

П.С. Бондаренко, Г.В. Горелова, И.А. Кацко

# ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Под редакцией проф. **И.А. Кацко** и проф. **А.И. Трубилина**

Допущено

Министерством сельского хозяйства Российской Федерации  
в качестве **учебного пособия**

для студентов высших аграрных учебных заведений,  
обучающихся по направлениям «Агрономия», «Садоводство»,  
«Продукты питания», «Экономика», «Бизнес-информатика»,  
«Прикладная информатика», «Информационные системы и технологии»,  
«Строительство», «Менеджмент»

**BOOK.ru**

ЭЛЕКТРОННО-БИБЛИОТЕЧНАЯ СИСТЕМА

**КНОРУС • МОСКВА • 2019**

УДК 591.2(075.8)  
ББК 22.17я73  
Б81

**Рецензенты:**

**Л.И. Ниворожкина**, заведующая кафедрой математической статистики, эконометрики и актуарных расчетов Ростовского государственного экономического университета (РИНХ), заслуженный деятель науки РФ, д-р экон. наук, проф.,  
**С.Г. Чефранов**, заведующий кафедрой информационной безопасности и прикладной информатики ФГБОУ ВО «МГТУ», д-р экон. наук, доц.

**Бондаренко, Петр Сергеевич.**

**Б81** Теория вероятностей и математическая статистика : учебное пособие / П.С. Бондаренко, Г.В. Горелова, И.А. Кацко ; под ред. И.А. Кацко, А.И. Трубилина. — Москва : КНОРУС, 2019. — 390 с. — (Бакалавриат).

**ISBN 978-5-406-06704-8**

Рассмотрены основные вопросы теории вероятностей и математической статистики в соответствии с программой подготовки бакалавров. Для углубленного изучения курса приведены приложения теории вероятностей в компьютерных науках (computer science), а также для отражения современных тенденций обработки данных, в известной мере опирающихся на идеологию теории вероятностей и математической статистики, дан обзор современных направлений анализа данных в контексте развития общества и информационных технологий. Пособие отражает опыт преподавания авторами дисциплины в Кубанском государственном аграрном университете и Инженерно-технологической академии Южного федерального университета.

Соответствует ФГОС ВО последнего поколения.

*Для студентов, обучающихся по экономическим и агроинженерным, а также связанным с ИТ направлениям, преподавателей, аспирантов и слушателей курсов повышения квалификации, применяющих в своей практике вероятностные и статистические методы.*

**УДК 591.2(075.8)  
ББК 22.17я73**

Бондаренко Петр Сергеевич  
Горелова Галина Виктровна  
Кацко Игорь Александрович

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ  
И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА**

Изд. № 18247. Формат 60×90/16. Гарнитура «Times New Roman».  
Усл. печ. л. 24,5. Уч.-изд. л. 21,0.

ООО «Издательство «КноРус».  
117218, г. Москва, ул. Кедрова, д. 14, корп. 2.  
Тел.: 8-495-741-46-28.  
E-mail: office@knorus.ru <http://www.knorus.ru>

Отпечатано в АО «Т8 Издательские Технологии».  
109316, г. Москва, Волгоградский проспект, д. 42, корп. 5.  
Тел.: 8-495-221-89-80.

**ISBN 978-5-406-06704-8**

© Бондаренко П.С., Горелова Г.В., Кацко И.А., 2019  
© ООО «Издательство «КноРус», 2019

## ОГЛАВЛЕНИЕ

Предисловие.....	6
Введение.....	9
<b>Часть I Теория вероятностей.....</b>	<b>12</b>
Глава 1 Случайные события.....	13
1.1 Алгебра событий.....	13
1.2 Вероятность события.....	17
1.3 Комбинаторика.....	21
1.4 Основные теоремы теории вероятностей.....	29
1.5 Формулы полной вероятности и вероятности гипотез.....	37
Глава 2 Повторные независимые испытания.....	45
2.1 Схемы повторных независимых испытаний и формула Бернулли.....	45
2.2 Приближенные формулы в схеме Бернулли.....	49
Глава 3 Дискретные случайные величины.....	59
3.1 Закон распределения дискретной случайной величины.....	59
3.2 Числовые характеристики дискретных случайных величин.....	62
3.3 Законы распределения дискретных случайных величин.....	65
3.4 Одинаково распределенные взаимно-независимые случайные величины.....	70
Глава 4 Непрерывные случайные величины.....	76
4.1 Функция распределения и ее свойства.....	76
4.2 Плотность распределения вероятностей непрерывной случайной величины.....	77
4.3 Числовые характеристики непрерывных случайных величин.....	78
Глава 5 Основные законы распределения непрерывных случайных величин.....	87
5.1 Равномерное распределение.....	87
5.2 Показательное распределение.....	89
5.3 Нормальное распределение.....	91
Глава 6 Система двух случайных величин.....	100
6.1 Понятие и закон распределения двумерной случайной величины.....	100
6.2 Функция распределения двумерной случайной величины.....	101
6.3 Плотность вероятности двумерной случайной величины.....	102
6.4 Числовые характеристики системы двух случайных величин. Коэффициент корреляции.....	104
Глава 7 Функции случайных величин.....	110
7.1 Закон распределения функции случайных величин.....	110
7.2 Композиция законов распределения.....	115
7.3 Специальные законы распределения.....	117
Глава 8 Закон больших чисел.....	122
8.1 Сущность закона больших чисел.....	122
8.2 Неравенство и теорема Чебышёва.....	124
8.3 Понятие о центральной предельной теореме.....	130
Глава 9 Цепи Маркова.....	133
Глава 10 Приложения теории вероятностей в компьютерных науках ( <i>computer science</i> ).....	138
10.1 Производящие функции.....	139
10.2 Вероятностный анализ скорости выполнения алгоритмов.....	143
10.3 Случайные числа, генераторы случайных чисел.....	151
10.4 Вероятностный подход к понятию информации.....	155

