

---

У Ч Е Б Н И К И  
ВЫСШЕЙ ШКОЛЫ ЭКОНОМИКИ

---

**ВШЭ**  
**HSE**

---

У Ч Е Б Н И К И  
ВЫСШЕЙ ШКОЛЫ ЭКОНОМИКИ

---

**ВШЭ**  
**HSE**

---

Ю.Н.Толстова

МАТЕМАТИКО-СТАТИСТИЧЕСКИЕ  
МОДЕЛИ

# В СОЦИОЛОГИИ

Математическая статистика  
для социологов

*Учебное пособие*

Второе издание



Издательский дом ГУ ВШЭ

---

Москва 2008

УДК 519.2(075)  
ББК 22.172  
Т52



Издание осуществлено в рамках  
Инновационной образовательной программы ГУ ВШЭ  
«Формирование системы аналитических компетенций  
для инноваций в бизнесе и государственном управлении»

**Рецензент:**  
кандидат физико-математических наук,  
профессор *А.А. Макаров*

ISBN 978-5-7598-0600-4

© Толстова Ю.Н., 2007  
© Оформление.  
Издательский дом ГУ ВШЭ, 2007

## ОГЛАВЛЕНИЕ

<b>Введение</b> .....	11
<b>В.1. Основная цель курса, адресат</b> .....	11
<b>В.2. Проблемы преподавания математических дисциплин студентам-социологам</b> .....	11
<b>В.3. Особенности курса</b> .....	12
<b>В.4. Общие организационные требования</b> .....	13
<b>В.5. Связь с курсом теории вероятностей</b> .....	15
<b>В.6. Специфика представления библиографии</b> .....	15
<b>Раздел I. Общее представление о математической статистике. Предварительные сведения об основном объекте ее изучения — случайных величинах (измерение, стандартизация, виды распределений, предельные теоремы)</b>	
Тема 1. Объект, предмет, цели и задачи математической статистики .....	19
1.1. Понятия выборки и генеральной совокупности.....	19
1.2. Понятие случайной величины.....	20
1.3. Понятие статистической закономерности.....	22
1.4. Объект изучения для математической статистики .....	29
1.5. Предмет изучения для математической статистики .....	29
1.6. Основная задача математической статистики .....	29
1.7. Методологические принципы использования математики в социологии .....	30
1.8. Некоторые замечания о терминах, используемых в западной литературе .....	33
<i>Повторение отдельных фрагментов курса по теории вероятностей</i> .....	36
<i>Примеры задач</i> .....	38
<i>Литература к теме 1</i> .....	40

Тема 2. Общее представление о социологических шкалах .....	43
2.1. Общие принципы понимания измерения в социологии .....	43
2.2. Определение номинальной, порядковой, интервальной шкалы .....	44
2.3. Проблема адекватности математического метода .....	46
<i>Примеры задач</i> .....	49
<i>Литература к теме 2</i> .....	50
Тема 3. Стандартизация значений случайных величин. Виды некоторых специфических распределений, использующихся при переносе результатов с выборки на генеральную совокупность....	51
3.1. Стандартизация (нормировка) значений случайной величины: способы и цели.....	51
3.2. Нормальное распределение (повторение) .....	52
3.3. Распределение $\chi^2$ .....	53
3.4. Распределение Стьюдента ( <i>t</i> -распределение).....	57
3.5. Распределение Фишера ( <i>F</i> -распределение, распределение дисперсионного отношения).....	59
<i>Повторение отдельных фрагментов курса     по теории вероятностей</i> .....	60
<i>Примеры задач</i> .....	62
Тема 4. Предельные теоремы .....	64
4.1. Центральная предельная теорема.....	64
4.2. Закон больших чисел .....	67
<i>Литература к теме 4</i> .....	68
<b>Раздел II. Оценивание параметров</b>	
Тема 5. Еще раз о задачах математической статистики. Точечное оценивание параметров .....	71
5.1. О задачах математической статистики .....	71
5.2. Точечные оценки параметров и предъявляемые к ним требования.....	72

