МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования
«Воронежский государственный архитектурно-строительный университет»

УТВЕРЖДАЮ
Декан строительно-технологического факультета
____________ Власов В.В.
«28» 06 2013 г

РАБОЧАЯ ПРОГРАММА
дисциплины

«Теория вероятностей и математическая статистика»

Направление подготовки 020300 «Химия, физика и механика материалов»

Квалификация (степень) выпускника бакалавр

Нормативный срок обучения 4 года

Форма обучения очная

Авторы программы: ____________ д. ф.-м. н., профессор А.В. Лобода

Программа обсуждена на заседании кафедры высшей математики
«29» 05 2013 года Протокол № 10

Зав. кафедрой: ____________ д. т. н., профессор В. Н. Колпачёв

Воронеж 2013

1. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

1.1. Цели дисциплины: развитие логического и алгоритмического мышления, выработка умения самостоятельно расширять и углублять математические знания; освоение необходимого математического аппарата, помогающего анализировать, моделировать и решать прикладные задачи; формирование у студента уровня математической культуры, достаточного для продолжения образования, научной работы или практической деятельности, методологических основ для формирования целостного научного мировоззрения, отвечающего современному уровню развития человеческой цивилизации.

1.2. Задачи освоения дисциплины:

- Выработка ясного понимания необходимости математического образования в подготовке бакалавра и представления о роли и месте вероятностных процессов в современной системе знаний и мировой культуре;
- Ознакомление с системой понятий, используемых для описания важнейших вероятностных и статистических моделей;
- Формирование конкретных практических приемов и навыков постановки и решения вероятностных и статистических задач, ориентированных на практическое применение при изучении дисциплин профессионального цикла;
- Овладение основными математическими методами, необходимыми для анализа вероятностных и статистических процессов и явлений при поиске оптимальных решений, обработки и анализа результатов экспериментов.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП

Дисциплина «Теория вероятности и математической статистики» относится к базовой (обязательной) части математического цикла учебного плана.

Студент, приступая к изучению дисциплины должен обладать знаниями, умениями и навыками в области математического анализа и геометрии.

Дисциплина «Теория вероятностей и математическая статистика» является предшествующей таким дисциплинам естественнонаучного цикла как: «Техническая механика», «Основы метрологии, стандартизации и сертификации», «Физика и химия поверхности».

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс изучения дисциплины «Теория вероятности и математической статистики» направлен на формирование следующих компетенций:

● наличие культуры мышления, способность к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения (ОК-1);

● использование основных законов естественнонаучных дисциплин в профессиональной деятельности, применение методов математического анализа и моделирования, теоретического и экспериментального исследования (ОК-10);

● способность использовать в познавательной и в профессиональной деятельности базовые знания в области математики и естественных наук (ПК-3);

● использование феноменологических, математических и численных (альтернативных) моделей для описания и прогнозирования различных явлений, осуществление их качественного и количественного анализа (ПК-11);

● наличие системных представлений о возможностях применения фундаментальных законов физики, химии, математики и механики для объяснения свойств и поведения широкого спектра разнообразных функциональных материалов и наноматериалов, предназначенных для электроники и здравоохранения (ПК-15);

● использование основ математического анализа; алгебры, геометрии и дискретной математики; теории дифференциальных уравнений и численных методов; теории вероятности и математической статистики; физических основ механики, физики колебаний и волн, статистической физики и термодинамики, электричества и магнетизма, квантовой физики, языков программирования и стандартного программного обеспечения для профессиональной деятельности (ПК-21).

В результате изучения дисциплины обучающийся должен:

знать:

- основы теории вероятностей и математической статистики, необходимые для решения экономических задач;

уметь:

- применять методы моделирования, теоретического и экспериментального исследования для решения экономических задач;

владеть:

- навыками применения современного математического инструментария для решения различных технических задач; методикой построения, анализа и применения математических моделей для оценки состояния и прогноза развития экономических явлений и процессов.

4. ОБЪЕМ ДИСЦИПЛИНЫ И ВИДЫ УЧЕБНОЙ РАБОТЫ

Общая трудоемкость дисциплины «Теория вероятности и математической статистики» составляет 3 зачетных единицы, 108 часа.

Вид учебной работы		Всего часов	Семестры
			4
Аудиторные занятия (всего)		36	36
В том числе:			
Лекции		18	18
Практические занятия (ПЗ)		18	18
Лабораторные работы (ЛР)			
Самостоятельная работа (всего)		36	36
В том числе:			
Курсовой проект			
Расчетно-графическая работа (количество)			
Вид промежуточной аттестации (зачет, экзамен)			зачет
Общая трудоемкость	час	72	72
	зач. ед.	2	2

5. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

5.1. Содержание разделов дисциплины

№ п/п	Наименование раздела дисциплины	Содержание раздела
1.	Случайные события	Пространство элементарных событий, алгебра событий. Основные формулы комбинаторики. Вероятность события и ее свойства. Условная вероятность. Формула умножения вероятностей. Статистическая зависимость между событиями. Формула сложения вероятностей. Формула полной вероятности и формула Байеса. Последовательность независимых испытаний. Схема и формула Бернулли и следствия из нее. Предельные теоремы Муавра-Лапласа и Пуассона.
2	Случайные величины	Определение, классификация (дискретные и непрерывные случайные величины), способы задания. Математические операции над случайными величинами. Функция распределения вероятностей случайной величины и ее свойства. Плотность распределения вероятностей случайной величины и ее свойства. Числовые характеристики случайных величин. Математическое ожидание случайной величины и его свойства. Дисперсия и среднее квадратическое отклонение случайной величины и их свойства. Моменты случайных величин. Скошенность и эксцесс. Примеры дис-

		кретных и непрерывных распределений. Биномиальное распределение и распределение Пуассона. Равномерное распределение вероятностей. Показательное распределение. Нормальный закон распределения вероятностей. Нормальная кривая. Функция Лапласа. Вычисление вероятности попадания в заданный интервал нормальной случайной величины. Правило трех сигм.
3	Многомерные случайные величины и функции случайных величин.	Определение, классификация, способы задания многомерных случайных величин. Функция распределения вероятностей двумерной случайной величины и ее свойства. Плотность распределения вероятностей двумерной случайной величины и ее свойства. Условные законы распределения вероятностей. Числовые характеристики многомерных случайных величин. Линии регрессии. Ковариация и коэффициент корреляции. Двумерное нормальное распределение. Функция одной случайной величины. Функция нескольких случайных величин. Теоремы о математических ожиданиях и дисперсиях функций от случайных величин. Некоторые специальные законы распределения, применяемые в математической статистике (распределение «*хи*-квадрат»; распределение Стьюдента; распределение Фишера-Снедекора).
4	Предельные теоремы теории вероятностей и элементы теории случайных процессов	Закон больших чисел в форме Чебышева. Неравенство Чебышева. Теорема Чебышева. Теорема Бернулли. Центральная предельная теорема Ляпунова. Понятие случайного процесса и случайной функции. Математическое ожидание, дисперсия и корреляционная функция случайного процесса. Стационарные случайные процессы. Примеры. Понятие марковского случайного процесса. Понятие о математическом моделировании случайных процессов.
5.	Математическая статистика и методы обработки экспериментальных данных	*Основы выборочного метода.* Выборка. Генеральная совокупность. Вариационный ряд. Эмпирическая функция распределения. Полигон и гистограмма. Точечные оценки параметров распределения. Свойства оценок. Понятие несмещенности, эффективности, состоятельности оценок. Несмещенность и состоятельность выборочного среднего как оценки математического ожидания. Смещенность выборочной дисперсии. Пример несмещенной оценки дисперсии. Методы нахождения оценок. Интервальное оценивание неизвестных параметров. Доверительная вероятность и доверительный интервал. Построение доверительных интервалов для параметров нормального распределе-

		ния. Построение доверительного интервала для неизвестной вероятности события. Оценки истинного значения измеряемой величины и точности измерений. *Проверка статистических гипотез.* Общие принципы проверки гипотез. Понятия статистической гипотезы, ошибок первого и второго рода, статистического критерия. Проверка гипотез о параметрах нормального распределения. Проверка гипотезы о виде распределения (критерий *хи*-квадрат). *Корреляционный и регрессионный анализ.* Понятие функциональной, статистической и корреляционной зависимости. Линейная парная регрессия. Выборочный коэффициент корреляции. Корреляционная таблица. Выборочное корреляционное отношение. Линейный множественный регрессионный анализ. Понятие о многомерном корреляционном анализе. Множественный регрессионный анализ. *Дисперсионный анализ.* Постановка задачи дисперсионного анализа. Однофакторный дисперсионный анализ. Понятие о двухфакторном дисперсионном анализе.

