Ordinary Least Squares.
- Transformations of regressors.
- Outliers.
- Forecasts.
- Regression residuals. Residuals and errors.
- Statistical properties of the estimators. Hypothesis testing.

(N Ch.12)

5. Analysis of variance (ANOVA)

- One-factor analysis of variance.
- Two-factors analysis of variance.
- Confidence intervals.

(N Ch.15)

Distribution of hours

#	Topic	Total hours	Contact hours		Self study
			Lectures	Seminars	
1.	Primary data analysis	26	6	6	14
2.	Elements of Probability theory	80	20	20	40
3.	Elements of Mathematical Statistics	110	30	30	50
4.	Models of simple regression	28	6	6	16
5.	Analysis of variance (ANOVA)	26	6	6	14
Total:		270	68	68	134