

**Федеральное государственное автономное образовательное учреждение
высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"**

**Факультет экономических наук
Департамент статистики и анализа данных**

Программа дисциплины

Теория вероятностей и математическая статистика

для образовательных программ «Экономика: исследовательская программа»
и «Статистическое моделирование и актуарные расчеты»
направления 38.04.01 «Экономика» подготовки магистра

Автор программы:

Пересецкий А.А., д.э.н., профессор департамента прикладной экономики, aperesetsky@hse.ru
Панов В.А., Ph.D., доцент департамента статистики и анализа данных, urpanov@hse.ru

Одобрена на заседании департамента статистики и анализа данных

«___» 2016 г.

Руководитель департамента

Мхитарян В.С. [подпись]

Утверждена Академическим советом образовательной программы

«Экономика: исследовательская программа»

«___» 2016 г., № протокола _____

Академический руководитель образовательной программы

Суворов А.Д. [подпись]

Утверждена Академическим советом образовательной программы

«Статистическое моделирование и актуарные расчеты»

«___» 2016 г., № протокола _____

Академический руководитель образовательной программы

Конаков В.Д. [подпись]

Москва, 2016

Настоящая программа не может быть использована другими подразделениями университета и другими вузами без разрешения кафедры-разработчика программы.



1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов, обучающихся по магистерской программе "080100 Экономика: исследовательская программа". Программа разработана в соответствии с:

- Образовательным стандартом федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский университет “Высшая школа экономики”» для направления 38.04.01 Экономика подготовки магистра;
- Образовательной программой 38.04.01 «Экономика» подготовки магистра;
- Рабочим учебным планом университета по направлению 38.04.01. Экономика» подготовки магистра утвержденным в 2014 г.

2 Цели освоения дисциплины

Курс «Теория вероятностей и математическая статистика» является первым в ряду вероятностно-эконометрических курсов ТВ и МС – Эконометрика.

Основные цели курса

- Изложить основные методы элементарной теории вероятностей в случае дискретного и непрерывного пространства элементарных исходов. Основные одномерные и многомерные распределения и их характеристики.
- изложить основные понятия и методы математической и прикладной статистики;
- дать практический навык применения статистических методов в прикладных исследованиях;
- дать понятие о прикладных методах многомерного статистического анализа.

Предполагается, что студенты владеют методами математического анализа, линейной в объеме стандартных университетских курсов.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- Знать основные методы элементарной теории вероятностей и уметь их применять к простейшим моделям типа случайных блужданий, ветвящихся процессов, экстремальных значений, многомерного гауссовского распределения и т.д.
- Получить навыки работы с производящими функциями, с преобразованиями Лапласа (Фурье) и вычисления вероятностей и моментов в типичных проблемах элементарной теории вероятностей.
- Получить навыки работы с основными типами распределений возникающих в элементарной теории вероятностей и математической статистики.
- Знать основные методы прикладной статистики, их диапазон применения, их математические основы и предположения при которых они применимы.
- Уметь выбирать инструментарий прикладной и математической статистики, адекватный экономической задаче. Уметь верно интерпретировать полученные результаты.
- Приобрести опыт прикладного статистического исследования с реальными данными.

В результате освоения дисциплины студент осваивает следующие компетенции:



1. ИК-ТВМС7.1 Умеет работать с основными типами распределений возникающих в элементарной теории вероятностей и математической статистики

2. ИК-ТВМС7.2 Умеет работать с производящими функциями, с преобразованиями Лапласа (Фурье) и вычисления вероятностей и моментов в типичных проблемах элементарной теории вероятностей.

1. ИК- ТВМС7.1. Способен выбрать статистическую модель адекватно описывающую проблему.

2. ИК- ТВМС7.2. Способен использовать прикладной статистический пакет для оценки параметров выбранной модели.

3. ИК- ТВМС7.3. Способен понять и правильно интерпретировать результаты программного пакета..