5.2 Разделы дисциплины и междисциплинарные связи с обеспечиваемыми (последующими) дисциплинами

№ п/п	Наименование обеспечиваемых (последующих) дисциплин	№ № разделов данной дисциплины, необходимых для изучения обеспечиваемых (последующих) дисциплин							
		1	**2**	**3**	**4**	**5**			
1.	Техническая механика	+	+	+	+	+			
2.	Основы метрологии, стандартизации и сертификации	+	+	+	+	+			
3.	Физика и химия поверхности	+	+	+	+	+			

5.3. Разделы дисциплин и виды занятий

№ п/п	Наименование раздела дисциплины	Лекц.	ПЗ	ЛР	СРС	Всего
1.	Случайные события	4	4	-	8	16
2	Случайные величины	4	4	-	8	16
3	Многомерные случайные величины и функции случайных величин.	4	4	-	8	16
4	Предельные теоремы теории вероятностей	4	4	-	8	16
5	Математическая статистика и методы обработки экспериментальных данных	2	2	-	4	8

5.4. ПРАКТИЧЕСКИЕ ЗАНЯТИЯ

№ п/п	Тематика практических занятий	Трудоемкость (час)
1	Вычисление вероятностей случайных событий. Освоение схемы Бернулли.	4
2	Описание законов распределений дискретных и непрерывных случайных величин, вычисление их числовых характеристик Решение задач, связанных с нормальным законом распределения.	4
3	Вычисление числовых характеристик многомерных случайных величин, ковариация и коэффициента корреляции. Построение линии регрессии. Решение задач, связанных с двумерным нормальным распределением.	2
4	Оценки числа испытаний, гарантирующих появление события с заданной вероятностью. Применение теорем Лапласа и вероятностных таблиц.	4
5	Составление точечных и интервальных распределений выборки, построение геометрических характеристик выборки, нахождение точечных и интервальных оценок генеральных параметров. Проверка статистических гипотез. Отыскание выборочного уравнения прямой линии регрессии.	4

6. ЛАБОРАТОРНЫЙ ПРАКТИКУМ

Лабораторный практикум учебным планом не предусмотрен.

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

7.1. Этап промежуточного контроля знаний

Результаты промежуточного контроля знаний (зачет) оцениваются по двухбалльной шкале с оценками:

- «зачтено»;
- «не зачтено».

Дескриптор компетенции	Показатель оценивания	Оценка	Критерий оценивания
Знает	основы математического анализа, теории вероятности и математической статистики (ПК-21).	зачтено	1. Студент демонстрирует полное понимание заданий. Все требования, предъявляе-
Умеет	использовать основные зако-		

Дескриптор компетенции	Показатель оценивания	Оценка	Критерий оценивания
	ны естественнонаучных дисциплин, примененять методы математического анализа и моделирования, теоретического и экспериментального исследования (ОК-10); использовать базовые знания в области математики и естественных наук (ПК-3); использовать феноменологические, математические и численные модели для описания и прогнозирования различных явлений, осуществить их качественный и количественный анализ (ПК-11);		мые к заданию выполнены. 2. Студент демонстрирует значительное понимание заданий. Все требования, предъявляемые к заданию выполнены. 3. Студент демонстрирует частичное понимание заданий. Большинство требований, предъявляемых к заданию выполнены.
Владеет	культурой мышления, способностью к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения (ОК-1); системными представлениями о возможностях применения фундаментальных законов физики, химии, математики и механики (ПК-15)		
Знает	основы математического анализа, теории вероятности и математической статистики (ПК-21).	не зачтено	1. Студент демонстрирует лишь поверхностное понимание заданий. Многие требования, предъявляемые к заданию не выполнены. 2. Студент демонстрирует непонимание заданий. 3. У студента нет ответа. Не было
Умеет	использовать основные законы естественнонаучных дисциплин, примененять методы математического анализа и моделирования, теоретического и экспериментального исследования (ОК-10); использовать базовые знания в области математики и естественных наук (ПК-3);		

Дескриптор компетенции	Показатель оценивания	Оценка	Критерий оценивания
	использовать феноменологические, математические и численные модели для описания и прогнозирования различных явлений, осуществить их качественный и количественный анализ (ПК-11);		попытки выполнить задание.
Владеет	культурой мышления, способностью к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения (ОК-1); системными представлениями о возможностях применения фундаментальных законов физики, химии, математики и механики (ПК-15)		

7.3.6. Паспорт фонда оценочных средств

№ п/п	Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или ее части)	Наименование оценочного средства
8	Теория вероятностей и основы математической статистики	ОК-1, ОК-10, ПК_3, ПК-11, ПК-15, ПК-21	Контрольная работа (КР) Тестирование (Т) Коллоквиум (КЛ) Зачет

8. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ ПРОЕКТОВ И КОНТРОЛЬНЫХ РАБОТ

Курсовой проект и контрольные работы учебным планом не предусмотрены.

9. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Текущий контроль успеваемости осуществляется на практических занятиях: в виде опроса теоретического материла и умения применять его к решению задач у доски, в виде проверки домашних заданий, в виде тестирования по отдельным темам.

Промежуточный контроль осуществляется проведением контрольных работ по отдельным разделам дисциплины, тестирования по разделам дисциплины, изученным студентом в период между аттестациями, проведением коллоквиумов по теоретическому материалу, выполнением расчетно- графических работ.

Изучение дисциплины заканчивается экзаменом.

9.1. Примерная тематика контрольных работ и расчетно- графических работ

Контрольные работы проводятся на практических занятиях или вне их, в рамках самостоятельной работы под контролем преподавателя.

Варианты расчетно - графических работ выдаются каждому студенту индивидуально.

1. Контрольная работа № 1: «Случайные события и случайные величины».
2. Контрольная работа № 2: «Корреляционный и регрессионный анализ.».
3. Расчетно- графическая работа № 1: «Теория вероятностей».
4. Расчетно- графическая работа № 2: «Математическая статистика».

9.2. Примерный перечень вопросов к коллоквиумам

1-й коллоквиум «Случайные события»

1. Опыт и событие. Классификация событий. Вероятность события. Классическое определение вероятности. Статистическая вероятность.
2. Сумма событий. Теорема сложения вероятностей несовместных событий.
3. Произведение событий. Условная вероятность. Теорема умножения вероятностей. Зависимость и независимость событий.
4. Теорема сложения вероятностей совместных событий. Вероятность появления хотя бы одного события.
5. Формула полной вероятности. Формулы Байеса.
6. Повторение испытаний. Формула Бернулли. Формула Пуассона. Локальная и интегральная теоремы Лапласа.

2-й коллоквиум «Случайные величины»

1. Понятие случайной величины. Дискретные и непрерывные случайные величины. Закон распределения вероятностей дискретной случайной величины.
2. Числовые характеристики дискретной случайной величины: математическое ожидание, дисперсия моменты, среднеквадратическое отклонение, их свойства, роль и назначение.
3. Основные распределения дискретных случайных величин (биномиальное распределение, распределение Пуассона).
4. Функция распределения вероятностей случайной величины, ее свойства.
5. Непрерывные случайные величины. Плотность распределения вероятности, ее свойства.
6. Числовые характеристики непрерывной случайной величины: математическое ожидание, дисперсия.
7. Распределения непрерывных случайных величин: равномерное, показательное, нормальное. Понятие о негауссовых распределениях.
8. Числовые характеристики нормального распределения, его свойства, интеграл вероятностей, правило 3-х сигм.