	10.5 Байесовские сети.....	158
<b>Часть II</b>	<b>Математическая статистика.....</b>	<b>170</b>
Глава 11	Вариационные ряды распределения.....	173
	11.1 Построение и графическое изображение вариационных рядов... ..	173
	11.2 Меры центральной тенденции.....	178
	11.3 Показатели вариации.....	181
Глава 12	Выборочный метод.....	190
	12.1 Понятие о выборочном методе.....	190
	12.2 Статистические оценки параметров генеральной совокупности..	193
	12.3 Оценка генеральной средней и дисперсии по выборочной средней и дисперсии.....	194
	12.4 Доверительные интервалы характеристик генеральной совокупности.....	196
Глава 13	Проверка статистических гипотез.....	204
	13.1 Понятие и виды статистических гипотез.....	204
	13.2 Проверка гипотезы о среднем значении нормально распределенной генеральной совокупности.....	208
	13.3 Проверка гипотезы о среднем значении генеральной доли.....	211
	13.4 Проверка гипотезы о дисперсиях нормально распределенных генеральных совокупностей.....	213
	13.5 Проверка гипотезы о равенстве двух средних независимых нормально распределенных генеральных совокупностей.....	214
	13.6 Проверка гипотезы о значимости средней разности двух зави- симых нормально распределенных генеральных совокупностей.....	217
	13.7 Проверка гипотезы о равенстве долей двух независимых нормально распределенных генеральных совокупностей.....	219
	13.8 Проверка гипотезы о виде распределения.....	220
Глава 14	Дисперсионный анализ.....	228
	14.1 Постановка задачи и сущность дисперсионного анализа.....	228
	14.2 Модели однофакторного и многофакторного дисперсионного анализа.....	232
	14.3 Примеры применения дисперсионного анализа.....	241
Глава 15	Корреляционно-регрессионный анализ.....	255
	15.1 Виды и формы связей между признаками.....	255
	15.2 Корреляционный анализ.....	260
	15.3 Однофакторный регрессионный анализ.....	269
	15.4 Множественный регрессионный анализ.....	290
Глава 16	Анализ временных рядов.....	307
<b>Часть III</b>	<b>Введение в анализ данных.....</b>	<b>317</b>
Глава 17	Анализ данных: современное состояние, перспективы.....	320
	17.1 Введение в методы анализа данных.....	320
	17.2 Системный подход как идеология анализа данных.....	327
	17.3 Представление данных.....	336
	17.4 Элементы анализа данных на современном этапе.....	341
	17.5 Анализ данных в контексте процесса формирования знаний.....	350
<b>Заключение</b>	.....	<b>364</b>
<b>Литература</b>	.....	<b>366</b>
<b>Ответы</b>	.....	<b>371</b>
<b>Приложения</b>	.....	<b>377</b>

*Замечательно, что науке, начинавшейся с рассмотрения азартных игр, суждено было стать важнейшим объектом человеческого знания...*  
*П. Лаплас*  
*Аналитическая теория вероятностей, 1812*

*При изучении наук примеры полезнее правил.*  
*И. Ньютон*

*... истинной логикой этого мира является исчисление вероятностей, занимающееся нахождением величин вероятностей, которые учитывает или должен учитывать любой здравомыслящий человек.*  
*Дж. Максвелл*

*Свои способности человек может узнать, только попытавшись приложить их.*  
*Сенека*

*Есть люди, полагающие, что математика – это нудное занятие, которое всегда уныло и скучно; мы же находим математику увлечением и не стыдимся признаться в этом.*  
*Р. Грэхем, Д. Кнут, О. Паташник.*  
*Конкретная математика. Основания информатики, 1998*

*На книжной полке рядом стоят два тома Пушкина: первый и второй. Страницы каждого тома имеют вместе толщину 2 см, а обложка каждая 2мм. Червь прогрыз (перпендикулярно страницам) от первой страницы первого тома до последней страницы второго тома. Какой путь он прогрыз? [Эта топологическая задача с невероятным ответом – 4 мм – совершенно недоступна академикам, но некоторые дошкольники легко справляются с ней.]*  
*В. И. Арнольд. Задачи для детей от 5 до 15, 2004*

*Конечно, мы будем учиться доказывать, но будем также учиться догадываться.*  
*Д. Пойа*

## Предисловие

На современном этапе развития Российского общества, с переходом к рыночным отношениям, резко повысилась управленческая роль руководителя организации. В связи с этим в нашей стране проводятся многочисленные исследования, перенимается и пропагандируется опыт зарубежных стран в области менеджмента и маркетинга. Одним из важнейших моментов в деятельности руководителя, менеджера, экономиста является принятие решений в условиях неопределённости. При этом наиболее разработанным инструментарием является математическая статистика, позволяющая решать *задачи принятия решений<sup>1</sup> в условиях вероятностной неопределённости* и имеющая достаточно развитое программное обеспечение (например, в *Excel*).

В процессе всей своей жизни человек часто сталкивается с событиями и явлениями, исход которых заранее не определен. Например, студент не знает, какие именно вопросы задаст экзаменатор, служащий – сколько времени у него займет дорога на работу завтра (через неделю), инвестор – окупятся ли его инвестиции, страховщик – причину и размер выплаты страхового вознаграждения и т. д. Тем не менее, в подобных ситуациях, связанных с неопределенностью, человеку необходимо принимать решения.

Обычно принятию решений предшествует анализ известных данных (на основании предшествующего опыта, здравого смысла, интуиции и т.д.). Первые известные примеры обработки данных описаны в Ветхом Завете (Книга Чисел). Стремясь увидеть и обосновать закономерности в неопределенных процессах, человечество выработало целый арсенал методов, совокупность которых называется математической статистикой (прикладной статистикой или анализом данных).

Кратко рассмотрим основные направления методов математической статистики.

Выборочное наблюдение – решает задачу обобщения на всю совокупность результатов, полученных при изучении ее части, например, рейтинг политиков, анкетирование, качество продукции и т. д.

Проверка статистических гипотез – позволяет ответить на вопрос о достоверности принимаемого решения (например, обоснованность рейтинга популярности).

Дисперсионный анализ – изучает влияние факторных признаков на результативный (например, зависит ли производительность труда рабочего от стажа, возраста, стажа и возраста).

Корреляционно-регрессионный анализ – позволяет выявить связи и построить модели зависимости (например, какая зависимость существует между расходами населения и сбережениями).

---

<sup>1</sup> О теории принятия решений см., например, учебник академика РАН О. И. Ларичева Теория и методы принятия решений, а также Хроника событий в Волшебных Странах. – М.: –Логос, 2000. – 296с. :ил.

Анализ временных рядов рассматривает последовательности чисел во времени и изучает их свойства (например, количество пятен на Солнце, характеризующее его активность по годам; курс доллара по дням и т. д.).

Последовательный анализ А. Вальда – один из первых разработанных методов, дающий возможность осуществлять контроль качества выпускаемой продукции.

Перечисленные выше (и другие) методы основываются на теоретических положениях теории вероятностей и являются классическими.

Существует ряд направлений, которые не используют предпосылки теории вероятностей. Наиболее известные из них указаны ниже.

Непараметрическая статистика, в отличие от классической (параметрической), не предполагает, что наблюдения подчиняются определенному закону распределения.

Бутстрап – способ обработки выборочных данных с помощью метода статистических испытаний (метода Монте-Карло), при котором выборку «размножают» и изучают устойчивость получаемых выводов.

Реальные данные – это не числа, а интервалы (результат измерения плюс – минус погрешность), поэтому возникает необходимость в соответствующей статистике.

Статистика объектов нечисловой природы. Например: качественные оценки признака (машина – очень плохая, плохая, хорошая, очень хорошая), ранжировка (распределение студентов по росту), классификация (группировка студентов курса не по одному, а по ряду признаков, например: интересы, достаток, оценки, симпатии, антипатии) и т.д.

Методы прикладной статистики быстро входят в нашу жизнь посредством пакетов прикладных математических (*Matcad, MatLab*), статистических (*Statistica, STADIA, SAS, IBM SPSS, AtteStat*) и других программ, в которых предусматриваются средства обработки данных. Здесь даже у серьезных людей возникает вопрос «А зачем знать теоретические положения? Главное уметь кнопки нажимать!...». Ответом на этот вопрос может служить притча, приводимая Ф. Фишером.

