Правительство Российской Федерации

Федеральное государственное автономное образовательное

учреждение высшего образования
«Национальный исследовательский университет
«Высшая школа экономики»

Факультет Компьютерных наук

Департамент математики на факультете экономических наук

УТВЕРЖДАЮ

Академический руководитель

образовательной программы

по направлению 01.03.02

«Прикладная математика и информатика»

А.С. Конушин

____

«___» _____________ 2017 г.

Программа дисциплины

Временные ряды и случайные процессы
(«Time Series and Stochastic Processes»)

Для направления 01.03.02 «Прикладная математика и информатика» подготовки бакалавров

Автор программы:

к. ф.-м. н. Горяинова Е.Р., доцент Департамента математики на ф-те экономических наук (el-goryainova@mail.ru)

Одобрена на заседании

Департамента математики на ф-те экономических наук «03» _03 _ 2017 г.

Зав. кафедрой Ф.Т. Алескеров

Рекомендована Академическим советом

образовательной программы

«Прикладная математика и информатика» «___»__________ 2017 г.

Менеджер департамента О.А. Колотвина

A Syllabus

1. Instructor and author

Elena R. Goryainova Associate Professor of the Department of Mathematics Faculty of Economic Sciences NRU HSE Moscow

2. Reference to regulatory document

This syllabus is prepared according to the Teaching standard «Образовательный стандарт федерального государственного автономного образовательного учреждения высшего образования национального исследовательского университета **«Высшая школа экономики»** по направлению подготовки 01.03.02 «Прикладная математика и информатика»

3. Abstract

This course presents an introduction to time series analysis and stochastic processes and their applications in operations research and management science.

Time series includes the description of the following models: white noise, Moving average models MA(q), Autoregressive models AR(p), Autoregressive-moving average ARMA(p,q) models, Nonlinear Autoregressive Conditional Heteroskedasticity (ARCH(p)) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH(p;q)) models and VAR models. Also, the solution of the problem of identification of the ARMA process, including the model selection, estimation of the model parameters and verification of the adequacy of the selected model, is given. Methods for reducing some non-stationary time series to stationary ones by removing trend and seasonal components are described. Then, the Dolado-Jenkinson-Sosvilla-Rivero procedure is presented to distinguish non-stationary time series such as Trend-stationarity (TSP) and Difference-stationarity (DSP). The procedure for diagnosing the presence of spurious regression is also considered.

Stochastic processes are discussed on a basic process Brownian motion and Poisson process. The method for constructing optimal forecasts for Gaussian stochastic processes and stationary time series is given.

At the end of the course Markov chains and continuous-time Markov chains are considered. For these models, the conditions for the existence of a stationary distribution are established. In particular, are found the final distribution for the processes of «birth and death» and for the queueing system M/M/n/r.

4. Learning Objectives.

To familiarize students with the concepts, models and statements of the theory of time series analysis and stochastic processes.

5. Learning Outcomes.

- Know basics of time series analysis and stochastic processes;
- Be able to choose adequate models in practical socio-economic problems;
- Have skills in model construction and solving problems of time series analysis and stochastic processes.

6. Student's competences after the course.

After completion of the course, the student must have the following universal competencies: the ability to identify the scientific essence of problems in the professional field and the ability to solve problems in professional activity on the basis of analysis and synthesis (These relate to items УК-2, УК-3 of the Teaching standard). The subject-oriented competencies include: the ability to describe problems and situations of professional activity using the language and mathematics software , the ability to understand, improve and apply the modern mathematical apparatus , the ability to develop a mathematical model and to analyze it for the assigned theoretical or applied task, the ability to conduct written and oral communication in English within the framework of professional and scientific communication (These relate to items ПК-1, ПК-3, ПК-8, ПК-11 of the Teaching standard).

7. Pre-requisites

-basic linear algebra;
-basic courses in Calculus;
-course Theory of Probability and Mathematical Statistics;
- Spoken English (intermediate level).

