

Государственный университет –

Высшая школа экономики

Международный институт экономики и финансов

Программа дисциплины

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ВВЕДЕНИЕ В СТАТИСТИКУ

для направления 080100.62 экономика

подготовки бакалавра

Автор Катышев П.К.

Москва 2012

1. Описание курса

Теория вероятностей и введение в статистику – двухсеместровый курс для студентов первого года обучения МИЭФ. Это вводный курс для студентов, специализирующихся в области экономики. Курс преподается на русском и английском языках.

В курсе изучаются базовые понятия статистики: описательные статистики, понятие генеральной совокупности и выборки, оценивание параметров, статистическая проверка гипотез и т.п. Кроме того, изучаются разделы теории вероятностей, необходимые для целостного изложения курса.

2. Цели курса

Основная цель курса – дать студентам систематические знания в области начальной статистики. Они должны понимать предмет и освоить основные методы статистического анализа. Студенты должны научиться проводить разведочный анализ данных (находить среднее, медиану, среднеквадратичное отклонение и другие описательные статистики), представлять данные графически (гистограммы, диаграммы ствол–листья, коробчатые диаграммы, точечные диаграммы).

Предполагается, что студенты должны получить начальное представление о методах сбора данных, о планировании и проведении эксперимента. Они должны научиться выявлять возможные источники смещений при сборе данных и проведении эксперимента.

У них должно сложиться понимание различия между генеральной совокупностью и выборкой и, соответственно, между теоретическими и выборочными характеристиками. Статистику даже начальную, нецелесообразно преподавать без элементов теории вероятностей, поэтому составной частью курса является изучение базовых понятий и результатов теории вероятностей. Студенты должны понимать, что такое вероятностное пространство, случайное события, вероятность события. Они должны научиться вычислять вероятности сложных событий, уметь решать простейшие комбинаторные задачи, освоить использование формулы полной вероятности и формулы Байеса. У студентов должно быть ясное представление о том, что такое случайная величина и ее распределение.

Студенты должны научиться формулировать и решать традиционные задачи начальной статистики: точечное и интервальное оценивание параметров, статистическая проверка гипотез, корреляционный анализ. Одна из целей курса – подготовить студентов к изучению эконометрики на основе изучения простейших моделей парной и множественной регрессии.

Курс не является математически строгим. Как следствие, доказательства многих теорем и даже точная формулировка результатов часто опускаются. Важной частью курса является решение задач. В основе задач – попытка проиллюстрировать различные способы применения теории на практике. В процессе обучения

студенты также выполняют компьютерные задания с реальными данными, вырабатывают практические навыки и интуицию.

По окончании курса студент должен понимать теорию, лежащую в основе статистической науки, уметь выполнять необходимые вычисления и применять стандартные методы на практике.

В конце первого года обучения студенты должны обладать набором знаний и навыков, необходимых для сдачи экзамена AP Statistics.

3. Тематический план учебной дисциплины

N	Наименование тем и разделов	ВСЕГО (часов)	Аудиторные занятия (час)		Самостоятельная работа
			в том числе		
			Лекции	Практические занятия	
1	Первичный анализ данных, описательные статистики	24	6	6	12
2	Сбор данных, планирование и проведение эксперимента	22	6	6	10
3	Элементы теории вероятностей	24	6	6	12
4	Дискретные случайные величины	16	4	4	8
5	Непрерывные случайные величины	16	4	4	8
6	Двумерные распределения	16	4	4	8
7	Предельные теоремы	16	4	4	8
8	Распределение выборочных статистик	16	4	4	8
9	Точечное оценивание параметров	24	6	6	12
10	Интервальное оценивание параметров	32	8	8	16
11	Статистическое тестирование гипотез	40	10	10	20
12	Модель парной регрессии	24	6	6	12
	ИТОГО	270	68	68	134

4. Методы

В курсе используются следующие методы и формы работы:

- Лекции (2 часа неделя)
- Семинары (2 часа в неделю)
- Еженедельные домашние задания
- Консультации преподавателя
- Самостоятельная работа с литературой.