В *N*-ской губернии разразилась эпидемия чумы. Крестьяне узнали, что можно «научно» выяснить причину болезни, для этого нужно посчитать коэффициент корреляции (глава 15). Таким образом, они выяснили, что между количеством заболеваний в деревнях и количеством врачей существует прямая корреляционная зависимость. Поэтому крестьяне решили избавиться от врачей, посчитав их причиной болезни...

Притча, скорее всего, вызывает улыбку, однако, в менее очевидном случае мы зачастую поступаем аналогично.

Поэтому, необходимо знать основные идеи и методы прикладной статистики, условия их применения, а затем «нажимать кнопки». Методически более целесообразно изучать анализ данных на компьютере в *Excel*, а затем, по мере возникновения соответствующих вопросов, переходить к профессиональным программам.

В тексте пособия кроме традиционных вопросов рассмотрены дополнительные разделы (гл. 10, часть III), а также вопросы, связанные с излагаемым материалом и выделенные в «замечания».

Первая часть пособия посвящена теории вероятностей, в которой изложены традиционные разделы: случайные события, случайные величины, основные законы распределения случайных величин, системы случайных величин, функции случайных величин, закон больших чисел. Для студентов, обучающихся по направлениям подготовки, связанным с ИТ предлагается глава 10 «Приложения теории вероятностей в компьютерных науках (*computer science*)».

Вторая часть пособия посвящена математической статистике. Рассмотрены следующие разделы: вариационные ряды, выборочный метод, проверка статистических гипотез, дисперсионный и корреляционно-регрессионный анализ, анализ временных рядов.

В третьей части изложен обзор современных направлений анализа данных. Она имеет цель показать преемственность курса теории вероятностей и математической статистики в современном контексте развития общества и информационных технологий как одной из ключевых составляющих современного понимания аналитики, знания возможностей которой совершенно необходимы для подготовки специалистов высокого уровня.

Авторы выражают благодарность:

- уважаемым рецензентам Л. И. Ниворожкиной, С. Г. Чефранову;
- нашим коллегам В. Н. Волковой, Ю. И. Бершицкому, В. Н. Лаптеву, Е. В. Луценко, Ю. И. Лыпарю, Н. Н. Лябах, А. М. Ляховецкому, А. И. Орлову, Н. Б. Паклину, а также всем участникам ежегодной международной научно-практической конференции «Системный анализ в проектировании и управлении» (СпбГПУ, рук. В. Н. Волкова, В. Н. Козлов, В. Е. Ланкин) многолетнее общение с которыми способствовало формированию взглядов, отраженных в настоящем пособии;

- членам «Российской ассоциации статистиков» (рук. И. И. Елисеева, А. Н. Пономоренко) К. А. Акопян, А. Я. Бурдяк, В. В. Глинскому, В. В. Демичеву, Е. С. Завариной, А. П. Зинченко, Т. Н. Лариной, И. П. Мамий, В. С. Мхитарян, В. Н. Салину, А. Н. Уколовой, А. В. Шаль, А. Е. Шибалкину и многим другим, принявшим очное и заочное участие в I Открытом российском статистическом конгрессе (Новосибирск 20-22 октября 2015 г.) общение, с которыми было очень продуктивно и полезно;

- коллегам по кафедре статистики и прикладной математики и студентам, принявшим участие в подготовке настоящего пособия к изданию: Н. Х. Вороковой (тесты); Т. В. Соловьевой (гл. 9, тесты); О.И. Швыревой, М. О. Киек (10.6); П. Ю. Величко, С. Г. Павлову (17.4), Ю. С. Михалевич, Е. В. Шкуропат (17.5).

Авторы будут признательны за замечания и пожелания, которые можно отправлять по адресу: stat @ kubsau.ru (для И. А. Кацко).



## Введение

Теория вероятностей – это математическая дисциплина, изучающая закономерности, происходящие в массовых однородных случайных явлениях и процессах.

С возникновением теории вероятностей наука получила мощный аппарат исследования случайных явлений и процессов, до этого исследовались лишь детерминированные явления и опыты, в которых первоначальные условия однозначно позволяли определить исход. Между тем случайные явления присутствуют во многих областях науки (биологии, генетике, агрономии, экономике, демографии, технике и т.д.), когда заранее невозможно предсказать результат опыта.

Исторически зарождение и развитие теории вероятностей связано с азартными играми, в которых требовалось обосновать то или иное решение. Классическим примером является задача рассматриваемая ниже.

*Задача.* Двое играют в безобидную игру (шансы выиграть у обоих одинаковы). Договариваются, что всю ставку забирает игрок, выигравший первым 6 партий. Как правильно разделить ставку, если игра остановилась при счете 5:3?

*Решение.* Для выигрыша первому игроку достаточно выиграть одну партию, второму игроку необходимо подряд выиграть три партии. Всего три партии предполагают  $2 \cdot 2 \cdot 2 = 8$  исходов (каждая партия имеет два исхода: выиграл, проиграл). В пользу второго игрока только один исход из 8 возможных, а в пользу первого – 7 исходов. Поэтому справедливо разделить ставку пропорционально шансам выиграть, то есть 7:1.

Решение подобных задач Б. Паскалем, П. Ферма, Х. Гюйгенсом послужило появлению теории вероятностей. Дальнейшее ее развитие связано с именами Я. Бернулли, С. Пуассона, А. Муавра, П. Лапласа, К. Гаусса, П. Чебышева, А. Маркова, А. Ляпунова, А. Хинчина, А. Колмогорова и др.

Как и всякая математическая теория, с точки зрения аксиоматического подхода, теория вероятностей занимается изучением соотношений между неопределяемыми объектами (понятиями). В геометрии неопределяемые понятия это точка, прямая и плоскость, аналогично, в теории вероятностей неопределяемые понятия это элементарные события (исходы) ( $\omega_i$ ) и пространство элементарных событий ( $\Omega = \{\omega_i\}$ ).

Пример 1. Монета подбрасывается один раз. Возможные элементарные исходы: выпал герб –  $\omega_1$ , выпала решка –  $\omega_2$ . Пространство элементарных событий  $\Omega = \{\omega_1, \omega_2\}$ .

Пример 2. Игральная кость подбрасывается один раз. Элементарные события:  $\omega_1$  – появление 1,  $\omega_2$  – 2,  $\omega_3$  – 3,  $\omega_4$  – 4,  $\omega_5$  – 5,  $\omega_6$  – 6. Пространство элементарных событий  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$ .

Вероятность события – это число, имеющее ту же природу, что и расстояние в геометрии или масса в теоретической механике. Она всегда связана с каким-либо пространством элементарных событий, природа которого не имеет значения. Понятие вероятности обычно строится на интуитивных соображениях (например, вероятность появления герба при подбрасывании симметричной монеты очевидно равна  $1/2$ ) и связано с статистической устойчивостью относительной частоты события при большом числе опытов (относительная частота события  $\mu(A) = m/n$ , где  $m$  – число появлений события  $A$  в серии из  $n$  опытов). При подбрасывании монеты достаточно большое число раз относительная частота появлений герба будет колебаться около  $0,5$ , следовательно, можно говорить, что вероятность появления герба равна  $0,5$  (табл.1). Устойчивость относительной частоты появления события позволяет судить о вероятности, как об объективной характеристике события в данном опыте, имеющей вполне определенное значение, независимо от того, будут проводиться опыты или нет. (То есть имеет место вероятностная неопределённость.)

