Approved
Academic Council of educational program " Economics: Research Program"
Protocol number 01, « 25» May 2017
Academic Council of educational program "Statistical Modelling and Actuarial Science"
Protocol number 01, « 22» May 2017

Author	A.A. Peresetsky, D.Sc., aperesetsky@hse.ru; J-F.M Jabir, PhD, jjabir@hse.ru
Credits	5
Contact Hours	96
Self-study Hours	94
Course	1

1 General information about the course.

This course aims to provide a solid introduction to probability theory and mathematical statistics. The fundamentals concepts and mathematical tools for the modeling and analysis of random phenomenon. Elementary models (in discrete and continuous settings) will be presented and discussed, meanwhile the last part of the course will be dedicated to a short introduction to discrete time stochastic processes and their applications in mathematical finance.

2 Course goals, learning objectives, expected learning outcomes

The course is first course in the sequence of probabilistic-econometrics courses
The main objectives of the course are:

- to provide students with the knowledge of the theoretical and modeling aspects related to probability theory;
- to provide students with the knowledge of elementary techniques to analyze probabilistic models;
- to present and study some fundamental distributions, random variable models,
- to develop students' ability to apply the knowledge acquired during the course to study and use probabilistic models in concrete situations, recognizing the appropriate frameworks and analytical tools related to the study.
- • to outline the basic concepts and methods of mathematical and applied statistics;
- • give practical skills in applying statistical methods in applied research;
- • give an idea of the applied methods of multidimensional statistical analysis.

It is assumed that students own methods of mathematical analysis, linear algebra in the volume of standard university courses.

3 Course Outline

Topic 1. Discrete probability. Sample space and operation on sets. Probability on a finite sample space. sigma algebra. Axioms of probability theory. Independent events. Conditional probability. Baye's theorem. Limit of events.

Topic 2. Discrete random variables. General definition of a random variable. Probability distribution of a discrete random variable. Cumulative Distribution Functions. Mean, variance and moments. Probability generating function. Independent random variables. Classical distributions: Bernoulli trials; Binomial distribution; Geometric distribution; Poisson distribution, ... Conditional expectation for discrete random variables.

Topic 3. Continuous random variables. Elements of measure theory. Cumulative distribution function and probability density function. Distribution of a function of a real random variable. Moment, variance and Characteristic function. Classical distributions: Uniform; Normal; Exponential, Cauchy, ...; Markov, Chebyshev and Jensen Inequalities. Conditional density function and Conditional expectation.

Topic 4. Random vectors. Multivariate distributions. Joint laws, Marginal distribution. Covariance and correlation. Gaussian vectors and linear transformation. Chi-deux, Student and Fisher distributions, …

Topic 5. Convergence in probability. Limit of a sequence of random variables: convergence almost sure, in probability, in distribution. Convergence in Lp. The Slutsky theorem. Law of Large Numbers. The Central Limit Theorem and its application.

Topic 6. Discrete time Stochastic Processes. Random walks. Martingale. Markov chains. Modeling in Mathematical Finance: Valuation and Hedging of an European option.

Topic 7. Statistical estimation of parameters. Samples. Property of estimators. Unbiasedness, efficiency, consitency.

Topic 8. Interval estimation. Standard confidence intervals for the parameters of a normal population. Confidence intervals for the mean, variance, difference of means, variances ratio, population proportion, the difference of proportions. The sample size.

Topic 9. Hypotheses testing. Type I, type II errors. P-value of the test. Tests on the values of the parameters of the normal population. Tests on mean, variance, differences, difference of means, variances ratio, population

Topic 10. Методы оценивания. Метод моментов. Метод максимального правдоподобия. Их свойства, примеры. Неравенство информации.

Estimation methods. The Method of Moments. Maximum Likelihood method. Their properties, examples. Information inequality (Fisher).

Topic 11. Test statistics, Neumann-Pearson Lemma. Likelihood ratio test. Wald test. Lagrange multiplier test.

Topic 12. Goodness-of-fit tests. Contingency tables. Kolmogorov-Smirnov test.

Topic 13. Bayes approach to estimation.