Курс включает 42 часа лекций и 42 часа семинаров.

5. Формы контроля по учебной дисциплине

- Домашние задания
- Контрольные работы
- Промежуточный осенний экзамен (180 мин.)
- Зимний экзамен (180 мин.)
- Промежуточный весенний экзамен (180 мин.)
- Внешний экзамен

6. Определение оценок

Студенты сдают в первом модуле промежуточный письменный экзамен, а также итоговый экзамен конце второго модуля. Экзамены включают вопросы множественного выбора и открытые вопросы.

В четвертом модуле студенты сдают промежуточный экзамен, имеющий аналогичный формат. После этого студенты сдают экзамен AP Statistics.

Оценка за первые два модуля складывается из

- оценки за промежуточный осенний экзамен (30%)
- оценки за зимний экзамен (60%)
- средней оценки за контрольные и домашние работы в 1-м и 2-м модулях (10%)

Оценка за год складывается из

- оценки за первые два модуля (25%)
- оценки за промежуточный весенний экзамен (35%)
- оценки за внешний экзамен (35%)
- средней оценки за контрольные и домашние работы в 3-м и 4-м модулях (5%)

7. Основная литература

1. Wonnacott R.J. and Wonnacott T.H. Introductory statistics for business and economics. John Wiley & Sons, fourth edition, 1985 (WW).

8. Дополнительная литература

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.

9. Интернет-ресурсы

1. <http://www.stat.yale.edu/Courses/1997-98/101/expdes.htm>
2. <http://www.itl.nist.gov/div898/handbook/pri/section3/pri3.htm>
3. http://www.cas.lancs.ac.uk/glossary_v1.1/main.html
4. <http://home.xnet.com/~fidler/triton/math/review/mat170/mat170.htm>
5. <http://www.statsoft.com/textbook/esc.html>
6. <http://mail.pittstate.edu/~winters/tutorial/>

10. Содержание курса

1. **Первичный анализ данных, описательные статистики (*WW Ch.2*)**
 - 1.1. Графическое представление одномерных данных. Точечный график. Диаграмма «ствол и листья». Гистограмма.
 - 1.2. Особенности данных. Выбросы. Кластеры. Форма гистограммы
 - 1.3. Описательные статистики. Меры среднего: арифметическое среднее, медиана, мода, среднее геометрическое. Меры разброса: размах, среднеквадратическое отклонение, Межквартильный размах, среднее абсолютное отклонение, среднее относительное отклонение. Представление данных с помощью коробчатой диаграммы.
 - 1.4. Преобразование элементарных статистик при линейном преобразовании данных.
 - 1.5. Меры положения наблюдения в выборке: квартили, перцентили, z-шкала.
 - 1.6. Вычисления с группированными данными.
2. **Сбор данных, планирование и проведение эксперимента (*WW, Ch.1*)**
 - 2.1 Методы сбора данных: перепись, выборочное обследование, эксперимент, наблюдения.
 - 2.2 Генеральная совокупность, выборка, случайная выборка.
 - 2.3 Источники смещения при выборке и обследовании.
 - 2.4 Типы выборки: простая случайная выборка, стратифицированная случайная выборка, кластерная выборка.
 - 2.5 Планирование и проведение эксперимента.
 - 2.6 Контрольные группы, случайное назначение, повторение.

- 2.7 Источники смещения при проведении эксперимента. Смешивающие факторы, эффект плацебо и «слепоты».
- 2.8 Полностью рандомизированный эксперимент. Разбиение на блоки.