Тема 6. Интервальное оценивание параметров .....	82
6.1. Понятие доверительного интервала и принципы его построения (на примере математического ожидания).....	82
6.2. Определение объема выборки .....	90
6.3. Доверительный интервал для медианы.....	92
6.4. Доверительный интервал для доли.....	93
6.5. Связь средних ошибок среднего арифметического и доли, обобщение этого факта на многомерный анализ .....	94
<i>Примеры задач</i> .....	95
<b>Раздел III. Проверка статистических гипотез</b>	
Тема 7. Общее представление о статистической гипотезе. Проверка статистической гипотезы об отсутствии связи (критерий $\chi^2$ ). .....	101
7.1. Общее представление о статистической гипотезе .....	101
7.2. Логика проверки статистической гипотезы. Использование принципа невозможности реализации маловероятных событий .....	103
7.3. Проверка гипотезы об отсутствии связи между номинальными признаками на основе критерия $\chi^2$ .....	108
<i>Примеры задач</i> .....	111
Тема 8. Проверка гипотезы о равенстве средних .....	113
8.1. Понятие зависимых и независимых выборок.....	113
8.2. Проверка гипотезы для независимых выборок .....	114
8.3. Проверка гипотезы для зависимых выборок .....	117
<i>Примеры задач</i> .....	118
Тема 9. Направленные и ненаправленные альтернативные гипотезы. Односторонние и двусторонние критерии .....	120
9.1. Направленные и ненаправленные альтернативные гипотезы .....	120
9.2. Односторонние и двусторонние критерии.....	121

Тема 10. Проверка статистических гипотез: о равномерности генерального распределения, равенстве дисперсий, равенстве нулю коэффициента корреляции, равенстве двух долей; ошибки первого и второго рода .....	123
10.1. Проверка гипотезы о равномерности генерального распределения с помощью критерия $\chi^2$ .....	123
10.2. Проверка гипотезы о равенстве двух дисперсий .....	125
10.3. Проверка гипотезы о равенстве нулю коэффициента корреляции .....	126
10.4. Проверка гипотезы о равенстве долей .....	127
Примеры задач.....	128
Тема 11. Методологические аспекты проверки математико-статистических гипотез.....	130
11.1. Ошибки первого и второго рода.....	130
11.2. Пример влияния содержательного характера задачи на выбор уровня значимости.....	134
11.3. Различие между статистической и содержательной гипотезой .....	136
<b>Раздел IV. Проблема изучения причинно-следственных отношений и эксперимент в социологии; основные идеи дисперсионного анализа</b>	
Тема 12. Методологические аспекты изучения причинно-следственных отношений с помощью математических методов. Эксперимент в социологии .....	144
12.1. Проблема изучения причинно-следственных отношений .....	144
12.2. Невозможность полностью формализовать понятия причины и следствия. Выделение двух основных направлений изучения причинно-следственных отношений: построение структурных уравнений и проведение эксперимента .....	154
12.3. Роль математической статистики в проведении эксперимента. Нестатистический (индуктивный) подход: эксперимент по Миллю.....	156

Примеры задач.....	162
Литература к теме 12.....	163
Тема 13. Корреляционное отношение.....	168
13.1. Линейная и нелинейная связи. Границы применимости коэффициента корреляции как показателя связи между изучаемыми переменными .....	168
13.2. Корреляционное отношение. Общее представление о внутригрупповом и межгрупповом разбросе.....	170
13.3. Проблемы, не решаемые с помощью корреляционного отношения .....	176
13.4. Соотношение между разными видами сумм квадратов .....	177
Примеры задач.....	178
Литература к теме 13.....	178
Тема 14. Однофакторный дисперсионный анализ .....	179
14.1. Однофакторный дисперсионный анализ как метод анализа результатов эксперимента при изучении причинно-следственных отношений.....	179
14.2. Модель однофакторного дисперсионного анализа.....	181
14.3. Однофакторный дисперсионный анализ как проверка статистической гипотезы .....	184
14.4. О понимании термина «влияет» (или что значит доказать наличие причинно-следственного отношения с помощью дисперсионного анализа) .....	187
14.5. Метод множественных сравнений для однофакторного дисперсионного анализа .....	187
Примеры задач.....	189
Тема 15. Двухфакторный дисперсионный анализ .....	191
15.1. Двухфакторный дисперсионный анализ как метод анализа результатов эксперимента при изучении причинно-следственных отношений .....	191
15.2. Модель двухфакторного дисперсионного анализа.....	191

15.3. Двухфакторный дисперсионный анализ как проверка статистических гипотез.....	195
<i>Примеры задач</i> .....	196
<i>Литература к темам 14 и 15</i> .....	197
Приложение 1. Литература .....	201
Приложение 2. Примерные экзаменационные вопросы .....	204
Приложение 3. Ориентировочные темы эссе (рефератов).....	207
Приложение 4. Таблицы распределений.....	222

## ВВЕДЕНИЕ

### В.1. Основная цель курса, адресат

Курс рассчитан на студентов-социологов и посвящен изложению основ математической статистики. По существу он является продолжением курса по теории вероятностей. Как известно, подобные курсы традиционно читаются студентам самых разных специальностей. Объясняется это тем, что изучение статистических закономерностей требуется практически в любой отрасли человеческого знания. В отечественной литературе в области статистики широко представлены учебники (в том числе переводные) и методические пособия самого разнообразного плана: с разной шириной охвата проблематики, рассчитанные на читателей с различным уровнем подготовки и т.д. Кажется бы, преподавание математической статистики для студентов-прикладников стало рутинным делом. Тем не менее, предлагаемое учебное пособие имеет ряд особенностей, позволяющих считать его в какой-то степени оригинальным (именно это обусловило изменение названия курса). Особенности эти вызваны желанием автора сделать курс хорошо воспринимаемым именно социологами.