9.3. Варианты тестов для контроля промежуточных знаний

1. Имеется три группы студентов: в первой 11 человек, во второй 18 человек, в третьей 20 человек. Количество способов выбора тройки студентов, в которой по одному студенту из каждой группы, равно…

1. $11\cdot18\cdot20$ 2. $\dfrac{11+18+20}{3}$ 3. $\dfrac{11\cdot18\cdot20}{3}$ 4. $11+18+20$

2. Число способов поставить 5 человек в очередь равно…

3. В слове «*WORD*» меняют местами буквы. Тогда количество всех возможных различных «слов» равно…

1. 8   2. 16   3. 4   4. 24

4. В коробке 6 цветных карандашей. Число способов выбрать три из них равно…

5. Число способов выбрать из группы в 20 студентов старосту и заместителя равно…

6. Из ящика, где находится 15 деталей, пронумерованных от 1 до 15, требуется вынуть 3 детали. Тогда количество всевозможных комбинаций номеров вынутых деталей равно…

1. $\dfrac{15!}{12!}$ 2. $\dfrac{15!}{3!\cdot12!}$ 3. $3!$ 4. $15!$

7. Число трехзначных чисел, которые можно составить из четырех карточек с цифрами 1, 2, 5, 7, равно…

8. Количество способов выбора стартовой пятерки из восьми игроков баскетбольной команды равно…

1. 120   2. 109   3. 336   4. 56

9. Решением уравнения $4C_{x+5}^{2}-A_{x+1}^{2}=x^{2}+74$ является…

1. 4   2. 5   3. 2   4. 8

10. В каком случае верно, что *A* влечет за собой *B* при бросании кости. Если:

1. *A* – появление четного числа очков, *B* – появление 6 очков
2. *A* – появление 4 очков, *B* – появление любого четного числа очков
3. *A* – выпадение любого нечетного числа очков, *B* – появление 3 очков

4. A – появление любой грани, кроме 6, B – появление 3 очков

11. Какое утверждение неверно, если говорят о противоположных событиях:

1. Событие, противоположное достоверному, есть невозможное
2. Сумма вероятностей двух противоположных событий равна единице
3. Если два события единственно возможны и несовместны, то их называют противоположными
4. Вероятность появления одного из противоположных событий всегда больше вероятности другого

12. Если два события A и B образуют полную группу, то для их вероятностей выполнено соотношение…

1. $p(A)=p(B)$ 3. $p(A)\cdot p(B)=0$
2. $p(A)=-p(B)$ 4. $p(A)=1-p(B)$

13. Если E – достоверное событие и события $A_1, A_2, ..., A_n$ образуют полную группу, то выполнено(ы) соотношение(я)…

1. $A_1 + A_2 + ... + A_n = E$ 3. $A_i + A_j = \emptyset$ для $i \neq j$
2. $A_i \cdot A_j = 1$ для $i \neq j$ 4. $A_1 \cdot A_2 \cdot ... \cdot A_n = E$

14. Бросают два кубика. События A – «на первом кубике выпала шестерка», B – «на втором кубике выпала шестерка» являются:

1. несовместными
2. совместными
3. независимыми
4. зависимыми

15. Из каждой из двух колод вынимают по одной карте. События A – «карта из первой колоды – красной масти» и B – «карта из второй колоды – бубновой масти» являются:

1. несовместными
2. совместными
3. независимыми
4. зависимыми

16. Случайные события A и B, удовлетворяющие условиям $P(A)=0{,}3$, $P(B)=0{,}4$, $P(AB)=0{,}2$, являются…

1. несовместными и зависимыми
2. совместными и независимыми
3. совместными и зависимыми
4. несовместными и независимыми

17. A и B – случайные события. A и B независимы, если выполнено…

1. $p(A)=p(B)$ 3. $p(A)=p(B)\cdot p(A/B)$
2. $p(AB)=\dfrac{p(A)}{p(B)}$ 4. $p(AB)=p(A)p(B)$

18. A и B – случайные события. Верным является утверждение…

1. $p(A+B)=p(A)+p(B)-p(AB)$
2. $p(A+B)=p(A)+p(B)-2p(AB)$
3. $p(A+B)=p(A)+p(B)+p(AB)$
4. $p(A+B)=p(A)\cdot p(B)$

19. Вероятность наступления некоторого события *не может* быть равна…

1. 1 2. 0 3. 4 4. 0,4

20. В урне находятся 6 шаров: 3 белых и 3 черных. Событие A – «Вынули белый шар». Событие B – «Вынули черный шар». Опыт состоит в выборе только одного шара. Тогда для этих событий *неверным* будет утверждение:

1. «События A и B несовместны»
2. «Вероятность события B равна $\frac{1}{2}$»
3. «Событие A невозможно»
4. «События A и B равновероятны»

21. Игральный кубик бросается один раз. Тогда вероятность того, что на верхней грани выпадет 2 очка, равна…

1. $\frac{1}{2}$ 2. $\frac{1}{6}$ 3. $\frac{1}{5}$ 4. $\frac{2}{3}$

22. Игральный кубик бросается один раз. Тогда вероятность того, что на верхней грани выпадет нечетное число очков, равна…

1. $\frac{1}{3}$ 2. $\frac{1}{6}$ 3. 0,1 4. $\frac{1}{2}$

23. Расположите случайные события в порядке возрастания их вероятностей:

A – при бросании кубика выпало не более 5 очков
B – при бросании кубика выпало нечетное число очков
C – при двух бросаниях кубика выпало в сумме не менее двух очков

24. В лотерее 1000 билетов. На один билет выпадает выигрыш 5000 рублей, на десять билетов – выигрыши по 1000 рублей, на пятьдесят билетов – выигрыши по 200 рублей, на сто билетов – выигрыши по 50 рублей; остальные билеты проигрышные. Покупается один билет. Тогда вероятность не выигрыша равна…

1. 0,839 2. $\frac{161}{839}$ 3. 0,849 4. 0,161.

25. В урне находится 5 белых и 3 черных шара. Из урны вынимаются четыре шара. Вероятность того, что три шара будут белыми, а один черным, равна…

1. $\frac{3}{7}$ 2. $\frac{1}{3}$ 3. $\frac{5}{8}$ 4. $\frac{3}{8}$

26. Два стрелка производят по одному выстрелу. Вероятности попадания в цель для первого и второго стрелков равны 0,7 и 0,2 соответственно. Тогда вероятность того, что в цель попадут оба стрелка, равна…

1. 0,9 2. 0,24 3. 0,15 4. 0,14

27. По оценкам экспертов вероятности банкротства для двух предприятий, производящих разнотипную продукцию, равны 0,3 и 0,5. Тогда вероятность банкротства *только одного* предприятия равна…

1. 0,80 2. 0,85 3. 0,52 4. 0,50

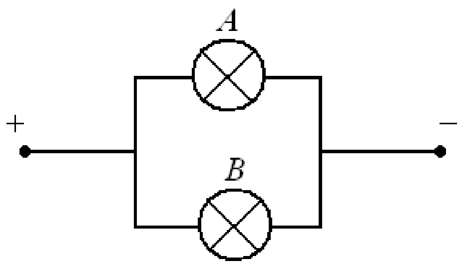
28. В урне из 8 шаров имеется 3 красных. Наудачу берут два шара. Тогда вероятность того, что среди них ровно один красный шар, равна…

1. $\frac{1}{15}$ 2. $\frac{15}{28}$ 3. $\frac{1}{4}$ 4. $\frac{15}{56}$

29. В урне лежит 3 белых и 3 черных шара. Последовательно, без возвращения и наудачу извлекают 3 шара. Тогда вероятность того, что все они будут белыми, равна…

1. $\frac{1}{9}$ 2. $\frac{1}{20}$ 3. $\frac{8}{27}$ 4. $\frac{6}{125}$

30. В электрическую цепь включены *параллельно* два прибора A и B. При подаче напряжения прибор A сгорает с вероятностью 0,01, прибор B – с вероятностью 0,05. Считаем, что через сгоревший прибор ток не идет. Тогда вероятность того, что при включении напряжения ток пройдет через цепь, равна…



1. 0,94 2. 0,95 3. 0,9405 4. 0,9995

31. Вероятность того, что один станок сломается в течение смены, равна 0,2. Тогда вероятность того, что в течение смены из трех станков откажет хотя бы один, равна…

1. 0,64 2. 0,2 3. 0,512 4. 0,488

32. Игральная кость брошена 3 раза. Тогда вероятность того, что хотя бы один раз выпадет число, делящееся на три, равна…