Предметом современной теории вероятностей является выявление общих закономерностей и зависимостей изучаемых явлений, а также изучение и описание различных явлений (физических, экономических, технологических и др.) с помощью абстрактных моделей.

Математическая статистика – это раздел математики, в котором изучаются математические методы сбора, систематизации, обработки и использования статистических данных для научных и практических выводов.

Математическая статистика использует математический аппарат и выводы теории вероятностей. Связующим звеном между теорией вероятностей и математической статистикой является закон больших чисел и так называемые предельные теоремы. В частности, закон больших чисел аргументирует применение средней арифметической в качестве оценки математического ожидания, относительной частоты появления события как оценки вероятности. Последнее обосновывает понятие статистической устойчивости.

Всю жизнь человек вынужден принимать решения: в личной сфере (в какой вуз поступать, с кем общаться, как учиться); в общественной (посещать вечера, театры, митинги, собрания, выборы); в производственной (определение факторов, существенно влияющих на урожайность, производительность труда качество материалов и т.д.); научной (выдвижение и проверка научных гипотез).

Принятие решений обычно преследует одну из целей: прогнозирование будущего состояния процесса (объекта); управление (т.е. как следует изменять одни параметры объекта (процесса), чтобы другие параметры приняли желаемое значение); объяснение внутренней структуры объекта (процесса).

Одним из основных подходов к обоснованию и последующему принятию решений является статистический.

Статистические методы обработки данных можно классифицировать по нижеследующим признакам.

По способу получения экспериментальных данных:

- активный эксперимент;
- пассивный эксперимент (выборочное или сплошное наблюдение).

По цели обработки данных:

- 1) описательные (получение и сравнение числовых характеристик экспериментальных данных) – анализ вариационных рядов, выборочный метод, проверка статистических гипотез и другие;
- 2) аналитические (количественная оценка и анализ зависимостей, описывающих изучаемые объекты (процессы) – дисперсионный анализ, регрессионный анализ, анализ рядов динамики и другие).

В математической статистике предполагается, что результаты опытов и наблюдений являются реализацией различных случайных процессов, имеющих те или иные законы распределения (причем неизвестные заранее), а иногда и детерминированные составляющие (регрессионный анализ). Отсюда вытекают основные задачи математической статистики:

- 1) организация наблюдений;
- 2) нахождение по результатам выборочных наблюдений оценок числовых характеристик всей совокупности и исследование точности их приближения (выборочный метод);
- 3) решение вопроса согласования результатов оценивания с опытными данными (проверка статистических гипотез);
- 4) оценка существенности влияния факторных признаков на результативный (дисперсионный анализ);
- 5) выявление аналитической зависимости между признаками (корреляционно-регрессионный анализ).

А. Вальд говорил, что «математическая статистика – это теория принятия решений в условиях неопределенности».

По существу, математическая статистика дает единственный, математически обоснованный, аппарат для решения задач управления и прогнозирования при отсутствии явных закономерностей (наличии случайностей) в изучаемых процессах.

# *Часть I*

## *Теория вероятностей*

# Глава 1 Случайные события

## 1.1 Алгебра событий

Одним из основных понятий теории вероятностей является опыт. Под опытом понимается выполнение комплекса условий, в результате которого происходят или не происходят определенные события (факты).

Простейшие неразложимые результаты опыта называются элементарными событиями ( $\omega_i$ ), а совокупность элементарных событий называется пространством элементарных событий  $\Omega = \{\omega_i\}$ . С каждым опытом связано пространство элементарных событий  $\Omega$  (введение, примеры 1, 2).

Любое конечное или счетное<sup>2</sup> подмножество  $\Omega$  называется событием. Различают три типа событий:

- 1) достоверные ( $\Omega$ ),
- 2) случайные,
- 3) невозможные ( $\emptyset$  или  $\bar{\Omega}$ ).

*Достоверным* называется событие, которое обязательно произойдет в данном опыте при выполнении комплекса условий. Например, при подбрасывании игральной кости выпадет не более шести очков.

*Невозможным* называется событие, которое в данном опыте не может произойти. Например, при подбрасывании игральной кости выпадет более шести очков.

*Случайным* называется событие, которое в данном опыте может либо произойти, либо не произойти. Например, выпадение шести очков при подбрасывании игральной кости.

События обычно обозначают первыми прописными буквами латинского алфавита:  $A, B, C, \dots$ . Например, в примере введения 2 событие  $A = \{\omega_2, \omega_4, \omega_6\}$  – появление четного числа при подбрасывании игральной кости.

События  $A$  и  $B$  несовместны, если в результате одного опыта они не могут происходить вместе, в противном случае – совместны. Например, при одном подбрасывании монеты не могут одновременно

---

<sup>2</sup>Счетным называется множество, элементам которого можно поставить в соответствие ряд натуральных чисел, например, последовательность  $\left\{\frac{1}{n}\right\}$ , где  $n \in N$  – счетное множество, так как можно установить соответствие между элементами последовательности и множеством натуральных чисел:

$$\begin{array}{cccccccc} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \dots & \frac{1}{n} & \dots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \dots & \updownarrow & \dots \\ 1 & 2 & 3 & 4 & 5 & \dots & n & \dots \end{array}$$

появиться герб и решка. Элементы последовательности событий  $A_1, A_2, \dots, A_n$  попарно несовместны, если любые два из них несовместны. Например, при подбрасывании игральной кости никакие два элементарных исхода (появление цифр 1, 2, 3, 4, 5, 6) не могут произойти одновременно.

Несколько событий равновозможны, если ни одно из них не имеет объективного преимущества перед другими. Например, элементарные исходы при подбрасывании монеты, игральной кости. Несовместные события  $A_1, A_2, \dots, A_n$  образуют полную группу, если в результате опыта кроме этих событий ничего не может произойти. Два события, образующие полную группу, называются противоположными. Например, попадание и промах по мишени при выстреле.

Обычно  $\Omega$  изображают на плоскости в виде некоторой области, а  $\omega_i$  в виде точек этой области, устанавливая, таким образом, соответствие между событиями и точечными множествами. Над событиями вводятся операции, совпадающие с операциями над множествами: сумма, произведение, отрицание.

1. Суммой событий  $A$  и  $B$  называется такое третье событие  $A + B$  (или  $A \cup B$ ), которое заключается в наступлении хотя бы одного из событий, т. е. или  $A$ , или  $B$ , или в совместном их появлении  $AB$ . Суммой двух или нескольких несовместных событий называется событие, заключающееся в появлении только одного из них (рисунок 1.1).

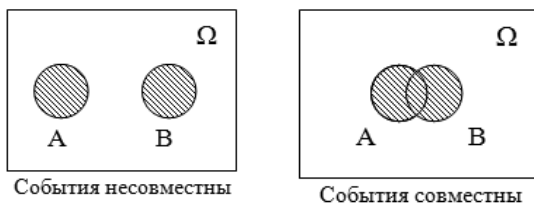


Рисунок 1.1 – Сумма событий

2. Произведением двух событий  $A$  и  $B$  называется такое третье событие  $AB$  (или  $A \cap B$ ), которое заключается в наступлении событий  $A$  и  $B$  вместе. Если события  $A$  и  $B$  несовместны, то  $AB = \emptyset$  (рисунок 1.2). Произведением нескольких событий называется событие, заключающееся в совместном появлении этих событий.