1. СК-3, СК-М3, способен к самостоятельному освоению новых методов исследования, изменению научного и научно-производственного профиля своей деятельности;

2. СК-6, СК-М6, способен анализировать, верифицировать, оценивать полноту информации в ходе профессиональной деятельности, при необходимости восполнять и синтезировать недостающую информацию и работать в условиях неопределенности;

3. ПК-4,ИК-М 3.2НИД_5.4, способен представлять результаты проведенного исследования научному сообществу в виде статьи или доклада ;

4. ПК-9,ИК-М4.1_4.4_4.6_АД_5.4, способен находить данные, необходимые для анализа и проведения экономических расчетов, используя различные источники информации;

5. ПК-10, ИК-М4.4АД_5.4: способен работать с большими массивами разнообразной информации, составлять прогноз основных социально-экономических показателей деятельности предприятия, отрасли, региона и экономики в целом, в т.ч. используя современные информационно-компьютерные технологии;

4 Место дисциплины в структуре образовательной программы

Для магистерской программы "080100 Экономика: исследовательская программа" настоящая дисциплина является обязательной.

Изучение данной дисциплины базируется на следующих дисциплинах:

Математический анализ

Линейная Алгебра

Успешное освоение этой дисциплины необходимо для следующих дисциплин:

Эконометрика

Микроэкономика

5 Тематический план учебной дисциплины

№	Название раздела	Всего часов	Аудиторные часы			Самостоятельная работа
			Лекции	Семинары	Практические занятия	
1	Аксиоматика Колмогорова. Вероятностное пространство, случайные события, случайные величины.	10	2	2	0	6
2	Функция распределения, мат. ожидание, дисперсия, корреляция, медиана, независимость, свёртка. Дискретные и непрерывные случайные величины. Сингулярные распределения.	14	2	2	0	10



3	Обзор наиболее популярных распределений. Распределения, ассоциированные с нормальным. Экспоненциальные семейства распределений.	12	2	2	0	8
4	Случайные вектора. Маргинальные распределения. Формула преобразования плотности. Многомерное нормальное распределение.	14	4	4	0	6
5	Гауссовский вектор. Связь между независимостью и некоррелированностью для гауссовых случайных величин.	6	2	2	0	2
6	Условное мат.ожидание и условная дисперсия.	12	2	2	0	8
7	Порядковые статистики. Распределение экстремальных значений. Смеси распределений.	9	2	2	0	5
8	Производящая функция, преобразование Лапласа, характеристическая функция.	12	4	4	0	4
9	Вероятностные неравенства. Виды сходимости случайных величин (по распределению, по вероятности, почти наверно, в среднем квадратическом).	8	1	1	0	6
10	Закон больших чисел и центральная предельная теорема.	9	2	2	0	5
11	Безгранично-делимые и устойчивые распределения.	2	1	1	0	0
1	Статистическое оценивание параметров. Выборки. Свойства оценок. Несмешенность, эффективность, состоятельность.	7	2	1		4
2	Интервальное оценивание. Стандартные доверительные интервалы для параметров нормальной генеральной совокупности. Доверительные интервалы для среднего, дисперсии, разности средних, отношения дисперсий, пропорции, разности пропорций. Размер выборки.	14	4	2		8
3	Тестирование гипотез. Ошибки I, II рода. Р-значение теста. Тесты на значения параметров нормальной генеральной совокупности. Тесты на значения среднего, дисперсии, разности средних, отношения дисперсий, пропорции, разности пропорций.	14	4	2		8
4	Методы оценивания. Метод моментов. Метод максимального правдоподобия. Их свойства, примеры. Неравенство информации.	7	2	1		4
5	Критические статистики. Лемма Неймана-Пирсона. Тест отношения правдоподобия	12	4	2		6
6	Критерии согласия. Таблицы сопряженности. Критерий Пирсона. Тест Колмогорова-Смирнова	7	2	1		4
7	Байесовский подход к оцениванию.	7	2	1		4
8	Одно- и двух-факторный дисперсионный анализ	7	2	1		4
9	Понятие о непараметрических методах. Тесты Вилкоксона, серий. Ранговые коэффициенты корреляции.	6	2	1		3
10	Методы классификации. Дискриминантный анализ. Разделение смеси распределений. Кластер-анализ. Снижение размерности. Метод главных компонент. Факторный анализ		4	2		7
11	Достаточная статистика. Минимальная достаточная статистика. Теорема Рао-Блэквелла. Полная статистика. Теорема Лемана-Шеффе.	14	4	2		8



6 Формы контроля знаний студентов

Курс состоит из 2- равноправных частей:
модуль 1 — Теория вероятностей
модуль 2 — Математическая статистика

В конце модуля 1 — контрольная работа № 1.
В конце модуля 2 — контрольная работа № 2.