1. $\frac{16}{27}$ 2. $\frac{19}{27}$ 3. $\frac{8}{27}$ 4. $\frac{1}{3}$

33. Два стрелка производят по одному выстрелу. Вероятность попадания в цель для первого и второго стрелков равны 0,4 и 0,9 соответственно. Тогда вероятность того, что цель будет поражена, равна…

1. 0,994 2. 0,36 3. 0,64 4. 0,94

34. По мишени производится четыре выстрела. Значение вероятности промаха при первом выстреле 0,5; при втором – 0,3; при третьем – 0,2, при четвертом – 0,1. Тогда вероятность того, что мишень *не будет поражена ни разу*, равна…

1. 0,275 2. 0,003 3. 1,1 4. 0,03

35. В урне находятся 2 белых, 1 красный, 2 зеленых и 3 черных шара. Из урны поочередно вынимают три шара, но после первого вынимания шар возвращается в урну, и шары в урне перемешиваются. Тогда значение вероятности того, что все извлеченные шары белые, равно…

1. $\frac{1}{112}$ 2. $\frac{1}{64}$ 3. $\frac{1}{128}$ 4. $\frac{1}{126}$

36. С первого станка на сборку поступает 40 %, со второго – 60 % всех деталей. Среди деталей, поступивших с первого станка, 5 % бракованных, со второго – 1 % бракованных. Тогда вероятность того, что поступившая на сборку деталь бракованная, равна…

1. 0,03 2. 0,06 3. 0,024 4. 0,026

37. Имеются две одинаковые на вид урны. В первой урне находятся два белых, два зеленых и три черных шара. Во второй урне – три белых два красных и три черных шара. Из наудачу взятой урны взяли одновременно два шара. Тогда вероятность того, что оба шара черные, равна…

1. $\frac{2}{15}$ 2. $\frac{2}{5}$ 3. $\frac{3}{28}$ 4. $\frac{1}{8}$

38. В первой урне 3 белых и 7 черных шаров. Во второй урне 6 белых и 4 черных шаров. Из наудачу взятой урны вынули один шар. Тогда вероятность того, что этот шар окажется белым, равна…

1. 0,45 2. 0,9 3. 0,5 4. 0,15

39. Событие A может наступить лишь при условии появления одного из двух несовместных событий B_1 и B_2, образующих полную группу событий. Известны вероятности $P(B_1)=\frac{1}{4}$, $P(A)=\frac{1}{6}$ и условная вероятность $P(A/B_1)=\frac{1}{3}$. Тогда условная вероятность $P(A/B_2)$ равна…

1. $\frac{5}{6}$ 2. $\frac{2}{3}$ 3. $\frac{3}{4}$ 4. $\frac{1}{9}$

40. С первого станка на сборку поступает 60 %, со второго – 40 % всех деталей. Среди деталей, поступивших с первого станка, 90 % стандартных, со второго – 80 %. Взятая наудачу деталь оказалась стандартной. Тогда вероятность того, что она изготовлена *на втором станке*, равна…

1. $\frac{16}{43}$ 2. $\frac{3}{7}$ 3. $\frac{8}{25}$ 4. $\frac{27}{43}$

41. Событие A может наступить лишь при условии появления одного из двух несовместных событий B_1 и B_2, образующих полную группу событий. Известны вероятности $P(B_1)=\frac{3}{4}$ и условные вероятности $P(A/B_1)=\frac{1}{4}$, $P(A/B_2)=\frac{1}{2}$. Тогда вероятность $P(A)$ равна…

1. $\frac{3}{4}$ 2. $\frac{1}{4}$ 3. $\frac{3}{16}$ 4. $\frac{5}{16}$

42. Монета брошена 4 раза. Тогда вероятность того, что «герб» выпадет ровно три раза, равна…

1. $\frac{1}{4}$ 2. $\frac{1}{8}$ 3. $\frac{3}{4}$ 4. $\frac{3}{8}$

43. Вероятность появления события A в 20 независимых испытаниях, проводимых по схеме Бернулли, равна 0,9. Тогда математическое ожидание числа появлений этого события равно…

1. 17,1 2. 1,8 3. 18 4. 2

44. Вероятность появления события A в 40 независимых испытаниях, проводимых по схеме Бернулли, равна 0,8. Тогда дисперсия числа появлений этого события равна…

1. 0,02 2. 0,64 3. 32 4. 6,4

45. Проводятся независимые испытания каждого из 12 элементов устройства. Вероятность, что элемент выдержит испытание, равна 0,8. Тогда наивероятнейшее число элементов, выдержавших испытание, равно…

1. 9 2. 11 3. 12 4. 10

46. Страхуется 1200 автомобилей; считается, что каждый из них может попасть в аварию с вероятностью 0,08. Для вычисления вероятности того, что количество аварий среди всех застрахованных автомобилей не превзойдет 100, следует использовать…

1. интегральную формулу Муавра-Лапласа
2. формулу Пуассона
3. формулу полной вероятности
4. формулу Байеса

47. Дискретная случайная величина X имеет закон распределения вероятностей:

X	–1	2
P	0,3	0,7

Тогда математическое ожидание $M(X)$ этой случайной величины равно…

1. 0,4 2. 1,7 3. 1 4. 1,1

48. Дискретная случайная величина X задана законом распределения вероятностей:

X	1	3	5	6
P	a	0,2	0,6	0,1

Пусть $M(X)$ – математическое ожидание. Тогда $10 \cdot M(X)$ равно…

49. Дискретная случайная величина X задана законом распределения вероятностей:

X	–1	0	2
P	0,1	0,3	0,6

Тогда математическое ожидание случайной величины $Y = 3X$ равно…

1. 3,9 2. 4,1 3. 3 4. 3,3

50. Функция распределения вероятностей дискретной случайной величины X имеет вид $F(X)=\begin{cases}0, & x\le 2,\\ 0{,}2, & 2<x\le 4,\\ 0{,}7, & 4<x\le 5,\\ 1, & x>5.\end{cases}$ Тогда вероятность $P(1\le X\le 3)$ равна…

1. 0,2 2. 0,5 3. 0,7 4. 0,9

51. Дискретная случайная величина X задана законом распределения вероятностей:

X	–5	–3	x_3
P	0,3	0,4	0,3

Если математическое ожидание $M(X)=-2{,}4$, то значение x_3 равно…

1. 0 2. 2 3. 1 4. –1

52. Функция распределения вероятностей дискретной случайной величины X имеет вид $F(X)=\begin{cases}0, & x\le 2,\\ 0{,}2, & 2<x\le 3,\\ 0{,}8, & 3<x\le 4,\\ 1, & x>4.\end{cases}$ Тогда математическое ожидание случайной величины X равно…

1. 3,8 2. 3 3. 2 4. 4,8

53. Дискретная случайная величина X задана законом распределения вероятностей:

X_i	0	2	4	6
P_i	0,1	0,1	0,1	0,7

Тогда значение интегральной функции распределения вероятностей $F(3)$ равно…

1. 0,1 2. 0,2 3. 0,3 4. 0,8

54. Дискретная случайная величина X задана законом распределения вероятностей:

X	–2	–1	0	1	2
P	0,1	0,2	0,2	0,4	0,1

Тогда вероятность $P(|X|\le 1)$ равна…

1. 0,3 2. 0,8 3. 0,9 4. 0,5

55. Дискретная случайная величина X задана законом распределения вероятностей:

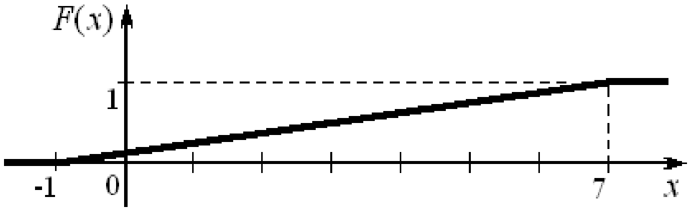
X	–1	1	2	4
P	0,2	0,1	a	b

Её математическое ожидание равно 2,3, если…

1. $a=0{,}4, b=0{,}3$ 3. $a=0{,}8, b=0{,}2$

2. $a=0{,}2, b=0{,}5$ 4. $a=0{,}5, b=0{,}2$

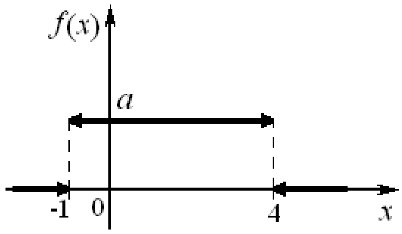
56. График функции распределения вероятностей непрерывной случайной величины X, распределенной равномерно в интервале $(-1;7)$, имеет вид:



Тогда математическое ожидание X равно…

1. 7 2. 4 3. 8 4. 3

57. График плотности распределения вероятностей непрерывной случайной величины X, распределенной равномерно в интервале $(-1;4)$, имеет вид:



Тогда значение a равно…

1. 0,20 2. 0,33 3. 0,25 4. 1

58. Непрерывная случайная величина X задана плотностью распределения вероятностей $f(x)=\frac{1}{2\sqrt{2\pi}}e^{-\frac{(x-3)^2}{8}}$. Тогда дисперсия этой нормально распределенной случайной величины равна…

1. 3 2. 2 3. 4 4. 8

59. Непрерывная случайная величина X задана плотностью распределения вероятностей $f(x)=\frac{1}{7\sqrt{2\pi}}e^{-\frac{(x-8)^2}{98}}$. Тогда математическое ожидание этой нормально распределенной случайной величины равно…

1. 8 2. 7 3. 49 4. 98

60. Точечная оценка параметра распределения равна 20. Тогда его интервальная оценка может иметь вид…

1. $(0; 20)$ 2. $(19; 21)$ 3. $(20; 21)$ 4. $(19; 20)$

61. Из генеральной совокупности извлечена выборка объема $n = 50$:

x_i	1	2	3	4
n_i	10	9	8	n_4

Тогда n_4 равно…

1. 7 2. 50 3. 23 4. 24

62. Проведено 5 измерений (без систематических ошибок) некоторой случайной величины (в мм): 4; 5; 8; 9; 11. Тогда несмещенная оценка математического ожидания равна…

1. 7,4 2. 9,25 3. 7,6 4. 8

63. Мода вариационного ряда 1, 4, 4, 5, 6, 8, 9 равна…

1. 4 2. 1 3. 9 4. 5

64. В результате измерений некоторой физической величины одним прибором (без систематических ошибок) получены следующие результаты (в мм): 11, 13, 15. Тогда несмещенная оценка дисперсии измерений равна…

1. 3 2. 8 3. 4 4. 13

65. Выборочное уравнение парной регрессии имеет вид $y = -3{,}2 + 1{,}6x$. Тогда выборочный коэффициент корреляции может быть равен…

1. 0,9 2. -3,2 3. -0,5 4. –0,9

66. При построении уравнения парной регрессии $y = \alpha + \beta x + \varepsilon$ были получены следующие результаты: $r_B = 0{,}8$, $\sigma_x = 2$, $\sigma_y = 1{,}5$. Тогда коэффициент регрессии β равен…

1. 0,6 2. 0,3 3. 0,75 4. 2,4

9.4. Примерный перечень вопросов к зачету

4-й семестр

1. Опыт и событие. Классификация событий. Вероятность события. Классическое определение вероятности. Статистическая вероятность.
2. Сумма событий. Теорема сложения вероятностей несовместных событий.
3. Произведение событий. Условная вероятность. Теорема умножения вероятностей. Зависимость и независимость событий.
4. Теорема сложения вероятностей совместных событий. Вероятность появления хотя бы одного события.
5. Формула полной вероятности. Формулы Байеса.
6. Повторение испытаний. Формула Бернулли. Формула Пуассона. Локальная и интегральная теоремы Лапласа.
7. Понятие случайной величины. Дискретные и непрерывные случайные величины. Закон распределения вероятностей дискретной случайной величины.
8. Числовые характеристики дискретной случайной величины: математическое ожидание, дисперсия моменты, среднеквадратическое отклонение, их свойства, роль и назначение.
9. Основные распределения дискретных случайных величин (биномиальное распределение, распределение Пуассона).
10. Функция распределения вероятностей случайной величины, ее свойства.
11. Непрерывные случайные величины. Плотность распределения вероятности, ее свойства.
12. Числовые характеристики непрерывной случайной величины: математическое ожидание, дисперсия.
13. Распределения непрерывных случайных величин: равномерное, показательное, нормальное. Понятие о негауссовых распределениях.
14. Числовые характеристики нормального распределение, его свойства, интеграл вероятностей, правило 3-х сигм.
15. Системы случайных величин, их числовые характеристики. Ковариация и коэффициент корреляции.
16. Функции случайных величин, их числовые характеристики.
17. Последовательность случайных величин. Сходимость последовательности по вероятности. Теорема Чебышева. Теорема Бернулли. Центральная предельная теорема.
18. Случайные процессы. Классификация случайных процессов. Потоки событий.
19. Задачи математической статистики. Выборочный метод. Генеральная совокупность, выборка. Причины применения выборочного метода. Случайный отбор.
20. Выборка. Эмпирическая функция распределения. Построение интервального вариационного ряда распределения. Графическое изображение вариационных рядов.

21. Точечные оценки числовых характеристик и параметров распределения генеральной совокупности. Требования к точечным оценкам: состоятельность, несмещенность, эффективность.
22. Состоятельность и несмещенность выборочного среднего как оценки генерального математического ожидания. Свойства выборочной дисперсии (смещенность, состоятельность). Несмещенная оценка дисперсии.
23. Методы моментов и максимального правдоподобия получения оценок параметров генерального распределения.
24. Доверительный интервал (интервальная оценка) числовой характеристики или параметра генерального распределения. Точность и надежность оценки.
25. Доверительный интервал для генеральной средней при известной генеральной дисперсии.
26. Доверительный интервал для генеральной средней при неизвестной генеральной дисперсии.
27. Статистическая гипотеза. Критерий проверки. Статистика критерия. Уровень значимости. Ошибки первого и второго рода. Критерий согласия. Общая схема проверки статистической гипотезы.
28. Проверка гипотез о сравнении характеристик положения и рассеяния (критерии Фишера, Стьюдента, непараметрические критерии).
29. Проверка гипотезы о принадлежности закону распределений (критерий согласия Пирсона).
30. Понятие функциональной, статистической и корреляционной зависимости. Линейная парная регрессия. Выборочный коэффициент корреляции. Анализ криволинейных связей. Корреляционная таблица. Выборочное корреляционное отношение.
31. Задача регрессии. Эмпирическая простая линейная регрессия. Метод наименьших квадратов построения регрессии.
32. Прямые линейной эмпирической регрессии «*Y* на *X*» и «*X* на *Y*». Проверка адекватности эмпирической простой линейной регрессии опытным данным.
33. Линейный множественный регрессионный анализ. Множественный корреляционный анализ.

10. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

10.1 Основная литература:

1 Гмурман В.Е./ Теория вероятностей и математическая статистика. –Высшее образование 2009.-478
2 Гмурман В.Е./ Теория вероятностей и математическая статистика. –Высшее образование 2010.-478
3 Гмурман В.Е./ Теория вероятностей и математическая статистика. –Высшее образование 2012.-478

4 Гмурман, Владимир Ефимович Теория вероятностей и математическая статистика: учебное пособие для бакалавров : рекомендовано Министерством образования и науки Российской Федерации. - 12-е изд.. - Москва : Юрайт , 2013 - 478, [1] с.

10.2 Дополнительная литература:

1. Дементьева, Александра Марковна Интегральное исчисление функций одной и нескольких переменных [Электронный ресурс] : учеб. пособие : рек. ВГАСУ / Дементьева, Александра Марковна, Артыщенко, Степан Владимирович, Попова, Виктория Анатольевна ; Воронеж. гос. архит.-строит. ун-т. - Воронеж : [б. и.], 2010. - 1 электрон. опт. диск.
2. Седаев, Александр Андреевич. Избранные главы курса математики в инженерном вузе (множества, графы, топология, функциональный анализ, вариационное исчисление) [Электронный ресурс] : учеб. пособие : рек. ВГАСУ / Седаев, Александр Андреевич, Стенюхин, Леонид Витальевич, Евченко, Валерия Константиновна ; под ред. С. М. Алейникова ; Воронеж. гос. архит.-строит. ун-т. - Воронеж : [б. и.], 2008. - 1 электрон. опт. диск (CD-R).
3. Теория вероятностей и математическая статистика. Учебник(2010, Балдин К.В., Дашков и К) - ЭБС IPRbooks.Теория вероятностей и математическая статистика. Учебное пособие (2013, Мхитарян В.С., Астафьева Е.В., Миронкина Ю.Н., Трошин Л.И., Московский финансово-промышленный университет "Синергия")- ЭБС IPRbooks.