8. Distribution of Hours

Unit	Topic title	Lectures	Classes	Self-study	Total
1	Basic concepts of the theory of stochastic processes	2	2	10	14
2	Some types of stochastic processes	2	2	6	10
3	Main models of stationary time series	2	2	10	14
4	Forecasting	2	2	12	16
5	Identification, estimation and testing of ARMA(p,q) models	4	5	24	33
6	Identification of nonstationary stochastic processes	4	4	12	20
7	Vector autoregressive models.Causality	2	2	10	24
8	Markov chains	2	2	10	24
9	Continuous-Time Markov Chains	2	3	12	17
Total		22	24	106	152

9. Course Plan.

Topic 1. Basic concepts of the theory of stochastic processes

Definitions of a stochastic process (SP), Time series, realizations (or sample-paths) of the process, finite-dimensional distribution functions of stochastic process. Kolmogorov consistency theorem (without prove). Main characteristics of time series (expectations, variance, moments, covariance function, correlation function. Properties of a covariance function. Examples.

Readings:

Main [1], [2],[5]. Additional [1],[2]

Topic 2. Some types of stochastic processes.

Strictly stationary stochastic process, weakly stationary stochastic process, relationship between weak and strict stationarity. Stochastic process with independent increments. stochastic process with orthogonal increments. Poisson stochastic process. Gaussian stochastic process, finite-dimensional density function of a Gaussian stochastic process. The Wiener process (Brownian Motion). Relationship between random walk and Brownian motion. Filtration problem.

Readings:

Main [4],[5]. Additional [6].

Topic 3. Main models of stationary time series.

Linear stochastic process. Lag (or back shift) operator. Discrete white noise. Moving average models MA(q). Condition of invertibility MA(q). Autoregressive models AR(p). Condition of stationarity AR(p). Autoregressive-moving average ARMA(p; q) models. Conditions of stationarity ARMA(p; q). Time-varying volatility. The notion of conditional volatility. Nonlinear Autoregressive Conditional Heteroskedasticity (ARCH(p)) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH(p;q)) models.

Readings:

Main [1], [2], [3]. Additional [1], [2], [3], [8], [10].

Topic 4. Forecasting.

An optimal in the mean square sense predictor. A mean square error of the predictor. The theorem on the best (in the mean square sense) predictor (with prove). Forecasting of Gaussian processes. Theorem on Normal Correlation. Forecasting of stationary time series.

Readings:

Main 4 Additional [1], [8], [10].

Topic 5. Identification, estimation and testing of ARMA(p,q) models

Sample autocorrelation function (ACF), sample partial autocorrelation function (PACF), correlograms. Statistical properties of sample ACF and sample PACF. Goodness of fit in time series models. Yule-Walker's method for estimating the parameters of AR(p) models. Backcasting procedure for estimating the parameters of MA(q) models. Recursive least squares (LS) method for estimating the parameters of ARMA(p,q) models. Distribution of LS estimates in ARMA(p,q) models. Check residuals for white noise. Akaike information criterion (AIC). Schwarz information criterion (SIC). Ljung-Box and Box-Pierce Q-tests. Jarque-Bera test for checking the normality of residuals.

Readings:

Main [1], [2], [3]. Additional [1], [2], [3], [6], [10].

Topic 6. Identification of nonstationary stochastic processes

Models with Trend and Seasonality. Box-Jenkins methodology. Difference operator. ARIMA models. Trend-stationarity stochastic process (TSP), Difference-stationarity stochastic process (DSP). Spurious regressions. Problem of the unit root. Dickey-Fuller test. Augmented Dickey-Fuller tests. Dolado-Jenkinson-Sosvilla-Rivero procedure.

Readings:

Main [1], [2]. Additional [1], [4], [5], [9].

Topic 7. Vector autoregressive models.Causality.

Vector autoregressive models. ADL models. Cointegrated series. The notion of causality. Granger causality.

Readings:

Main [1], [3]. Additional [2], [3], [4], [5].

Topic 8. Markov chains.