Topic 14. One- and Two-way ANOVA.

Topic 15. Some concepts of non-parametric methods. Wilcoxon tests, run test. Rank correlation coefficients.

Topic 16. Classification methods. Discriminant analysis. Separation of the mixture of distributions. Cluster-analysis. Reduction of dimension. Principal component analysis.

Topic 17. Sufficient statistic. Minimal sufficient statistic. Rao-Blackwell Theorem. Complete statistics.Lehmann-Sheffe Theorems.

4 Description of course methodology and forms of assessment to be used

The course consists of two equal parts:
module 1 — Probability Theory
module 2 — Mathematical Statistics

At the end of module 1 — test # 1.
At the end of module 2 — test # 2.
During the course the students carry out home work and short (10-20 min) quizzes at the lectures. Quizzes and home assignments are graded.
Activity of students in seminars is also graded.

4.1 Criteria for the evaluation of knowledge, skills

Tests, home assignments, are evaluated on a 100-point scale. With certain weights, these grades, as well as scores for activity at the seminars, are aggregated into a final grade (also on a 100-point scale), which is then converted into a 10-point scale. The conversion thresholds are not fixed in advance and are determined by the distribution of estimates in a 100-point scale.

All elements of intermediate control are focused at understanding the mathematical foundations of the studied methods, their range of application, understanding of the numerical results of model estimation and practical work with real data.

4.2 Evaluation

Accumulated score is exhibited separately at first (probability theory) and second (mathematical statistics) parts of the course. Accumulated score for the whole course is calculated as the arithmetic mean of the two scores.
Accumulated score on the first part of the course consists of the following components:
60% — score for the final test # 1;
20% — intermediate test;
20% — 20%-work at seminars;

Accumulated score for the second part of the course consists of the following components (each is scored on a 100-point system):
40% — grades for homework and quizzes
60% — score for the test # 2,
To total amount points are added for the activity at the seminars (i.e. theoretically the final score can be more than 100 points).

According to the 10-point grading scale officially adopted at NRU HSE, the cumulative grade obtained by a student on a 100-point system is converted into a score on a 10-point system. Thresholds are not fixed in advance. Below is an example of thresholds for conversion.
If a student in test # 2 scored less than 25 points, then the cumulative score from the results of current control is equal to this number of points.
Retake of components of current control is not allowed, with the exception of test # 2.

One retake of the second test is allowed, for those who missed it with the medical certificate and for those who received less than 25 points for this test score.

Retake is organized during teacher consultations. The time and place of the retake will be agreed between the students and the teacher.

When retake, the setting of the final grade takes place using a similar procedure, but the thresholds for converting the score from one hundred point to ten points scale do not necessarily coincide with the original thresholds, since the complexity of the tasks of the test during the retake may vary.

Thresholds example:

For grading on a 10-point scale, it is possible to use the following basic scale of converting the grades from 100-point scale in grades on a ten-point scale:
0-11% – 1; 12-22% – 2; 23-33% – 3; 34-40% – 4; 41-47% – 5; 48-55% – 6; 56-64% – 7; 65-72% – 8; 73-80% – 9; 81-100% – 10.

In this assessment, "unsatisfactory" correspond to points 1, 2, 3; "satisfactory" correspond to points 4, 5; “good” correspond to points 6, 7; and "excellent" correspond to points 8, 9, 10.

Depending on the complexity of the course and the complexity of the tests, the threshold gradations can vary by a teacher in the range of 15 percentage points.

5 Texts, readings and other informational resources

Main

F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester (2005): *A Modern Introduction to Probability and Statistics*.

G. Grimmett and D. Stirzaker (2001): *One Thousand Exercises in Probability*.

M. Kelbert and Y. Suhov (2005): *Probability and Statistics by Example: Volume 1, Basic Probability and Statistics.*

Hogg R.V. and Tanis E.A. (2009). *Probability and statistical inference*, 8th edition. Prentice Hall.