10.3 Программное обеспечение и Интернет-ресурсы:

Для работы в сети рекомендуется использовать сайты:
- http://encycl.yandex.ru (Энциклопедии и словари).
- http://www.intuit.ru/department/mathematics/intmath/ (Вводный курс в высшую математику. Рассматриваются основы высшей математики для «нематематических» специальностей. Изложение сопровождается большим количеством специально подобранных примеров, поясняющих суть исследуемых понятий и фактов).
- http://mathelp.spb.ru (Лекции, учебники on-line, web-сервисы по высшей математике в помощь студентам).
- http://mathem.by.ru (Справочная информация по математическим дисциплинам).
- http://www.exponenta.ru (Материалы по высшей математике).
- http://teorver-online.narod.ru/teorver73.html (Манита А. Д. Теория вероятностей и математическая статистика. Интернет-учебник).

http://eqworld.ipmnet.ru/ru/library/mathematics.htm (Электронные учебники)

Для работы с электронными учебниками требуется наличие таких программных средств, как Adobe Reader для Windows и DjVuBrowserPlugin.

11. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Для проведения ряда лекционных занятий по дисциплине необходимы аудитории, оснащенные презентационным оборудованием (компьютер с ОС Windows и программой PowerPoint, мультимедийный проектор и экран).

Для обеспечения практических занятий требуется компьютерный класс с комплектом лицензионного программного обеспечения (при использовании электронных изданий – компьютерный класс с выходом в Интернет).

12. МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ ИЗУЧЕНИЯ ДИСЦИПЛИНЫ (образовательные технологии)

Для более эффективного усвоения курса «Теория вероятностей и математическая статистика» рекомендуется использовать на лекциях и практических занятиях видеоматериалы, обобщающие таблицы и др.

№	Темы учебных занятий, проводимых в интерактивных формах	Объем занятий
1.	*Лекции с элементами проблемного обучения* с использованием ПК, мультимедиапроектора и комплекта презентаций по темам: «Схема Бернулли», «Основные законы распределения случайных величин», «Корреляционный и регрессионный анализ».	8
2.	*Лекции – учебные дискуссии* (с использованием рабочих тетрадей, содержащих опорные конспекты изучаемых тем и пропущенные смысловые места для заметок, поправок, примеров) по темам «Основные формулы и правила комбинаторики», «Схема Бернулли»	4
3.	*Практические занятия* (*с элементами компьютерных симуляций и дидактических игр*) в компьютерном классе с использованием программного комплекса Maple для выполнения профессионально ориентированных (индивидуальных) заданий, связанных с расчетами, по теме: «Вычисление числовых характеристик случайных величин».	4
	Всего, час / удельный вес, %	16 / 21

Для повышения интереса к дисциплине и развития математической культуры целесообразно сообщать на лекциях сведения из истории теории вероятностей и информацию о вкладе российских ученых в этот раздел математической науки.

Важным условием успешного освоения дисциплины «Теория вероятностей и математическая статистика» является самостоятельная работа студентов. Для осуществления индивидуального подхода к студентам и создания условий ритмичности учебного процесса рекомендуются индивидуальные расчетно-графические работы (РГР) в группах, коллоквиумы и контрольные работы (КР). Коллоквиум и контрольная работа являются не только формами промежуточного контроля, но и формами обучения, так как позволяют своевременно определить уровень усвоения студентами разделов программы и провести дополнительную работу.

Программа составлена в соответствии с требованиями ФГОС ВО с учетом рекомендаций и ПрОПОП ВО по направлению подготовки _ 04.03.02 Химия, физика и механика материалов

Руководитель основной
образовательной программы
доцент кафедры химии,
к.х.н., доц. О.В. Артамонова
(занимаемая должность, ученая степень и звание) (подпись) (инициалы, фамилия)

Рабочая программа одобрена учебно-методической комиссией строительно-технологического института

«_____»____________201 г., протокол № ________.

Председатель д.т.н., доцент___________________ ____Г.С. Славчева
учёная степень и звание, подпись инициалы, фамилия

Эксперт

_____________________ ____________________ ___________________________
(место работы) (занимаемая должность) (подпись) (инициалы, фамилия)

М П организации

Конспект лекций

Теория вероятностей и математическая статистика

Предмет теории вероятностей

Теория вероятностей – математическая наука, изучающая закономерности случайных явлений. В природе, технике и экономике нет явлений, в которых не присутствовал бы элемент случайности. Разработкой методов изучения закономерностей, наблюдаемых в случайных явлениях, занимается теория вероятностей.

Первые работы по теории вероятностей появились в XVI – XVII вв., они принадлежали Б. Паскалю, П. Ферма, Х. Гюйгенсу и др. при попытке создания теории азартных игр. Далее в XVIII в. Я. Бернулли доказал теорему, названную законом больших чисел, которая обосновала накопленные ранее факты. В XVIII – XIX вв. теория вероятностей была развита А. Муавром, П. Лапласом, К. Гауссом, С. Пуассоном и др. Русские математики П.Л. Чебышев, А.М. Ляпунов и А.А. Марков сделали большой вклад в «математику случайного» в XIX и в начале XX в. В XX в. Основной вклад в последующее развитие теории вероятностей внесли российские математики А.Н. Колмогоров, С.Н. Бернштейн, В.М. Романовский и др., а также ученые англо-американской школы Р. Фишер, Э. Пирсон, Е. Нейман, Стьюдент (В. Госсет) и др. Широкому внедрению приложений вероятностных методов (статистических методов) способствовало появление во второй половине XX в. ЭВМ, в частности, персональных компьютеров.

Основные понятия и определения

Всякий факт, который может наблюдаться при наличии некоторых условий, будем называть *событием*. Условия, при наличии которых может произойти событие, будем называть *опытом* или *испытанием*. Например, подбрасывание монеты – это опыт, а появление герба в результате подбрасывания – это событие; сдача экзамена – это опыт, а получение пятерки – это событие и т.д.

Замечание. В дальнейшем, говоря о событии, будем подразумевать, что речь идет о конкретном опыте.

Определение 1. Событие называют случайным, если в результате опыта оно может появиться или не появиться.

Определение 2. Событие называют невозможным, если в результате опыта оно не может произойти.

Определение 3. Событие называют достоверным, если в результате опыта оно обязательно произойдет.

Например, при подбрасывании игральной кости выпадение пяти очков – случайное событие, выпадение нуля очков – невозможное событие, а выпадение не более шести очков – достоверное событие.

Случайные события обозначают заглавными буквами латинского алфавита. Например, A – выпадение герба при броске монеты, B – попадание в десятку при выстреле в мишень и т.д. Достоверное событие будем обозначать буквой U, невозможное – буквой V.

Определение 4. События A и B называются несовместными, если появление одного из них в данном опыте исключает появление другого. В противном случае A и B называют совместными.

Пример. Из колоды карт извлекается наудачу одна карта. Пусть A – это карта-картинка, B – это карта бубновой масти, C – это карта черной масти. Тогда A и B – совместные события, так как карта может быть бубновой картинкой, а B и C – несовместные события.

Определение 5. События A1,… , An называют полной группой событий, если в результате опыта появляется хотя бы одно из этих событий.

Заметим, что если события, образующие полную группу, попарно несовместны, то в результате испытания появится одно и только одно из этих событий.

Например, при сдаче экзамена события: A – получение пятерки, B – четверки, C – тройки, D – двойки образуют полную группу событий.

Определение 6. Два события называют противоположными, если они несовместны и образуют полную группу. Если одно из них обозначено A , то противоположное ему обозначают $\bar{A}$.

Например, сдача экзамена и получение двойки, попадание и промах при выстреле в мишень и т.д.

В дальнейшем любое событие, которое может появиться в результате опыта, будем называть исходом опыта. Если нельзя отдать предпочтение ни одному из исходов в смысле возможности его появления, то исходы называют равновозможными.