3. Отрицанием события  $A$  называется событие  $\bar{A}$  (не  $A$ ), заключающееся в не наступлении события  $A$  ( $A + \bar{A} = \Omega, A\bar{A} = \emptyset$ ). Причем, если в результате опыта может произойти событие  $A$ , то может произойти и обратное ему событие  $\bar{A}$  (рисунок 1.3).

Если наступление события  $A$  приводит к наступлению события  $B$  и наоборот (наступление  $B$  влечет наступление  $A$ ), то события  $A$  и  $B$  равны ( $A = B$ ).

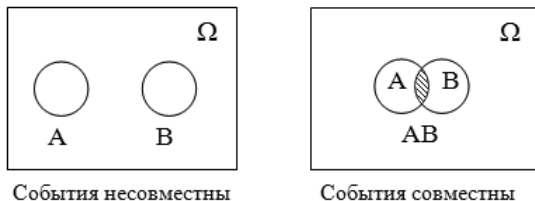


Рисунок 1.2 – Произведение событий

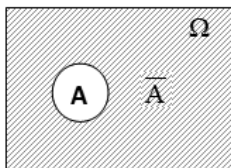


Рисунок 1.3 –  $\bar{A}$ – отрицание события  $A$

Пусть  $S$  – множество всех подмножеств  $\Omega$ , для которого выполняются следующие свойства:

- 1) если  $A \in S$  и  $B \in S$ , то  $A + B = A \cup B \in S$ ,
- 2) если  $A \in S$  и  $B \in S$ , то  $AB = A \cap B \in S$ ,
- 3) если  $A \in S$ , то  $\bar{A} \in S$ ,

тогда множество  $S$  называется алгеброй событий.

**Замечание.**

1. При более точном подходе достаточно одного из свойств 1) или 2), так как одно из них следует из другого.

2. При расширении операций сложения и умножения на случай счетного множества событий, алгебра событий  $S$  называется  $\sigma$ -алгеброй или борелевской алгеброй:

если  $A_i \in S, i \in N$ , то  $\sum_{i=1}^{\infty} A_i \in S, \prod_{i=1}^{\infty} A_i \in S$ .

**Пример 1.1.** Стрелок произвел 3 выстрела по мишени, элементарные события:

$A_1$  – попал при 1-ом выстреле;  $\bar{A}_1$  – не попал при 1-ом, выстреле;

$A_2$  – попал при 2-ом выстреле;  $\bar{A}_2$  – не попал при 2-ом, выстреле;

$A_3$  – попал при 3-ем выстреле;  $\bar{A}_3$  – не попал при 3-ем, выстреле.

Выразим через  $A_1, A_2, A_3$  и их отрицания следующие события:

а) одно попадание:  $A = A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3$ ,

б) три промаха:  $B = \bar{A}_1\bar{A}_2\bar{A}_3$ ,

в) три попадания:  $C = A_1A_2A_3$ ,

г) хотя бы один промах:  $D = A + B + A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3$ ,  
или  $D = \bar{A}_1 + \bar{A}_2 + \bar{A}_3$ ,

д) не менее двух попаданий:  $E = A_1A_2\bar{A}_3 + A_1\bar{A}_2A_3 + \bar{A}_1A_2A_3 + C$ ,

- е) не более одного попадания:  $F = A + B$ ,  
 ж) попадание в мишень после первого выстрела:  
 $G = \bar{A}_1(\bar{A}_2A_3 + A_2\bar{A}_3 + A_2A_3)$ .

### Контрольные задания

- Являются ли несовместными следующие события:
  - опыт – бросание двух монет. События:  
 $A_1$  – появление двух гербов,  $A_2$  – появление двух цифр;
  - опыт – три выстрела по мишени. События:  
 $B_1$  – хотя бы одно попадание,  $B_2$  – хотя бы один промах;
  - опыт – бросание двух игральных костей. События:  
 $C_1$  – хотя бы на одной кости появилось три очка,  
 $C_2$  – появление четного числа очков на каждой кости;
  - опыт – извлечение двух шаров из урны, содержащей белые и черные шары. События:  
 $D_1$  – взято два белых шара,  $D_2$  – оба извлеченных шара одного цвета;
  - опыт – покупка двух лотерейных билетов. События:  
 $E_1$  – выиграно два билета,  $E_2$  – выигран хотя бы один билет,  
 $E_3$  – выигран только один лотерейный билет;
  - опыт – лифт отправляется с 10 пассажирами и останавливается на пяти этажах. События:  
 $F_1$  – на первых четырех остановках вышло не более 9 человек,  
 $F_2$  – на последней остановке вышел хотя бы один человек.
- Образуют ли полную группу следующие события:
  - опыт – два выстрела по мишени. События:  
 $A_1$  – два попадания в мишень,  $A_2$  – хотя бы один промах по мишени;
  - опыт – бросание двух игральных костей. События:  
 $B_1$  – сумма очков на верхних гранях больше 3,  
 $B_2$  – сумма очков на верхних гранях равна 3;
  - опыт – посажено четыре зерна. События:  
 $C_1$  – взошло одно зерно,  $C_2$  – взошло два зерна,  
 $C_3$  – взошло три зерна,  $C_4$  – взошло четыре зерна.
  - покупатель посещает три магазина. События:  
 $D_1$  – покупатель купит товар хотя бы в одном магазине,  
 $D_2$  – покупатель не купит товар ни в одном магазине.
- Являются ли равновероятными следующие события:
  - опыт – выстрел по мишени. События:  
 $A_1$  – попадание при выстреле,  $A_2$  – промах при выстреле;
  - опыт – бросание двух игральных костей. События:  
 $B_1$  – произведение очков на верхних гранях равно 12,  
 $B_2$  – сумма очков на верхних гранях равна 9;
  - опыт – бросание двух монет. События:  
 $C_1$  – появление двух гербов,  $C_2$  – появление двух цифр,  
 $C_3$  – появление одного герба и одной цифры.

**Замечание.** В данном случае (пункт в)) возможно несколько подходов, связанных с различными статистическими моделями физических частиц.

- 1) Модель (статистика) Больцмана-Максвелла (частицы различимы, хотя известно, что таких частиц в природе не существует). Для опыта с монетами:  $\Gamma_1\Gamma_2, \Gamma_1P_2, \Gamma_2P_1, P_1P_2$  (монеты различимы).



## Заключение

*Ещё многое имею сказать вам,  
но вы теперь не можете вместить  
Евангелие от Иоанна , 16:12*

*Тебе бы опыт сделать не мешало –  
Ведь он для вас источник всех наук...*  
Гете

В настоящем пособии рассмотрено введение в ряд классических разделов теории вероятностей и математической статистики, основывающейся на предпосылках нормальности распределения изучаемых случайных величин, которые стали основой развития методов анализа данных.

Успешная практическая деятельность человека все в большей мере зависит от организации сбора и обработки информации. Совершенствование технологий записи и хранения данных – создание баз данных и знаний во всех сферах деятельности человека предъявляет новые требования к уровню подготовки специалистов. Большой объем информации, которой сопровождается деятельность практически любого предприятия и учреждения, обычно содержит полезные сведения, благодаря которым, можно значительно повысить эффективность работы, совершенствуя технологию, управление и т.д. И завтра сегодняшним студентам новые идеи анализа данных будут очевидны и необходимы на практике.