Markov processes as generalizations of IID variables and of deterministic dynamical systems. The Markov property and the strong Markov property. Classifications of States of Markov chain. Ergodic Markov chain. Limiting distribution of Markov chain. The Classical Ruin Problem.

Readings:

Main [4]. Additional [11].

Topic 9. Continuous-Time Markov Chains.

A series of events. Chapman-Kolmogorov Equations. Ergodic properties of homogeneous Markov chains. Birth and Death Processes. Queuing theory.

Readings:

Main [2], [4]. Additional [11].

Mid-term exam (approximate version)

1. A stochastic sequence $\xi_n,\ n \in \mathbb{Z}$, satisfies a stationary second-order autoregression equation with known parameters. There are observations $\xi_{50}, \xi_{51}, \xi_{52}$. Construct the best (in the mean square sense) forecast for the ξ_{53}.

2. Put by definition $\xi(t) = \begin{cases} 1, & t < \tau \\ -1, & t \geq \tau \end{cases}, t \geq 0, \tau \sim E(\alpha)$.

Find one-dimensional and two-dimensional distributions of the process $\xi(t)$.

3. Assume that X and Y are independent random variables, X has a normal distribution with mean 0 and variance 1, Y has an uniform distribution on $[-\pi, \pi]$. Find the covariance functions of the stochastic process $\xi(t) = cos(t+Y)$, $-\infty < t < \infty$. Is the stochastic process $\xi(t)$ a stationary process?

4. Let $W(t)$ be a standard Wiener process. You have two observations of this process at points t_1 and t_2. Give the best (in the mean square sense) estimate of $W(t)$ at a point $t = t_3$, $0 < t_1 < t_2 < t_3$. Find the mean square error $\Delta = E(W(t_3) - \widehat{W}(t_3))^2$, of this estimate.

Final exam (approximate version)

1. The values of the sample autocorrelation function ACF (k) and the values of the sample partial autocorrelation function PACF (k), k=1,...,12. of the observed time series X(t), t=1,...,225 are given in the table:

k	1	2	3	4	5	6	7	8	9	10	11	12
ACF	-0,757	0,502	-0,31	0,222	-0,183	0.156	-0,135	0,124	-0,117	0,014	-0,110	0,096
PACF	-0,757	-0,166	0,019	0,089	-0,027	0,005	0,071	-0,006	0,013	0,004	-0,007	0,004

Make identification of this time series. What further steps will you take to build the statistical model of the X(t) series?

2. There are 100 observations of the time series X_t. The estimates and the mean-square deviations of the estimates of the parameters of the model $X_t = \mu + \beta t + \alpha X_{t-1} + +\theta \Delta X_{t-1} + \epsilon_t$ were computed by the least squares method: The results of the estimation are given in the table:

μ	β	α	θ
53,02 (13,18)	1,2 (0,24)	0,751 (0,061)	0,41(0,1)

Lags of higher order are not included in the model, because the coefficients at lags of a higher order are considered insignificant. Find out if this series is a random walk.

3. There are 500 observations of the time series X_t By the method of least squares, estimates and mean-square deviations of the estimates of the parameters of the following models were calculated: $X_t = \mu + \beta t + \alpha X_{t-1} + \epsilon_t;\ \ \Delta X_t = \ \mu + \beta t + \epsilon_t;\ \ X_t = \mu + \alpha X_{t-1} + \epsilon_t,\ \ \Delta X_t = \ \mu +$

ϵ_t and $X_t = \alpha X_{t-1} + \epsilon_t$. According to White's test, Breusch-Godfrey test and Jarque-Bera test, the residuals of each model are recognized as homoscedastic, uncorrelated and Gaussian, respectively. The results of the evaluation are given in the table:

Model	μ	β	α
1	0,124(0,05)	0,00025(0,00011	0,97(0,013)
2	0,029(0,013)	0,0007(0,0006)	
3	0,09(0,05)		0,94(0,026)
4	0,08(0,06)		
5			0,95 (0,019)

Can you assume that the series X_t is a DSP series? The answer is justified.