Например, при подбрасывании игрального кубика выпадение одного, двух, …, шести очков – равновозможные исходы, а попадание и промах при броске в корзину мяча баскетболистом, вообще говоря, не равновозможные исходы (попадание или промах зависит от тренированности баскетболиста).

Возможность появления того или иного события характеризуется числом, называемым вероятностью этого события. По мере развития теории вероятностей предлагались разные способы
вычисления вероятности случайного события. Единый теоретико-множественный подход к определению вероятности был предложен в 1933 г. советским академиком А.Н. Колмогоровым. Однако при решении несложных задач достаточно использовать более простые определения вероятности.

Классическое определение вероятности

Рассмотрим опыт, в результате которого может появиться событие A. Пусть известно, что этот опыт имеет n равновозможных, несовместных исходов, образующих полную группу. Исход, при котором появляется событие A, будем называть благоприятствующим этому событию. Пусть количество исходов, благоприятствующих событию A, равно m. В этом случае вероятность события A (обозначается $P(A)$) равна отношению числа исходов, благоприятствующих событию A, к числу всех исходов опыта

$$P(A)=\frac{m}{n}. \qquad (1.1)$$

Пример 1. Вычислить вероятность выпадения шести очков и вероятность выпадения четного числа очков при броске игральной кости. Очевидно, n=6, так как у кости 6 граней и появление каждой из них равновозможное. Обозначим через A появление 6 очков, через B – появление четного числа очков. Тогда m=1 для события A, а m=3 для события B. Поэтому

$$P(A) = \frac{1}{6},\ P(B) = \frac{3}{6} = \frac{1}{2}.$$

Пример 2. В урне 4 белых и 6 красных шаров. Наудачу извлекается один шар. Найти вероятность того, что шар белый (красный). Очевидно, n=10. Если A – шар белый, а B– шар красный, то m=4 для события A, а m=6 для события B. Поэтому

$$P(A) = \frac{4}{10} = \frac{2}{5},\ P(B) = \frac{6}{10} = \frac{3}{5}.$$

Из классического определения вероятности видны ее свойства.

Свойства вероятности

1. Вероятность достоверного события равна 1 (n=m), т.е. $P(U)$=1.

2. Вероятность невозможного события равна 0 (m=0), т.е. $P(V)$=0.

3. Для любого события $0 \le P(A) \le 1$.

Замечание. При использовании классического определения вероятности нахождение значений m и n является достаточно трудным даже в простых ситуациях. Например, число исходов опыта при вытаскивании произвольных пяти карт из колоды. Здесь число равновозможных исходов – это число различных комбинаций из пяти карт, которые можно составить из данной колоды. Подсчитать такое число комбинаций позволяют методы *комбинаторики*.

Основные понятия комбинаторики

Пусть имеется *n* пронумерованных элементов.

1. *Перестановками* из *n* элементов называют комбинации всех *n* элементов, отличающиеся только порядком элементов.

Число возможных перестановок из *n* элементов обозначают *Pn* . Можно доказать, что

$$P_n = n!. \qquad (1.2)$$

Пример. Пусть элементы – это три цифры: 1,2,3. Тогда перестановки из этих цифр – трехзначные числа: 123, 132, 213, 231, 312 и 321 (всего 6).

Этот результат можно получить, используя формулу (1.2): в рассмотренном примере *n*=3, $P_n = 3! = 6$.

2. *Размещения* из *n* элементов по *m* элементов – это комбинации, состоящие из *m* элементов, отличающиеся либо элементами, либо их порядком. Число возможных размещений из *n* элементов по *m* элементов обозначают A_n^m. Справедлива формула

$$A_n^m = \frac{n!}{(n-m)!} = n(n-1)...(n-m+1). \qquad (1.3)$$

Пример. Пусть элементы – это снова три цифры: 1, 2, 3. Тогда размещения по два элемента из этих трех – это двузначные числа 12, 13, 21, 23, 31, 32 (всего 6). Формула (1.3) приводит к такому же результату:

$$A_n^m = \frac{3!}{(3-2)!} = 6.$$

3. *Сочетаниями* из n элементов по m элементов называют комбинации, состоящие из m элементов, отличающиеся хотя бы одним элементом (т.е. порядок элементов здесь не важен).

Пример. Снова рассмотрим три цифры: 1, 2, 3. Сочетания из этих трех элементов по два – это комбинации 12, 13, 23 (комбинация 21, например, – это то же сочетание, что и 12).

Число возможных сочетаний из n элементов по m элементов обозначают C_n^m. Справедливо соотношение $A_n^m = C_n^m \cdot P_m$, поэтому

$$C_n^m = \frac{A_n^m}{P_m} = \frac{n!}{m!(n-m)!}. \qquad (1.4)$$

В рассмотренном выше примере $C_3^2 = \dfrac{3!}{2!(3-2)!} = 3$.

Геометрическое определение вероятности

Классическое определение вероятности применимо, когда число исходов опыта конечно. Если опыт имеет бесконечное множество несовместных, равновозможных исходов и эти исходы можно сопоставить с точками области G, то вероятность события A определяется формулой

$$P(A) = \frac{\mu(g)}{\mu(G)}. \qquad (1.5)$$

Здесь μ – это мера (длина, площадь или объем) области; g – это область, соответствующая исходам, благоприятствующим A, и содержащаяся в G.

Для иллюстрации случайных событий удобно изображение множества исходов опыта в виде области на плоскости. Например, хорошо иллюстрируются совместные и несовместные события.
На рис. 1 изображены множества исходов $g(A)$ и $g(B)$, благоприятствующие несовместным событиям A и B, а на рис. 2 – совместным.

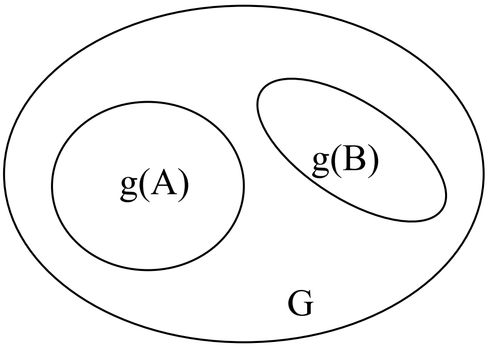


Рис. 1

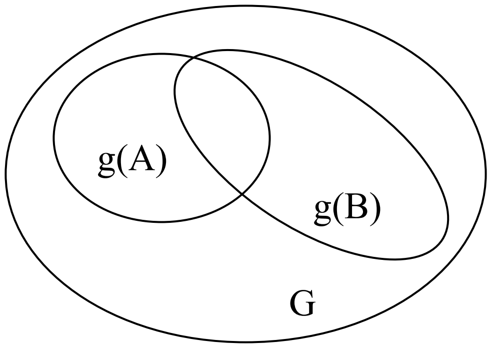


Рис. 2

Рассмотрим на примере использование геометрического определения вероятности.

Пример. Известно, что автобус № 5 отправляется с остановки строго по расписанию каждые 20 мин. Пассажир, не знающий расписания, приходит на остановку. Какова вероятность того, что он будет ждать автобус не более 2 мин.
В этой задаче можно считать, что приход человека на остановку (испытание) будет равновозможным в любой момент времени 20-минутного интервала между приходами автобуса. Следовательно, каждый исход опыта сопоставим с точкой отрезка G=[0, 20] на числовой оси. Пусть событие A – приход на остановку не более чем за 2 мин до появления автобуса. Исходы, благоприятствующие событию A, соответствуют точкам отрезка g=[18, 20] на числовой оси. Вероятность события A равна отношению длины отрезка g=[18, 20] к длине отрезка G=[0, 20], т.е.

$$P(A) = \frac{20-18}{20-0} = \frac{2}{20} = 0{,}1.$$

Классическое и геометрическое определения вероятности применимы лишь в случае, когда событие может появиться в опыте, имеющем равновозможные исходы. Но достаточно часто это не выполнено. Например, нужно определить вероятность попадания в мишень при одном выстреле или вероятность рождения мальчика, или вероятность изготовления станком-автоматом стандартной детали и т.д. В этих примерах опыты – выстрел в мишень, рождение ребенка, изготовление детали. Очевидно, исходы в этих опытах нельзя считать равновозможными. Поэтомуклассическое и геометрическое определения вероятности здесь неприменимы.