Современные коммерческие организации интенсивно внедряют информационные хранилища (банки) данных и знаний, многие из которых содержат средства интеллектуального анализа данных или предполагают возможность их применения. Деловые люди осознали, что применение методов *анализа данных*, использующих вероятностную и геометрическую природу данных, может дать ощутимое преимущество в конкурентной борьбе.

Более подробно ознакомиться с современными методами прикладной статистики и *интеллектуального анализа данных* можно, например, по электронным учебникам фирмы *StatSoft*, которые в основном соответствуют описанию статистических модулей пакета *STATISTICA*, а также на сайте фирмы *BaseGroup Labs*, создателя аналитической платформы *Deductor* – флагмана российской бизнес-аналитики. Эти направления мы предлагаем для дальнейшего самостоятельного изучения и применения на практике для анализа структурированных данных, с помощью прикладных пакетов статистических, *DATA MINING* и *KDD* (*knowledge discovery in databases*) и других программ, позволяющих анализировать слабоструктурированные и неструктурированные данные

Следует отметить, что на современном этапе управление производством, фирмой, районом, регионом практически невозможно без си-

темного подхода [51-53], разрабатывающего методики анализа целей, методы и модели совершенствования организационной структуры, управления функционированием социально-экономических объектов.

В зависимости от априорной информации об изучаемом объекте применяют следующие методы: мозговой атаки, построения сценариев, экспертной оценки, построения дерева целей, математической логики, теории множеств, теории игр, прикладной статистики, математического программирования, интеллектуального анализа данных и т.д. Разумеется, большинство методов пересекается. При этом рассматриваемые в книге вероятностные методы в рамках системного анализа являются одним из возможных подходов перевода вербального (словесного) описания модели изучаемого объекта в формальное, для решения задач управления и принятия решений.

Сегодня современные специалисты развивают системный подход в рамках необходимости изучения кроме механических и биологических систем мультиразумных (социальных), которые характеризуются взаимозависимыми переменными. Этот подход получил название *системного мышления*.

Безусловно, не существует абсолютных рецептов построения моделей сложных процессов. Моделирование в большей степени искусство, овладеть которым можно только, решая практические задачи. Целью настоящего изложения является не сама истина, которая со временем обычно изменяется, а возможный путь её поиска, двигаясь по которому можно приобрести неоценимый опыт.

*Природа – древний храм, где строй живых колонн  
Обрывки смутных фраз роняет временами.  
Мы входим в этот храм в смятении, а за нами  
Лес символов немых следит со всех сторон.*

Бодлер

## Литература

### Часть I, II

1. Айвазян С.А. Прикладная статистика. Основы моделирования и первичная обработка данных / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – Финансы и статистика. – 1985. – 472 с.
2. Айвазян С.А. Прикладная статистика. Исследование зависимостей / С.А. Айвазян, И.С. Енюков, Л.Д. Мешалкин. – Финансы и статистика. – 1985. – 488 с.
3. Айвазян С.А. Прикладная статистика. Классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков и др. – Финансы и статистика. – 1989. – 607 с.: ил.
4. Андрухаев Х.М. Сборник задач по теории вероятностей / Х.М. Андрухаев. – Просвещение. – 1985. – 160 с.
5. Бондаренко П.С. Теория вероятностей и математическая статистика: учеб. пособие для бакалавров / П.С. Бондаренко, Г.В. Горелова, И.А. Кацко. – Кубанский ГАУ. – 2013. – 340 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете)
6. Бондаренко П.С. Практикум по эконометрике: учеб.-практ. пособие для бакалавров / П.С. Бондаренко и др. – Кубанский ГАУ. – 2013. – 164 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете)
7. Венецкий Н.Г. Теория вероятностей и математическая статистика изд. 3/ Н.Г. Венецкий, Г.Е. Кильдишев. – «Статистика», 1975. – 264 с.
8. Вентцель Е.С. Теория вероятностей: Учеб. для вузов. 5-е изд. стер./ Е.С. Вентцель. – Высш. шк., 1998. – 576 с., ил.
9. Вентцель Е.С. Теория вероятностей и её инженерные приложения/ Е.С. Вентцель, Л.А. Овчаров. – Высшая школа. – 2000. – 480 с.
10. Вентцель Е.С. Задачи и упражнения по теории вероятностей: Учеб. пособие для вузов изд. 3 / Е.С. Вентцель, Л.А. Овчаров. – Высш.шк. – 2000. – 366с.: ил.
11. Герасимович А.И. Математическая статистика / А.И. Герасимович, Я.И. Матвеева. – Высш. шк.- 1978. – 200 с.
12. Гмурман В.Е. Теория вероятностей и математическая статистика. Изд. 6 - е, стер./ В.Е. Гмурман. – Юрайт - 2016. – 480 с.
13. Гмурман Д.Е. Руководство к решению задач по теории вероятностей и математической статистике: Учебное пособие – 4-е изд., стереотип / В.Е. Гмурман. – Юрайт, 2016. – 404 с.
14. Гнеденко Б.В. Курс теории вероятностей: Учебник – Изд. 6-е, перераб. и доп. – Б.В. Гнеденко // Наука. – 1988. – 448 с.
15. Горелова Г.В. Теория вероятностей и математическая статистика в примерах и задачах с применением EXCEL. Учеб. пособие для вузов. Издание 4-е испр. и доп./ Г.В. Горелова, И.А. Кацко. – Феникс. – 2006. – 475с.:ил.
16. Грэхэм Р. Конкретная математика. Основания информатики: Пер с англ./ Р. Грэхэм, Д. Кнут, О. Паташник. – Мир, 1998. – 703 с., ил.
17. Демиденко Е. З. Линейная и нелинейная регрессия/ Е. З. Демиденко. – Финансы и статистика. – 1981. – 302 с., ил.
18. Дрейпер И. Прикладной регрессионный анализ: В 2-х кн. Пер. с англ. – 2-е изд. перераб. и доп. / И. Дрейпер, Г. Смит. – Финансы и статистика. – 1986 Кн. 1. – 366 с., ил; 1987. Кн.2. – 351 с., ил.
19. Езикал М., Фокс К. Методы анализа корреляций и регрессий: линейных и криволинейных. – М.: 1966. – 557с. (с.211).