В течении курса студенты выполняют домашние работы и короткие (10-20 мин.) тесты (квизы) на лекциях.

Квизы и домашние задания оцениваются. Оценивается активность студентов на семинарах.

6.1 Критерии оценки знаний, навыков

Контрольные работы, домашние задания, оцениваются по 100-балльной шкале. С определенными весами эти оценки, а также баллы за активность на семинарах, агрегируются в итоговую оценку (тоже в 100-балльной шкале), которая затем конвертируется в 10-балльную шкалу. Пороги конвертирования заранее не фиксированы и определяются по распределению оценок в 100-балльной шкале.

Все элементы промежуточного контроля и направлены на понимание математических основ изучаемых методов, их диапазона применения, понимания численных результатов оценки моделей и умение практической работы с реальными данными.

6.2 Порядок формирования оценок по дисциплине

Накопленная оценка выставляется отдельно по первой (теория вероятностей) и второй (математическая статистика) частям курса. Накопленная оценка по всему курсу рассчитывается как среднее арифметическое из полученных оценок.

Накопленная оценка по первой части курса складывается из следующих компонент:

- 50% - оценка за итоговую контрольную работу №1;
- 30% - домашние работы;
- 20% - работа на семинарах.

По каждой компоненте ставится оценка от 1 до 10. Накопленная оценка умножается на 10 - в итоге студент получает оценку по 100-балльной системе.

Накопленная оценка по второй части курса складывается из следующих компонент (каждая оценивается по 100-балльной системе):

- 40% — оценки за домашние задания и квизы (письменные 15-мин. опросы на лекциях),
- 60% — оценка за контрольную работу № 2,

К итоговой сумме добавляются баллы за активность на семинарах (т.е. теоретически итоговая оценка может быть более 100 баллов).

Все домашние задания и «проект» принимаются только в бумажном виде и только в указанные сроки.

В связи с официально принятой в НИУ ВШЭ десятибалльной шкалой оценок, полученная студентом накопленная оценка по 100-балльной системе, конвертируются в оценку по 10-балльной системе. Пороги заранее не фиксированы. Ниже приведен пример порогов для конвертации.



Не допускается пересдача компонентов текущего контроля, за исключением контрольных работ № 1 и № 2.

В первой контрольной работе определяется одна или несколько базовых задач, решение которых необходимо для получения оценки выше или равной 4 баллам. Если базовые задачи не решены, то остальная часть работы не проверяется, и студенту ставится оценка менее 4 баллов. Допускается одна пересдача первой контрольной работы, для тех, кто пропустил ее по болезни и для тех, кто получил за эту контрольную оценку менее 4 баллов.

Если студент по контрольной работе № 2 набрал менее 25 баллов, то накопленная оценка по результатам текущего контроля приравнивается к этому числу баллов. Допускается одна пересдача второй контрольной работы, для тех, кто пропустил ее по болезни и для тех, кто получил за эту контрольную оценку менее 25 баллов.

Пересдача организуется во время консультаций преподавателя. Время и место проведения пересдачи согласуется между студентами и преподавателем.

При пересдаче выставление итоговой оценки происходит с использованием аналогичной процедуры, но пороги конвертации оценки из стобалльной в десятибалльную не обязательно совпадают с первоначальными порогами, т.к. сложность заданий контрольной работы во время пересдачи может отличаться.

Пример порогов:

Для выставления оценок по 10-балльной шкале возможно применение следующей базовой шкалы пересчета выставляемых в процентах оценок в оценки по десятибалльной шкале:
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.

При этом оценке "неудовлетворительно" соответствуют баллы 1, 2, 3; оценке "удовлетворительно" — 4, 5; оценке "хорошо" — 6, 7; оценке "отлично" — 8, 9, 10.

В зависимости от сложности курса и сложности контрольных работ пороговые градации могут изменяться преподавателем в диапазоне 5 процентных пунктов.

Базовые учебники

Ширяев А.Н. (2011) Вероятность. Том 1. М.МЦНМО.