Complementary

S. Shreve (2004): *Stochastic Calculus for Finance I: The Binomial Asset Pricing Model*.

R. Durrett (2012): *Essentials of Stochastic Processes*.

D. Lamberton and D. Lapyere (2008): *Introduction to Stochastic Calculus applied to Finance*.

Johnston A.R. and Bhattacharyya G.K. (2009). *Statistics. Principles and methods*. 6rd edition, Wiley.

Hogg R.V., McKean J.W. and Craig Allen T. (2012). *Introduction to mathematical statistics*, 7th edition, Pearson Prentice Hall.

Casella G., and Berger R.L. (2012). *Statistical inference*. 2nd edition. Duxbury.

Wackerly D.D., Mendenhall W. III, Scheaffer R.L. (2008). *Mathematical statistics with applications*. 7th edition, Thompson.

6 Technology

Mandatory seminars on mathematical statistics in computer classes with examples on real data and solving theoretical problems. EViews, STATA packages (optional-R)

7 Samples of some basic questions treated in seminars:

- Give an example of an event which is independent of himself.
- Assume that P(A)=0,6 and P(B)=0,45. Show that 0,05 $\leq P(A \quad B) \leq 0.25$.
- Propose a probabilistic model for a sequence of n independent Bernoulli trials.
- Determine the asymptotic distribution of a sequence of binomial distribution B(n,1/n).
- Determine the mean and variance of the Bernoulli, geometric and Poisson distribution by using their moment generating function.
- Given X a real random variable with continuously differentiable and (strictly) increasing cumulative distribution function F, determine the distribution of the r.v. F^{-1}(X).
- Compute the kth moment of a centered gaussian random variable.
- Given a sequence of r.v. Compute the
- Given a centered real r.v. X with standard deviation 0,01, give an estimate of **P(2≤|X|).**
- Given a gaussian vector X~N(0,M), compute the distribution of Y=AX+B.
- Give an example of a sequence of random variables which converges in probability but does not converge almost surely.

7.1 Sample test # 1.

Model examination paper

Solve at least 4 problems from 6 for the full mark 100

Problem 1 (a) Two players with initial capitals a and b toss a fair coin ($p = q = \frac{1}{2}$). One persons bets for a head, another bets for a tail. Let τ be a random moment when one of the players is out of money. Find $\mathrm{E}\tau$.
(b) Find $\mathrm{E}\tau$ if $p > q$.

Problem 2 Let X_n be a Galton-Watson branching process with the offspring distribution $p_0 = \frac{1}{3}, p_2 = \frac{2}{3}$. (a) Find the survival probability

$$\pi = \lim_{n\to\infty} \mathbf{P}(X_n > 0).$$

(b) Let $X_0 = 1$. Find the distribution of X_2, i.e. the number of species in generation 2.

Problem 3 Let $U_1, U_2, U_3 \sim$U$[0,1]$ be IID RVs. (a) Find distribution of

$$|U_1 - U_2|.$$

. (b) Find the distribution of $U_1 + U_2 + U_3$.

Problem 4 Let $X_1, \ldots, X_n \sim \mathrm{N}(0,1)$ be IID RVs. Let

$$S = (X_1 - X_2)^2 + (X_2 - X_3)^2 + \ldots + (X_{n-1} - X_n)^2.$$

Find $\mathrm{E}S$, $\mathrm{Var}(S)$.

Problem 5 Let $X_1, \ldots, X_n \sim$Exp(1) be IID RVs and $X_{(1)} < \ldots < X_{(n)}$ be the order statistics. (a) Find the joint distribution of $X_{(n-1)}, X_{(n)}$. Find the joint distribution of $X_{(1)}, X_{(n)}$.
(b) Find the distribution of the span $X_{(n)} - X_{(1)}$.

Problem 6 (a) Let $U_1, \ldots, U_n \sim$U$[0,1]$ be IID RVs. Find the distribution of $-2\ln(U_1 U_2 \ldots U_n)$.
(b) Indicate at least one other way to simulate a RV with the same distribution.

….

7.2 sample test # 2

Problem 1 (20 points)

A sample of 100 Moscow citizens was asked about voting at mayor elections. 51 voted for Sobyanin, 29 — for Navalny, others — for other candidates. The goal is to test is Sobyanin got more votes than Navalny.