### В.2. Проблемы преподавания математических дисциплин студентам-социологам

Опыт показывает, что студенты-социологи часто бывают настроены на «гуманитарный лад», и либо отрицают необходимость серьезного рассмотрения каких бы то ни было математических методов, либо делают это формально, в глубине души считая соответствующие знания для себя лишними. В результате — если не отсутствие знаний, то освоение материала на «абстрактном» уровне, без всякого сопряжения с практикой проведения социологических исследований. Во всяком случае, автору неоднократно приходилось наблюдать, что даже добросовестные студенты плохо представляют себе, как использовать знания, полученные ими в курсах теории вероят-

ностей и математической статистики, на практике. Преодолеть соответствующую проблему, на наш взгляд, можно путем определенной «привязки» курса к социологическим проблемам.

### В.3. Особенности курса

Основной чертой предлагаемого учебного пособия, отличающего его от других учебников соответствующей направленности, является прежде всего то, что все вводимые теоретические положения сопровождаются иллюстрациями их использования в социологических исследованиях. В качестве примеров случайных событий служат события, каждое из которых состоит в том, что какой-либо респондент обладает определенным сочетанием значений рассматриваемых признаков. Сами признаки служат примерами случайных величин (вместо вероятностей в примерах, естественно, фигурируют относительные частоты).

Еще одна особенность работы состоит в том, что в ней большое внимание уделяется проблеме измерения исходных данных. Дело в том, что в социологии проблемы выбора способа получения данных и метода их анализа (в том числе и с помощью алгоритмов математической статистики) не могут решаться отдельно, поскольку отражают две стороны одного и того же процесса. В предлагаемом курсе это проявляется прежде всего в том, что, говоря о параметрах распределений, мы соотносим их с типами шкал, использованных при получении исходных данных.

Существенное внимание в курсе уделяется описанию роли статистического подхода в социологии; обсуждается возможность обеспечения того комплекса условий, реализация которого приводит к появлению интересующих социолога случайных событий; в частности, затрагивается проблема существования случайных величин. Рассматриваются часто встречающиеся в социологии ситуации, в которых не выполняются условия реализации известных математико-статистических методов, показывается, как может действовать социолог в таких случаях.

В определенной мере затрагивается история применения статистического подхода к изучению социальных явлений. Потребность в этом объясняется следующими обстоятельствами. Статистический

подход, зародившись в XVII в. именно при изучении общества, потом, на стыке XIX и XX вв., начал необоснованно отвергаться некоторыми обществоведами. В какой-то мере возникший кризис был преодолен. Но сейчас, через сто лет, история повторяется. И ретроспективный анализ работ наших предшественников оказывается весьма полезным для современной ситуации.

Одним из проявлений кризисности современной ситуации с использованием математического языка в социологии является присущая многим социологам механистичность использования методов, отсутствие потребности в анализе задействованных в методах моделей, в сопряжении их со смыслом решаемой задачи. Для исправления такого положения и представляется полезным обращение к истокам, корням, рассмотрению тех обстоятельств, которые привели к появлению того или иного метода.

Разговор о роли математической статистики в социальных исследованиях в данной работе ведется на фоне обсуждения общих принципов использования математики в социологии. И в качестве основного пласта содержательных задач, выбранного для иллюстраций, используются задачи изучения причинно-следственных отношений. Это представляется естественным, поскольку в содержательном плане методы математической статистики в значительной мере направлены именно на решение соответствующих проблем. Для обеспечения возможности серьезного разговора по поводу связи содержания социологической задачи и математического формализма в работе коротко рассматривается развитие понятия причины и анализируется роль статистического (и нестатистического) подхода к ее изучению. И здесь мы также пытаемся обратиться к истокам соответствующих теоретических положений, руководствуясь сформулированным выше принципом: в кризисной ситуации эффективным может быть обращение к «корням».