Тутубалин, В.Н. (1992). Теория вероятностей и случайные процессы. М.: Издательство Московского университета.

Боровков А.А. (2016). Теория вероятностей. М.: URSS.

Billingsley, P. (1995). Probability and measure. John Wiley and Sons.

Grimmett G.R., Stirzaker D.R. (2003). One thousand exercises in probability. Oxford University Press.

Айвазян С.А., Мхитарян В.С. (2001). *Прикладная статистика и основы эконометрики*. (2-е издание). Том 1: Теория вероятностей и прикладная статистика. М.: ЮНИТИ.

Айвазян С.А., Мхитарян В.С. (2001). *Прикладная статистика в задачах и упражнениях*. М.: ЮНИТИ.

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

Дополнительная литература

Feller, W. (1968). An introduction to probability measure and its applications (in 2 volumes). John Wiley and Sons.



- Прохоров, А.В., Ушаков, В.Г. и Ушаков, Н.Г. Задачи по теории вероятностей. М.: Наука.
- Johnston A.R. and Bhattacharyya G.K. (2009). *Statistics. Principles and methods*. 6th 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.
- Гмурман В.Е. (2005). Теория вероятностей и математическая статистика. Москва. Высшая школа.
- Шведов А.С. (2005). Теория вероятностей и математическая статистика. Москва. ГУ ВШЭ.
- Wackerly D.D., Mendenhall W. III, Scheaffer R.L. (2008). *Mathematical statistics with applications*. 7th edition, Thompson.
- Boos D. D., and L. A. Stefanski (2013). Essential statistical inference: Theory and methods (Springer Texts in Statistics)

7 Образовательные технологии

Обязательное проведение семинаров по Математической статистике в компьютерных классах с примерами на реальных данных и решением теоретических задач. Пакеты STATA, EViews (опционально — R).



8 Оценочные средства для текущего контроля и аттестации студента

8.1 Примерный вариант контрольной работы №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 $E\tau$.

(b) Find $E\tau$ 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

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

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

Problem 3 Let $U_1, U_2, U_3 \sim \text{U}[0, 1]$ 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 $X_1, \dots, X_n \sim \text{N}(0, 1)$ be IID RVs. Let

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

Find $ES, \text{Var}(S)$.

Problem 5 Let $X_1, \dots, X_n \sim \text{Exp}(1)$ be IID RVs and $X_{(1)} < \dots < X_{(n)}$ 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 $U_1, \dots, U_n \sim \text{U}[0, 1]$ be IID RVs. Find the distribution of $-2 \ln(U_1 U_2 \dots U_n)$.

(b) Indicate at least one other way to simulate a RV with the same distribution.

....



8.2 Примерный вариант контрольной работы №2.

Задача 1 (20 баллов)

После голосования на выборах мэра Москвы было опрошено 100 человек, 51 из которых проголосовали за Собянина, 29 — за Навального, остальные — за других кандидатов. Наша цель — проверить, верно ли, что Собянин набрал больше голосов, чем Навальный.

(а) (4 балла) Можно ли для этого воспользоваться стандартным тестом сравнения двух пропорций? Объясните ответ.

(б) (4 балла) Пусть случайно выбранный избиратель с вероятностью p_1 голосует за Собянина, а с вероятностью p_2 — за Навального. Рассмотрите случайную величину

$$\xi = \begin{cases} 1, & \text{если избиратель голосует за Собянина,} \\ -1, & \text{если избиратель голосует за Навального,} \\ 0, & \text{иначе.} \end{cases}$$

Найдите $E(\xi)$, $V(\xi)$.

(в) (4 балла) Постройте асимптотический тест для проверки того, верно ли, что Собянин набрал больше голосов, чем Навальный. Какие дополнительные предположения вы при этом сделали?

(г) (4 балла) Протестируйте интересующую нас гипотезу на 5%-ным уровне значимости на данных условия задачи.

(д) (4 балла) Если применить здесь стандартный тест на равенство двух пропорций, чаще или реже чем следует, будет отвергаться нулевая гипотеза?

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.

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

(б) (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.

Задача 3 (20 баллов)

Пусть есть X_1, \dots, X_n — выборка значений дискретной случайной величины, принимающей значения $-1, 0$ и 1 с вероятностями θ^2 , $2\theta(1-\theta)$ и $(1-\theta)^2$ соответственно, где $0 < \theta < 1$.