(a) (4 points) Can you use for this purpose a standard test comparing two proportions? Explain your answer.

(б) (4 points) Let randomly selected voter with probability p_1 voted for Sobyanin, and with probability p_2 — for Navalny. Consider a random variable

$$\xi = \begin{cases} 1, & voted\ for\ Sobyanin, \\ -1, & voted\ for\ Navalny, \\ 0, & other. \end{cases}$$

Find $E(\xi),\ V(\xi)$.

(в) (4 points) Design an asymptotic test to check whether it was true that Sobyanin scored more votes than Navalny. What additional assumptions you have made?

(г) (4 points) Test the hypothesis of interest at 5% significance level with given survey results.

(д) (4 points) If you apply a standard test here on the equality of the two proportions, more often or less often than should the null hypothesis will be rejected?

Problem 2 (10 points)

The Newspaper reported that an opinion poll demonstrates that $45\% \pm 1.5\%$ of voters are going to vote in favour of the candidate P.

(a) (5 points) Assuming that they use 95% confidence, what was the sample size?

(b) (5 points) Given that information find interval estimate for the probability that in a sample of size 100 of arbitrary chosen voters, 50 or more will vote for the candidate P.

Problem 3 (20 points)

Let $X_1, \ldots, X_n$ be a sample from a discrete random variable which takes values –1, 0 and 1 wit the probabilities θ^2, $2\theta(1-\theta)$ and $(1-\theta)^2$ respectively, were $0 < \theta < 1$.

(c) (4 points) Find estimator for θ with Method of Moments.

(d) (4 points) Derive asymptotic distribution of this estimator.

(e) (4 points) Find ML estimator for θ.

(f) (4 points) Derive asymptotic distribution of this estimator.

(g) (4 points) Let prior distribution of θ is uniform at the interval $[0, 1]$. Find posterior distribution of θ and find Bayes point estimator as mean of the posterior distribution.

Problem 4 (20 points)

Let $X_1, \ldots, X_n$ be random sample from random variable $X \in \{0, 1, 2, \ldots\}$ with parameter $0 < \theta < 1$ and $P(X = k) = \dfrac{\theta(-\ln\theta)^k}{k!}$, $k = 0, 1, 2, \ldots$.

(a) (5 points) Find sufficient and complete statistics for θ. Give the proves.

(б) (5 points) Show that

$$\hat{\theta} = \begin{cases} 1, & if \ X_1 = 0, \\ 0, & elsewere. \end{cases}$$

Is an unbiased estimator for θ.

(в) (5 points) Find Rao-Blackwellization of $\hat{\theta}$. Prove that it is UMVUE for θ.

(г) (5 points) Find the lower boundary of the unbiased variance parameter estimator of θ, with Rao-Cramer inequality. If there exists an estimator of the parameter θ, for which this lower bound is achieved?

Problem 5 (12 points)

The takings of five different vendors at a street market were monitored over the seven different days of the same week. The total revenue each vendor made over the week (total in seven days) measured in pounds was 1017 for A, 965 for B, 477 for C and 408 for D. The value for vendor E is not given. The following information is provided as well.

Source	degrees of freedom	sum of squares	mean square	F-value
Day				4.17
Vendor				
Error			3876	
Total		303678		

(a) (4 points) Complete the table using the information provided above.

(b) (4 points) Is there a significant difference between the daily takings of different vendors? What about the takings of all vendors on different days?

(c) (4 points) Construct a 90% confidence interval for the difference between the daily takings of vendors A and D. Would you say there is evidence of a difference? (use all available information)

Problem 6 (18 points)

X is a random variable with density

$$f(x,\theta) = \begin{cases} \dfrac{1}{\theta} e^{-x/\theta} & 0 \le x, \\ 0, & \textit{иначе.} \end{cases}$$

Let we have a sample of size $n = 2$.

(a) (9 points) Find most powerful test for testing $H_0 : \theta = 1$ against the alternative $H_1 : \theta = 2$ at significance level $\alpha = 0.05$.

(b) (9 points) Find the power of this test.