4. A matrix of transition probabilities for a homogeneous Markov chain is given

$$\begin{pmatrix} 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}.$$

Is the chain ergodic? Find all stationary distributions of the chain.

5. The queuing system consists of two identical service channels with service rates μ. The stream of applications is the simplest Poisson flow with intensity λ. It is required: 1) to construct a stochastic graph of the service process; 2) write the Kolmogorov equations for calculating the probabilities of the states of the process; 3) find the stationary distribution of the process; 4) calculate the probability of idle time and the probability of full occupancy of channels in the case when λ = 2μ.

Exam paper questions

Topic 1.

1. Give the definition of a stochastic process.
2. Give the definition of a time series.
3. Give the definition of a realizations of the stochastic process.
4. Give the definition of finite-dimensional distribution functions of stochastic process.
5. What are the main characteristics of stochastic processes you know?
6. Formulate and prove the basic properties of a covariance function.

Topic 2.

1. Give the definition of a strictly stationary stochastic process.
2. Give the definition of a weakly stationary stochastic process.
3. What is the relationship between weak and strict stationarity?
4. Give the definition of a Gaussian stochastic process.
5. Give the definition of the Wiener process (Brownian Motion).
6. What is the relationship between random walk and Brownian motion?
7. Write n-dimensional density function of a Gaussian stochastic process.

Topic 3.

1. Give the definition of a discrete white noise
2. What time series are called linear stochastic processes?
3. What basic models of stationary time series do you know?
4. What are the conditions of the invertibility of a moving average process?
5. Find autocorrelation function of a moving average MA(q) process.
6. Find autocorrelation function of AR(1) process
7. What are the causes of nonlinear models?
8. Describe the nonlinear stationary models Autoregressive Conditional Heteroskedasticity (ARCH(p)) and Generalised Autoregressive Conditional Heteroskedasticity (GARCH(p;q))
9. What is called volatility of time series?

Topic 4.

1. Give the definition of an optimal in the mean square sense predictor.
2. What is called a mean square error of the predictor?
3. Specify the explicit form of the best(in the mean square sense) predictor.
4. Describe the procedure for constructing the best forecast for a Gaussian process.

Topic 5.

1. What are the statistical properties of the sample autocorrelation function (ACF)?
2. What are the statistical properties of the sample partial autocorrelation function (PACF)?
3. What is the asymptotic distribution of the sample autocorrelation function (ACF) for white noise?
4. What is the asymptotic distribution of the sample autocorrelation function (ACF) of a moving average MA(q) process?
5. What is the asymptotic distribution of the sample partial autocorrelation function (PACF) of an autoregressive AR(p) process?
6. Describe methods for estimating the parameters of the autoregression equation.
7. Describe the procedure for determining the order of the model of the autoregression process.
8. What criteria do you know about the adequacy of the selected ARMA model?
9. Describe the procedure for checking the Gaussianity of the residuals of the ARMA model.

Topic 6.

1. Describe Box-Jenkins methodology.
2. Describe the main stages of the information of the non-stationary time series to the stationary one.
3. What methods of estimating and excluding trend and seasonal components are known to you? Show on examples.
4. What are the main differences between non-stationary series such as TSP and DSP.
5. Describe Dickey-Fuller test and Augmented Dickey-Fuller tests.
6. Describe the Dolado-Jenkinson-Sosvilla-Rivero procedure. What is this procedure for?

Topic 7.

1. Specify vector autoregressive models
2. How to identify spurious regression?
3. What time series are called cointegrated?
4. How to identify the cointegration of two series?
5. What is Granger’s causality?

Topic 8.

1. What is the Markov chain?
2. Which chain is called homogeneous?
3. What states are called essential, communicating, aperiodic?
4. State the ergodic theorem.
5. How can one find the stationary distribution of a homogeneous Markov chain if it exists?

Topic 9.