20. Елисеева И.И. Теория статистики с основами теории вероятностей: Учеб. пособие для вузов / И.И. Елисеева, В.С. Князевский, Л.И. Ниворожкина, З.А. Морозова; Под ред. И.И. Елисеевой. – ЮНИТИ – ДАНА. – 2001. – 446 с.
21. Закс Л. Статистическое оценивание. Пер. с нем. В.Н. Варыгина. / Л. Закс, под ред. Ю.П. Адлера, В. Г. Горского. – М., «Статистика», 1976. – 598 с., ил.
22. Йейтс Ф. Выборочный метод в переписях и обследованиях. Пер. с англ. Е.И. Арона./ Ф. Йетс, под ред. А.Г. Волкова. – «Статистика». – 1965. – 434 с.
23. Кацко И.А. Моделирование современных тенденций риск-ориентированного аудита с применением теоремы Байеса / И.А. Кацко, О.И. Швырева, М.О. Киек // Научный журнал КубГАУ [Электронный ресурс]. – Краснодар: КубГАУ, 2016. – №115(01). – Режим доступа: <http://ej.kubagro.ru>.
24. Кендэл М. Временные ряды. Пер. с англ. Ю.П. Лукашина./ М. Кендэл. – Финансы и статистика. – 1981. – 199 с., ил.
25. Кокрен У. Методы выборочного исследования. Пер. с англ. И.М. Сонина. / У. Кокрен, под ред. А.Г. Волкова. – «Статистика». – 1976. – 440 с.
26. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – 2-е изд., перераб. и доп. – М.: ЮНИТИ – ДАНА, 2004. – 573 с.
27. Кнут Д. Искусство программирования, том 1. Основные алгоритмы, 3-е изд.: Пер. с англ./ Д.Кнут. – Вильямс. – 2001. – 720 с.
28. Митропольский А.К. Техника статистических вычислений / А.К. Митропольский. – Наука. – 1971., 576с.:ил.
29. Ниворожкина Л.И. Математическая статистика с элементами теории вероятностей в задачах с решениями: Учебное пособие для бакалавров / Л.И. Ниворожкина, З.А. Морозова, И.Э. Гурьянова; под ред. проф. Л.И. Ниворожкиной. – 2-е изд., перераб. и доп. – М.: Издательско-торговая корпорация «Дашков и К», 2015. – 480 с.
30. Пугачев В.С. Теории вероятностей и математическая статистика: Учеб. Пособие. – 2-е изд., исправл. и допол./В.С. Пугачев. – ФИЗМАТЛИТ, 2002.–496 с.
31. Пустыльник Е.И. Статистические методы анализа и обработки наблюдений / Е.И. Пустыльник. – Наука, 1968.-288 с.,ил.
32. Реньи А. Трилогия о математике. (Диалоги о математике. Письма о вероятности. Дневник. – Записки студента по теории информации.); Пер. с венгер. / А. Реньи. – Мир. – 1980. – 375 с.
33. Свешников А.А. Сборник задач по теории вероятностей, математической статистике и теории случайных функций/ А.А./ под редакцией А.А. Свешникова. – «Наука», Главная редакция физико-математической литературы.- 1970.- 656 с.
34. Смирнов Н.В. Курс теории вероятностей и математической статистики для технических приложений / Н.В. Смирнов, И.В. Дунин – Барковский. – Наука. – 1969. – 511 с
35. Сошникова Л.А. Многомерный статистический анализ в экономике: Учеб. пособие для вузов/ Л.А. Сошникова, В.Н. Тамашевич, Г. Уебе, М. Шеффер, под ред. проф. В.Н. Тамашевича. – ЮНИТИ – ДАНА. – 1999. – 598 с.
36. Стивенс Р. Visual Basic. Готовые алгоритмы: Пер. с англ./ Р. Стивенс. – ДМК Пресс, 2000. – 384 с.: ил. (Серия «Для программистов»).
37. Фелер В. Введение в теорию вероятностей и ее приложения. В 2-х томах. Т.1: Пер. с англ./ В. Фелер // Мир, 1984. – 528 с., ил.
38. Хабаров С. Экспертные системы.: Уч. пос./ С. Хабаров. – Кафедра Информатики и Информационных Систем, ФЭУ, С-Пб. ЛТА.
39. Хьюстон А. Дисперсионный анализ. Пер. с англ. А.Г. Кругликова. / А. Хьюстон. – «Статистика» – 1971. – 88 с.
40. Четыркин Е.М. Вероятность и статистика / Е.М. Четыркин, И.Л. Калихман. – Финансы и статистика. – 1982. – 319 с.:ил.

41. Чистяков В.П. Курс теории вероятностей. 4-е издание / В.П. Чистяков. – Агар.-1996. – 256 с.
42. Шеффе Г. Дисперсионный анализ / Шеффе Г. // Физматгиз. – 1963. – 628 с., ил.
43. Юденков В.А. Дисперсионный анализ. Уч. пос. / В.А. Юденков. – Вышэйш. шк. 1982. – 95с.

### Часть III

44. Арнольд В.И. «Жесткие» и «мягкие» математические модели / В.И. Арнольд. – МЦНМО. 2008. – 32 с.
45. Акофф Р. От данных к мудрости. – *Journal of Applied Systems Analysis*. 1989. Vol. 18. P. 3-9. (Пер.: <http://cyberleninka.ru/article/n/>).
46. Аффифи А. Статистический анализ: Подход с использованием ЭВМ / А. Аффифи, С. Эйзен; Пер. с англ. – Мир, 1982. – 488 с., ил.
47. Борисов В.В. Компьютерная поддержка сложных организационно-технических систем / В.В. Борисов, И.А. Бычков и др. – Горячая линия – Телеком, 2002. – 154 с.: ил.
48. Боровиков В.П. *STATISTICA* – Статистический анализ и обработка данных в среде Windows. Издание 2-е, стереотипное / В.П. Боровиков, И.П. Боровиков. – Информационно-издательский дом «Филинь», 1998. – 608 с.
49. Боровиков В.П. *STATISTICA*. Искусство анализа данных на компьютере: Для профессионалов 2-е изд. / В.П. Боровиков. – Питер, 2003. – 688с.:ил.
50. Брусиловский П. Data Mining vs. Statistics – <http://www.bisolutions.us/Data-Mining-vs-Statistics.php>.
51. Волкова В.Н. Основы теории систем и системного анализа. / В. Н. Волкова, А. А. Денисов. – СПб.:Изд. СПбГТУ, 1997. – 510с. :ил.
52. Волкова В. Н. Постепенная формализация моделей принятия решений / В. Н. Волкова. – Изд-во Политехн. ун-та, 2006. – 120 с.: ил.
53. Гараедаги Дж. Системное мышление: Как управлять хаосом и сложными процессами: Платформа для моделирования архитектуры бизнеса / Дж. Гараедаги. – Гревцов Букс, 2010. – 480 с.
54. Горелова Г.В. Когнитивный анализ и моделирование устойчивого развития социально-экономических систем. / Г.В. Горелова, Е.Н. Захарова, Л.А. Гинис. . – Изд-во Рост, ун-та, 2005. – 288 с.
55. Дойч Д. Начало бесконечности. Объяснения, которые меняют мир / Д. Дойч. пер. с англ. – Альпина нон-фикшн, 2016. – 582 с.
56. Дубров А.М., Многомерные статистические методы: Учебник / А.М. Дубров, В.С. Мхитарян, Л.И. Трошин. – Финансы и статистика, 1998. – 352 с.: ил.
57. Дюк В. *Data Mining*: учебный курс / В. Дюк, А. В. Самойленко. – СПб, 2001. – 368 с.:ил.
58. Загоруйко Н.Г. Прикладные методы анализа данных и знаний /Н.Г. Загоруйко. – ИМ СО РАН, 1999. — 270 с.
59. Карр Н. Великий переход: что готовит революция облачных технологий / Н. Карр; пер. с англ. А. Баранова. – Манн, Иванов и Фарбер, 2014. – 272 с.
60. Касти Дж. Большие системы. Связность, сложность и катастрофы / Дж. Касти Пер. с англ. – 1982. – 216 с.
61. Кацко И.А. К вопросу о современных проблемах образования / И.А. Кацко, С.А. Кацко. – Системный анализ в проектировании и управлении: сб. науч. тр. XIX междунар. науч. – практ. конф. 1–3 июля 2015 года. Ч. 2. // Изд-во Политехн. ун-та, 2015. – с. 269-273.
62. Кацко И.А. Роль анализа данных формировании мировоззрения современных студентов / И.А. Кацко, Ю.С. Михалевич, Е. В. Шкурюпат – Мы продолжаем традиции российской статистики: Материалы I Открытого статистического конгресса (Новосибирск, 20-22 октября 2015 г.) // Новосибирск: НГУЭУ, 2015.