3. Элементы теории вероятностей (WW Ch. 3, 4)

- 3.1. Эксперимент со случайным исходом. Понятие о вероятности. Вероятность события как предел относительной частоты его появления в серии независимых испытаний. Понятие закона больших чисел.
- 3.2. Пространство элементарных исходов как математическая модель эксперимента со случайным исходом. Алгебра событий. Несовместные события.
- 3.3. Вероятность в пространстве элементарных исходов. Классическая вероятность. Простейшие комбинаторные схемы. Вероятность суммы событий.
- 3.4. Условные вероятности. Вероятность произведения событий. Независимые события.
- 3.5. Формула полной вероятности. Формула Байеса

4. Дискретные случайные величины (WW Ch. 3, 4)

- 4.1. Примеры дискретных случайных величин. Распределение дискретных случайных величин. Частоты и кумулятивные частоты.
- 4.2. Среднее значение (математическое ожидание). Дисперсия. Стандартное отклонение.
- 4.3. Последовательность независимых испытаний. Биномиальное распределение. Геометрическое распределение.

5. Непрерывные случайные величины (WW Ch. 3, 4)

- 5.1. Примеры непрерывных случайных величин. Функция распределения. Плотность распределения. Среднее значение (математическое ожидание). Дисперсия. Стандартное отклонение.
- 5.2. Нормальное распределение, его свойства. Таблицы нормального распределения.
- 5.3. Линейное преобразование случайной величины.

6. Двумерные распределения

- 6.1. Совместное распределение двух дискретных случайных величин. Маргинальные распределения. Условные распределения. Условное среднее.
- 6.2. Независимые случайные величины.
- 6.3. Коэффициент ковариации. Коэффициент корреляции как мера линейной связи двух случайных величин. Некоррелированность и независимость. Математическое ожидание и дисперсия линейной комбинации двух случайных величин.

7. Предельные теоремы

- 7.1. Закон больших чисел.
- 7.2. Центральная предельная теорема. Аппроксимация биномиального распределения нормальным.

8. Распределение выборочных статистик

- 8.1. Распределение выборочного среднего, выборочной пропорции.
- 8.2. Распределение разности двух пропорций. Распределение разности двух независимых выборочных средних.
- 8.3. t -распределение (распределение Стьюдента), хи-квадрат распределение.

9. Точечное оценивание параметров (WW Ch. 6, 7, 8, 9, 17)

- 9.1. Понятие точечного оценивания параметров генеральной совокупности. Примеры точечных оценок: выборочное среднее и выборочная дисперсия.
- 9.2. Свойства оценок: несмещенность, эффективность, состоятельность.
- 9.3. Оценки среднего и дисперсии.
- 9.4. Оценка пропорции.

10. Интервальное оценивание параметров

- 10.1. Понятие доверительного интервала. Доверительный интервал для среднего значения генеральной совокупности. Нормальная аппроксимация при больших выборках. Случай малых выборок (распределение Стьюдента).
- 10.2. Доверительный интервал для разности двух средних (парные и независимые выборки).
- 10.3. Доверительный интервал для разности двух пропорций.
- 10.4. Двусторонние и односторонние доверительные интервалы.

11. Статистическое тестирование гипотез

- 11.1. Понятие гипотезы и статистического теста. Тестирование гипотез о среднем генеральной совокупности. Использование доверительных интервалов и тест-статистик.
- 11.2. Двусторонние и односторонние тесты. Понятие P -значения теста.
- 11.3. Понятие об ошибках первого и второго рода. Значимость теста, мощность теста.
- 11.4. Стандартные тесты:
 - 11.4.1. среднее генеральной совокупности
 - 11.4.2. пропорция
 - 11.4.3. разность средних независимых и парных генеральных совокупностей
 - 11.4.4. разность пропорций
- 11.5. Критерий согласия Пирсона. Таблицы сопряженности

12. Модель парной регрессии (WW Ch.11, 12)

- 12.1. X - Y график. Подгонка прямой. Метод наименьших квадратов.
- 12.2. Преобразования, приводящие к линейной модели
- 12.3. Выбросы.
- 12.4. Прогнозные значения.
- 12.5. Остатки и ошибки регрессии.

12.6. Статистические свойства оценок. Доверительный интервал для коэффициента наклона. Тестирование гипотез о коэффициенте наклона.