(с) (4 балла) Найдите оценку θ методом моментов, используя первый момент

(д) (4 балла) Выведите асимптотическое распределение для оценки метода моментов

(е) (4 балла) Найдите оценку θ ММП.

(ф) (4 балла) Выведите асимптотическое распределение для оценки θ ММП

(г) (4 балла) Пусть априорное распределение θ — равномерное на отрезке $[0, 1]$. Найдите апостериорное распределение θ и найдите байесовскую точечную оценку как среднее по апостериорному распределению.



Задача 4 (20 баллов)

Дана случайная выборка X_1, \dots, X_n значений случайной величины $X \in \{0, 1, 2, \dots\}$ с параметром $0 < \theta < 1$ и

$$P(X = k) = \frac{\theta(-\ln \theta)^k}{k!}, \quad k = 0, 1, 2, \dots$$

(а) (5 баллов) Найдите полную достаточную статистику для параметра θ . Докажите ее достаточность и полноту.

(б) (5 баллов) Покажите, что статистика

$$\tilde{\theta} = \begin{cases} 1, & \text{если } X_1 = 0, \\ 0, & \text{иначе.} \end{cases}$$

является несмешенной оценкой параметра θ .

(в) (5 баллов) Постройте улучшение оценки $\tilde{\theta}$. Докажите, что полученная статистика является UMVUE для θ .

(г) (5 баллов) Найдите нижнюю границу дисперсии несмешенной оценки параметра θ , воспользовавшись неравенством Рао–Крамера. Существует ли несмешенная оценка параметра θ , для которой эта нижняя граница достигается?

Задача 5 (12 баллов)

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		

(а) (4 балла) Complete the table using the information provided above.

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

(с) (4 балла) 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)

Задача 6 (18 баллов)

Случайная величина X имеет плотность распределения

$$f(x, \theta) = \begin{cases} \frac{1}{\theta} e^{-x/\theta} & 0 \leq x, \\ 0, & \text{иначе.} \end{cases}$$

Имеется выборка размера $n = 2$.

(а) (9 баллов) Найдите наиболее мощный тест для тестирования гипотезы $H_0: \theta = 1$ против альтернативной гипотезы $H_1: \theta = 2$ с уровнем значимости $\alpha = 0.05$.

(б) (9 баллов) Найдите мощность этого теста.

9 Учебно-методическое и информационное обеспечение дисциплины

Базовые учебники

Ширяев А.Н. (2011) Вероятность. Том 1. М.МЦНМО.



Тутубалин, В.Н. (1992). Теория вероятностей и случайные процессы. М.: Издательство Московского университета.

Billingsley, P. (1995). *Probability and measure*. John Wiley and Sons.

Боровков А.А. (2016). Теория вероятностей. М.: URSS.

Grimmett G.R., Stirzaker D.R. (2003). *One thousand exercises in probability*. Oxford University Press.

Айвазян С.А., Мхитарян В.С. (2001). *Прикладная статистика и основы эконометрики*. (2-е издание). Том 1: Теория вероятностей и прикладная статистика. М.: ЮНИТИ.

Айвазян С.А., Мхитарян В.С. (2001). *Прикладная статистика в задачах и упражнениях*. М.: ЮНИТИ.

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

Дополнительная литература

Feller, W. (1968). *An introduction to probability measure and its applications* (in 2 volumes). John Wiley and Sons.

Прохоров, А.В., Ушаков, В.Г. и Ушаков, Н.Г. Задачи по теории вероятностей. М.: Наука.

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.

Гмурман В.Е. (2005). Теория вероятностей и математическая статистика. Москва. Высшая школа.

Шведов А.С. (2005). Теория вероятностей и математическая статистика. Москва. ГУ ВШЭ.

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

Boos D. D., and L. A. Stefanski (2013). *Essential statistical inference: Theory and methods* (Springer Texts in Statistics)

Программные средства

STATA, Eviews, (опционально — R)

10 Материально-техническое обеспечение дисциплины

Компьютерный класс для семинаров.

Большая (обязательно) доска в лекционном зале и компьютерный проектор.