1. Give the definition of the Poisson stochastic process
2. Which stream of events is called the simple Poisson stream?
3. Define the Markov property for a CS with continuous time and a finite set of states.
4. Write down the Kolmogorov algebraic equation system for a homogeneous Markov process.
5. Determine the intensity of the transition from state to state.
6. What kind of stochastic graph has the birth-death process?

Main reading

1. Brockwell, P.J., and R.A. Davis, 2003, Introduction to Time Series and Forecasting, Springer Publ., 2nd ed.
2. Kirchgässner G., Wolters J. Introduction to modern time series analysis. Springer, Berlin, 2007.
3. Enders, W., 2003, Applied Econometric Time Series, Wiley Publ., 2nd ed.
4. Shiryaev A. N. Essentials of stochastic finance. Facts, models, theory: Advanced Series on Statistical Science and Applied Probability, 3. World Scientific Publishing Co., Inc., River Edge, NJ, 1999.
5. Kannan D. An introduction to stochastic processes. North Holland Series in Probability and Applied Mathematics. North-Holland, New York-Oxford, 1979.

Additional reading

1. Box, G.E.P., G.M. Jenkins, and G.C. Reinsel, 2008, Time Series Analysis: Forecasting and Control (Wiley Series in Probability and Statistics), Wiley Publ., 4th ed. – 784 pages. ISBN: 978-0-470-27284-8.
2. Lütkepohl, H., &Krätzig, M., 2004, Applied time series econometrics. Cambridge University Press.
3. Patterson, K.D., 2000, Introduction to Applied Econometrics: A Time Series Approach, Palgrave Macmillan.
4. Maddala, G.S. and In-Moo Kim, 1999, Unit Roots, Cointegration, and Structural Changes (Themes in Modern Econometrics), Cambridge University Press.
5. Banerjee, A., J. Dolado, J.W. Galbraith, and D.F. Hendry, 1993, Co-integration, error-correction, and the econometric analysis of non -stationary data, N.Y., OxfordUniv. Press
6. Harvey, A.C., 1993, Time Series Models, Harvester Wheatsheaf, 2nd ed.
7. Maronna R.A., Martin D., Yohai V. Robust statistics. Theory and methods. Wiley Series in Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 2006.
8. Williams R.J. Introduction to the Mathematics of Finance. – Providence, Rhode Island, American Mathematical Society, 2006.
9. Dolado J.J., Jenkinson T. and Sosvilla-Rivero S. Cointegration and Unit Roots // Journal of Economic Survey. 1990. Vol. 4. P. 249-73.
10. Hamilton J.D., 1994, Time Series Analysis, Princeton University Press
11. Feller W. An introduction to probability theory and its applications, Vol.1 3ed., Wiley, 1968.

10. Forms of knowledge assessment and grading procedures

Knowledge assessment is conducted as a four-step controlling procedure: 1) Participation in the statistical game; 2) Homework ; 3) Mid-term exam; 4) Final exam.

When conducting a statistical game, the teacher presents to the student the graphics of the sample ACF and sample PACF of a certain ARMA process. The student's task is to correctly indicate the mathematical model of the time series.

Homework is to make identification and forecasting of a real or simulated time series of the TSP type. Identification includes: estimation and removal of the trend, selection of the model for the series residuals, estimation of the parameters of the selected model and checking the adequacy of the selected model. Using the Dolado- Jenkinson- Sosvilla-Rivero procedure, establish the belonging of this series to the TSP type.

The mid-term exam lasts 80 minutes. The mid-term exam includes 4 problems. Typical tasks are given above.

The final exam lasts 100 minutes. The final exam includes 5 problems. Typical tasks are given above.

The **final mark** is computed according to formula **FM=0.1*SG+0.2*HW+0.3*MEX +0.4*EX**, where SG is Participation in the statistical game, HW is Homework, MEX is mid-term exam, EX is final exam.

11.Methods of Instruction

The discipline is delivered through lectures seminars, including computer classes.

12. Special Equipment and Software Support (if required): Computer classes