Syllabus
«Bayesian data analysis»

Lecturer: *Irina Kalinina*
PhD, Associate Professor, Associate Professor of the Intelligent Information Systems Department

Expected learning outcomes
As a result of studying the discipline, the student
must know:

- mathematical formalization and peculiarities of application of Bayesian methodology of data analysis;
- the main types of probabilistic statistical models used in Bayesian data analysis;
- methods and procedures for taking into account uncertainties in probabilistic and statistical Bayesian modelling;
- mathematical formalization and features of application of Bayesian theorem and network for discrete and continuous data;
- features of probabilistic methods and approaches to in-depth data analysis, performance of theoretical and applied scientific research, improvement of knowledge and skills regarding the practical use of obtained scientific results to support the adoption of appropriate decisions of a systemic nature in accordance with problem statements;
- features of comparative analysis of results obtained by other researchers, creation of new concepts, methods, theoretical bases and algorithms of probabilistic-statistical modelling, forecasting and decision-making support in the economy, finance, business and production.

must be able to:

- independently carry out research activities using probabilistic and statistical Bayesian methods of data processing, including in-depth

The scope of the discipline: 5 ECTS credits (16 hours of lectures, 48 hours of practice).

Purpose: development of general and professional competences related to data collection and analysis using Bayesian probabilistic statistical methods, mastering analytical probabilistic methods and approaches to in-depth data analysis, performing theoretical and applied scientific research, improving knowledge and skills related to the practical use of scientific results to support appropriate decision-making of a systemic nature according to the statements of tasks; comparative analysis of results obtained by other researchers, creation of new concepts, methods, theoretical bases and algorithms of probabilistic-statistical modelling, forecasting and decision-making support in the economy, finance, business and production.

The content of the discipline

Topic 1. Introduction to the Bayesian methodology of data analysis. Bayesian programming methodology. Structural components of the Bayesian methodology of data analysis. Peculiarities of the Bayesian methodology of data analysis and expert evaluations. Comparison of classical probabilistic data analysis and Bayesian approach.
Topic 2. Basic types of probabilistic statistical models used in Bayesian data analysis. Peculiarities of using Bayesian models of various types. Definition of probabilistic statistical model. The structure of the Bayesian model and its evaluation. The sequence of building a probabilistic-statistical model of the Bayesian type.
Topic 3. Uncertainties in probabilistic statistical modelling. Connection with the principles of system analysis. Types of uncertainties encountered in probabilistic statistical modelling. Identification of uncertainties related to statistical data and expert judgments. Analysis of uncertainties in the methodology of system analysis.
Topic 4. Methods and procedures for taking into account uncertainties in probabilistic and statistical Bayesian modelling. Consideration of probabilistic uncertainties. Statistical uncertainties and their consideration in modelling procedures. Interval assessment and its

analysis of tasks, problems, setting goals and objectives, choosing means and methods of research, as well as analysing assessments of its quality;
- initiate, plan, implement and adjust a consistent process of scientific research using probabilistic-statistical Bayesian modelling, forecasting and decision support.

Prerequisites

The teaching of the academic discipline is based on the knowledge obtained as a result of the study of previous academic disciplines and the acquisition of competences after completing studies at the bachelor's level, it requires basic knowledge of mathematical and informational disciplines, sufficient for the perception of methods and models based on the use of a probabilistic-statistical Bayesian approach to modelling, forecasting and decision support. In particular, these are "Higher mathematics", "System analysis", "Intelligent data analysis", "Probability theory and mathematical statistics", "Decision-making theory", "Methods and systems of machine learning".

Consequences

The discipline provides applicants with the third degree of higher education with the necessary knowledge and practical skills for the accumulation and analysis of scientific information regarding probabilistic-statistical methods of statistical/experimental data analysis on the topic of qualification work, construction and analysis of the adequacy of created probabilistic-statistical models intended for solving forecasting problems short- and medium-term forecasting of the development of processes and events and supporting the adoption of relevant management decisions.

Technical support

Practices on the discipline is carried out in computer classes using languages and software environments: R Studio, R, Python.

Deadline policy

Works that are submitted in violation of deadlines without good reason are evaluated at a lower grade.

use.

Topic 5. Function and principle of plausibility, definition and application. Calculation examples. Peculiarities of probabilistic statistical modelling using probability functions. Using the principle of plausibility in methods of system analysis. Calculation of optimal estimates of parameters of mathematical models.

Topic 6. Bayes theorem for discrete data and discrete parameters. Bayes theorem for continuous data and discrete parameters. Formulation, interpretation and application possibilities. Analysis of the constituent elements of the theorem. Analysis of application examples.

Topic 7. Bayesian networks for discrete and continuous data and expert evaluations. Formulation, construction, interpretation and application possibilities. The sequence of building a probabilistic statistical model in the form of a Bayesian network. Discretization of continuous data. Formation of a probabilistic conclusion based on the network and its interpretation. Application examples.

Topic 8. Representation of models in the space of states. Advantages and disadvantages. State space methods: Kalman filters for a linear Gaussian model; hidden Markov models; Bayesian structural time series. Examples of representing models in state space.

Topic 9. Bayesian structural time series (BSTS). Bayesian principles of time series analysis. BSTS model learning algorithm with and without predictors.

Evaluation criteria of laboratory works / practices / individual works / reports / projects

Maximum number of points a student with high quality independently completed the entire amount of work, answers all questions related to the completed work, and makes additional calculations offered by the teacher. The teacher has no complaints about the implementation and requirements for the performance of the work.

Approximately 70%-99% of the maximum number of points – a student completed all tasks with sufficient quality, but in the process of work he made some mistakes, which, after being pointed out by the teacher, he corrected himself. He answers some questions incorrectly. Additional calculations proposed by the teacher are done with some effort. Not all requirements for performance of work are met.

Approximately 40%-69% of the maximum number of points – a student independently completed all the work, but the quality of the implementation is insufficient (calculation errors, not all work requirements are met). The answers to questions about the performance

Academic Integrity Policy

Provides independent performance of practices. Write-off (including using mobile devices) is prohibited. The work is not credited in case of detection of plagiarism or write-off.

of work are not quite clear. There are mistakes in the answers.

Approximately 1%-39% of the maximum number of points – a student did not complete all the work independently, while the quality of the implementation was insufficient (errors in calculations, does not comply with the requirements for the design of the work). He does not answer questions about the performance of work clearly. There are gross mistakes in the answers.

0 points – a student did not complete the entire amount of work, or did it with gross errors. He has problems with calculations, does not know the theoretical material, the software implementation does not meet the requirements.