### В.4. Общие организационные требования

Курс длится два модуля. В конце первого модуля — контрольная работа, в конце второго — экзамен<sup>1</sup>. К середине второго модуля не-

<sup>1</sup> Хочется отметить негативное отношение автора к современной тенденции

обходимо сдать реферат (эссе). Примерные темы приведены в Приложении 3. Они охватывают следующие направления: анализ исторических корней математической статистики, их общности с корнями эмпирической социологии; изучение методических достижений русской земской статистики; оценка гносеологических аспектов использования статистических закономерностей в социологии; рассмотрение проблемы построения выборки; осмысление методологических вопросов, возникающих при получении нового социологического знания с помощью математических методов.

Чтение лекций сопровождается семинарскими занятиями, примерная тематика которых отражена в списках ориентировочных задач, приведенных после раскрытия большинства тем. Кроме того, на семинаре должна осуществляться та связь с курсом теории вероятностей, о которой идет речь ниже.

---

безоговорочной замены устных экзаменов письменными. У устного и письменного экзаменов примерно те же достоинства и недостатки, что и у мягких и жестких методов сбора данных. Устный экзамен позволяет преподавателю более адекватно понять, что студент знает, а что не знает. Случайная ошибка студента здесь играет гораздо меньшую роль, чем в письменной работе. Списанная или механически повторенная фраза быстро обретает свою истинную значимость в глазах преподавателя. Экзаменатор и экзаменуемый в процессе устного экзамена взаимно воздействуют друг на друга, каждый из них имеет шанс получить в процессе этого общения нечто новое и полезное для себя. Но, с другой стороны, конечно, скорость проведения письменного экзамена намного выше скорости устного. Здесь необходимо упомянуть о тестовом опросе, который в особой мере повышает эту скорость. У нас имеется глубокое убеждение в том, что такой опрос в принципе не может позволить в должной мере оценить глубину понимания материала студентами. Напротив, он стимулирует студента осваивать материал весьма поверхностно. Использование тестов может быть полезно только на этапах промежуточного контроля. Для нас важнее адекватная оценка степени понимания студентом пройденного материала (вкуче с наличием контакта с экзаменуемым), а не возможность опросить как можно больше студентов за единицу времени. О преимуществах и недостатках мягких и жестких методов сбора данных много сказано в социологической литературе, но почему-то до сих пор не проведено профессиональных исследований, направленных на оценку возможности переноса этих результатов на экзаменационную ситуацию.

## В.5. Связь с курсом теории вероятностей

Успешное освоение предлагаемого курса возможно только после знакомства студента с элементами теории вероятностей в том объеме, который обычно предполагается рассчитанными на социологов учебными программами соответствующей дисциплины. Ниже указывается, какие именно знания по теории вероятностей требуются для освоения того или иного фрагмента настоящего курса. Такие указания оформлены в виде специальных рубрик (как было сказано, это — материал для семинарских занятий).

## В.6. Специфика представления библиографии

В отечественной литературе имеется очень много работ (в том числе переводных), прекрасно описывающих основные положения математической статистики (см. Приложение 1). В них можно найти материал почти по всем темам. К *основной* литературе мы отнесли работы, либо выпущенные в последние годы, либо ориентированные на социологов или по способу изложения, или по специфике рассматриваемых аспектов (к сожалению, эти работы зачастую опубликованы достаточно давно).

После некоторых лекций приведены списки книг, содержание которых более узко касается только рассматриваемой лекции.

Указываются отдельные работы и внутри текста (имеются в виду работы, содержание которых выходит за пределы стандартных курсов по теории вероятностей и математической статистике).

Т52 **Толстова, Ю. Н.** Математико-статистические модели в социологии (математическая статистика для социологов) : учеб. пособие / Ю. Н. Толстова ; Гос. ун-т — Высшая школа экономики. — 2-е изд. — М. : Изд. дом ГУ ВШЭ, 2008. — 243,[1] с. — (Учебники Высшей школы экономики). — 1000 экз. — ISBN 978-5-7598-0600-4 (в обл.).

УДК 519.2(075)  
ББК 22.172

*Учебное издание*

Серия «Учебники Высшей школы экономики»

Толстова Юлиана Николаевна

**Математико-статистические модели в социологии  
(математическая статистика для социологов)**

*Второе издание*

Зав. редакцией *О.А. Шестопалова*

Редактор *Л.И. Кузнецова*

Художественный редактор *А.М. Павлов*

Компьютерная верстка и графика: *Ю.Н. Петрина*

Корректор *Н.А. Балашова*

Издано при содействии ООО «МАКС Пресс»

Подписано в печать 23.11.07. Формат 60×88<sup>1/16</sup>  
Гарнитура Newton C. Усл. печ. л. 14,79. Уч.-изд. л. 12,79  
Тираж 1000 экз. Заказ № Изд. № 964

ГУ ВШЭ. 125319, Москва, Кочновский проезд, 3  
Тел./факс: (495) 772-95-71