- Труды I Всероссийского статистического конгресса 2015 г., с. 388. (Том 3. Статистические измерения в социальных и гуманитарных науках. Доклады международной научно-практической конференции « Открытый российский статистический конгресс». В четырех томах.)
63. Кречетов Н. Продукты для интеллектуального анализа данных/ Н. Кречетов // Рынок программных средств, N14-15\_97, с. 32-39.
  64. Киселев М. Средства добычи знаний в бизнесе и финансах / М. Киселев, Е. Соломатин // Открытые системы, № 4, 1997, с. 41-44.
  65. Кацко И.А. Модели и методы прикладных системных исследований (практикум): (учебное пособие) / Под ред. А.И. Трубилина, И.А. Кацко. – Кубанский ГАУ. – 2014. – 449 с., ил. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете)
  66. Ляховецкий А.М. Статистика: Учебное пособие для бакалавров / А.М. Ляховецкий, Н.В. Климова, Е.В. Кремьянская, под редакцией В.И. Нечаева. – Кубанский ГАУ. – 2013. – 359 с. (Серия: Вероятность, статистика и прикладные исследования в аграрном университете)
  67. Мостеллер Ф. Анализ данных и регрессия / Ф. Мостеллер, Дж. Тьюки. – «Финансы и статистика», 1982. Вып. 1 – 317 с., вып. 2 – 239 с.
  68. Нейсбит Дж. Мегатренды / Джон Нейсбит. – АСТ, Ермак, 2003. – 380 с.
  69. Нейсбит Дж. Высокая технология, глубокая гуманность: технологии и наши поиски смысла / Джон Нейсбит. – АСТ: Транзиткнига, 2005. – 381 с.
  70. Орлов А.И. Современная прикладная статистика (обобщающая статья) / А.И. Орлов. – Заводская лаборатория. –1998. т.64. №3. – 52–60 с.
  71. Орлов А.И. Прикладная статистика / А.И. Орлов. – Издательство «Экзамен», 2004. – 656 с.
  72. Орлов А.И. Организационно-экономическое моделирование: учебник: в 3 ч. / А.И. Орлов. – М.:Изд-во МГТУ им. Н.Э. Баумана, Ч. 1: Нечисловая статистика. – 2009. – 541 с. : ил.
  73. Орлов А.И. Организационно-экономическое моделирование: учебник: в 3 ч. / А.И. Орлов. – М.: Изд-во МГТУ им. Н.Э. Баумана, Ч. 2: Экспрессные оценки. – 2011. – 486 с. : ил.
  74. Орлов А.И. Организационно-экономическое моделирование: учебник: в 3 ч. / А.И. Орлов. – М.: Изд-во МГТУ им. Н.Э. Баумана, Ч. 3: Статистические методы анализа данных. – 2012. – 623 с. : ил.
  75. Паклин Н. Б. Бизнес-аналитика: от данных к знаниям. Учебное пособие. 2-у изд., перераб. и доп. / Н. Б. Паклин, В. И. Орешков. – СПб.: Изд. Питер. – 2012. – 704 с.
  76. Покорный Ю.В. Унижение математики? / Ю.В. Покорный. – Воронежский государственный университет, 2006. – 336 с.
  77. Тоффлер Э. Метаморфозы власти: Знание, богатство и сила на пороге XXI века / Э. Тоффлер / Пер. с англ. Науч. ред., предисл. П.С. Гуревича. – ООО «Издательство АСТ», 2002. – 670 с.
  78. Трахтенгерц Э.А. Компьютерная поддержка принятия решений / Э.А. Трахтенгерц. – СИНТЕГ, 1998. – 376 с.
  79. Тума Т. Эксклюзивный опрос в странах группы G20 / Т. Тума, Д. Хайльман. *WEB*: <http://inosmi.ru/politic/20160328/235888545.html>
  80. Тьюки Дж. Анализ результатов наблюдений. Разведочный анализ / Дж. Тьюки. – Мир, 1981. – 693 с. :ил.
  81. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере / Ю.Н. Тюрин, А.А. Макаров / Под. ред. В.Э. Фигурнова. – 3-е изд., перераб. и доп. – ИНФРА – М, 2003. – 544 с., ил.
  82. Форман Дж. Много цифр: Анализ больших данных при помощи Excel / Дж. Форман; пер. с англ. А. Соколовой. – Альпина Паблишер, 2016. – 461с.

83. Фрэнкс Б. Укрощение больших данных: как извлекать знания из больших массивов информации с помощью глубокой аналитики / Б. Фрэнкс; пер. с англ. А. Баранова. – Манн, Иванов и Фарбер, 2014. – 352 с.
84. Фрэнкс Б. Революция в аналитике: Как в эпоху *Big Data* улучшить ваш бизнес с помощью операционной аналитики / Б. Фрэнкс; пер. с англ. – Альпина Паблишер, 2016. – 316 с.
85. Хаббард Д. Как измерить все, что угодно. Оценка стоимости нематериального в бизнесе / Д. Хаббард; пер с англ. Е. Пестеревой. – ЗАО «Олимп-Бизнес», 2009. – 320 с.: илл.
86. Хэммонд Дж. Правильный выбор по принятию решений / Дж. Хэммонд, Р. Кини, Г. Райфа; пер с англ. В.Н. Егорова. – Бином. Лаборатория знаний, 2014. – 254 с.: ил.
87. Шарф К. Где место разуму? / К. Шарф.  
WEB: <http://inosmi.ru/science/20160329/235898814.html>
88. StatSoft, Inc. (2012). Электронный учебник по статистике. Москва, StatSoft.  
WEB: <http://www.statsoft.ru/home/textbook/default.htm>.
89. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей: Пер. с нем. / Ахим Бююль, Петер Цёфель. – СПб.: ООО «ДиаСофтЮП», 2002. – 608 с.
90. Berry M.W. *Survey of Text Mining I: Clustering, Classification, and Retrieval*. – Springer, 2003. – P. 261.
91. Bing B. *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*. – Springer, 2011. – P. 642.
92. Cronin B. *Annual Review of Information Science and Technology*. — ARIST, 2004. – P. 674.
93. Sivaramakrishnan J., Balakrishnan V. *Web Mining Functions in an Academic Search Application*. – 2009.
94. Srivastava J. Desikan P. Kumar V. *Web Mining – Concepts, Applications, and Research Directions*. – 2004.