Статистическое определение вероятности

Пусть событие A появляется в результате опыта, который можно повторять сколь угодно большое количество раз. Предположим, что проведено N опытов и событие A появилось в M опытах. *Относительной частотой* события A в проведенных опытах называется число $W(A)$, равное отношению числа появлений события A к числу проведенных опытов, т.е.

$$W(A) = \frac{M}{N}.$$

Например, произведено 10 выстрелов в мишень, из которых промахов – 2. Тогда относительная частота промахов в этой серии будет W=0,2.

Очевидно, в разных сериях опытов значения относительной частоты, как правило, различны. Однако при многократном повторении серий опытов с увеличением числа опытов в сериях значения относительной частоты $W(A)$, меняясь, сгущаются около некоторого значения. Таким образом, можно дать следующее *статистическое определение вероятности* события A: вероятностью события называется число, около которого группируются значения относительной частоты (или предельное значение относительной частоты).

Для подтверждения правомочности такого определения математики рассматривали в качестве A – появление герба при бросании монеты ($P(A)$=0,5 по классическому определению вероятности). Они провели множество серий бросков и получили результаты, приведенные в табл. 1.

Таблица 1

Автор	N	M	$W(A)=M/N$
Бюффон	4040	2048	0,5069
Феллер	10000	4979	0,4979
Джевонс	12000	6019	0,5016
Пирсон	24000	12012	0,5005

Недостаток статистического определения вероятности – в нестрогой формулировке и в необходимости проведения большого количества опытов. Поэтому им пользуются в случаях, когда изучаются события, наблюдаемые в многократно повторяемых опытах.

Например, рождение мальчика (A). Для определения вероятности события A можно использовать статистические данные о рождении детей за большой период времени.

Замечание. Легко видеть, что вероятность события, найденная с помощью геометрического или статистического определения, обладает теми же свойствами, которые были отмечены ранее для классического определения.

Вычисление вероятности события – достаточно сложная задача. поэтому естественно научиться искать вероятности событий, связанных с событиями, вероятности которых известны.

Операции над случайными событиями

Определение 1. Суммой событий A и B называется событие C, обозначаемое $C=A+B$, представляющее собой появление хотя бы одного из событий A, B.

Изобразим исходы опыта и исходы, благоприятствующие событиям A и B, соответствующими областями (рис. 3).

а) б)

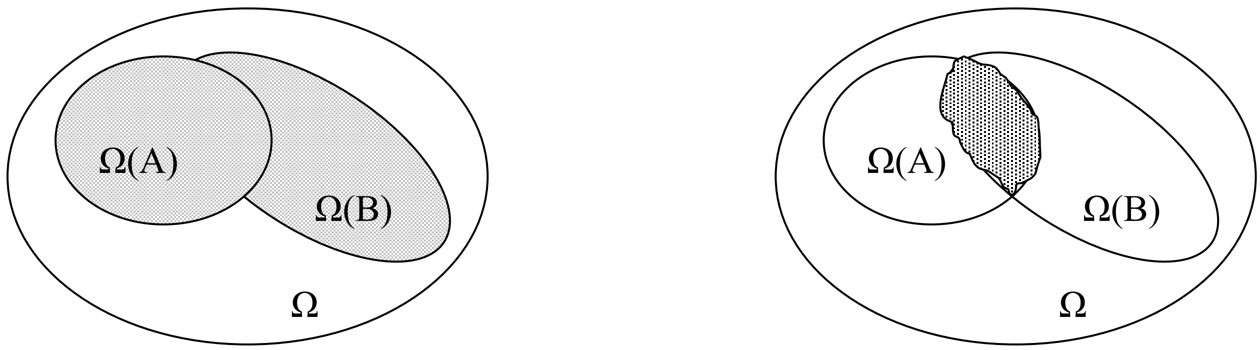


Рис. 3

На приведенном рисунке область Ω – исходы опыта; $\Omega(A)$ – исходы, благоприятствующие A; $\Omega(B)$ – исходы, благоприятствующие B; $\Omega(A+B)= =\Omega(A)\cup\Omega(B)$ – исходы , благоприятствующие $A+B$ (рис. 3,*а*).

Определение 2. Произведением событий A и B называют событие C, обозначаемое $C=AB$, представляющее собой появление обоих событий A и B. На рис. 3,*б* множество исходов, благоприятствующих AB, изображено областью $\Omega(AB)= =\Omega(A)\cap\Omega(B)$.

Пример. Два стрелка стреляют в одну мишень. Пусть событие A – поражение мишени первым стрелком, событие B – вторым стрелком. Тогда событие $A+B$ – мишень поражена (т.е. хотя бы один стрелок попал), событие AB – мишень поражена двумя стрелками.

Определение 3. События A и B называют независимыми, если вероятность любого из них не зависит от того, произошло другое событие или нет. В противном случае события A и B называют зависимыми.

Пример 1. Бросают две монеты. Очевидно, появление герба на каждой из монет не зависит от того, что выпадет на другой. Поэтому события $A1$ и $A2$ – появление герба на первой и второй монетах – независимы и

$$P(A_1)=\frac{1}{2},\ P(A_2)=\frac{1}{2}.$$

Пример 2. Из урны с пятью синими и десятью красными шарами извлекают последовательно два шара. Событие $A1$ – первый шар красный, $A2$ – второй шар красный. Очевидно, $A1$ и $A2$ – зависимые события, так как при появлении события $A1$ вероятность P($A2$)=9/14, а при непоявлении $A1$ вероятность $P(A2)$=10/14.

Определение 4. Вероятность события *A*, вычисленная в предположении, что произошло событие *B*, обозначают $P(A/B)$ и называют условной вероятностью события *A*. Вероятность $P(A)$ называют безусловной вероятностью события *A*.

Для независимых событий *A* и *B* условная вероятность $P(A/B)$=P(A), а для зависимых $P(A/B) \neq P(A)$.

Пример. Бросают две кости. Пусть *A* – выпадение 5 очков на первой кости, *B* – выпадение 10 очков на двух костях в сумме. Тогда $P(B)$=1/12 (n=36, m=3), $P(B/A)$=1/6 (здесь *A* и *B* – зависимые события).

Покажем, как определяются условная и безусловная вероятности геометрически, используя рис.4. Пусть Ω – множество исходов опыта, ΩA – исходы, благоприятствующие *A*, ΩB – исходы, благоприятствующие *B*.

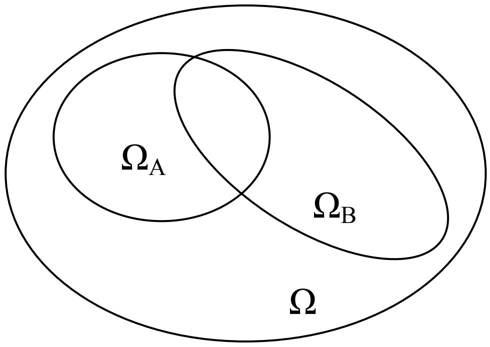


Рис. 4

Согласно геометрическому определению вероятности

$$P(A)=\frac{S(\Omega_A)}{S(\Omega)},\ P(A/B)=\frac{S(\Omega_{AB})}{S(\Omega_B)},\ P(B/A)=\frac{S(\Omega_{AB})}{S(\Omega_A)}.$$

где S обозначает площадь фигур.

Эти формулы используют для доказательства следующей теоремы.

Теорема умножения вероятностей

Вероятность произведения двух событий А и В равна произведению вероятности события А и вероятности события В при условии, что событие А произошло:

$$P(AB)=P(A)P(B/A). \qquad (2.1)$$

Для доказательства воспользуемся рис. 4 и геометрическим определением вероятности:

$$P(AB)=\frac{S(\Omega_{AB})}{S(\Omega)}=\frac{S(\Omega_A)}{S(\Omega)}\frac{S(\Omega_{AB})}{S(\Omega_A)}=P(A)P(B/A).$$

Заметим, что из очевидного равенства $AB=BA$ формула (2.1) может записываться и в виде $P(AB)=P(B)\,P(A/B)$.

Следствие. Вероятность произведения независимых событий равна произведению их вероятностей.

Действительно, если A и B независимы, то $P(B/A)=P(B)$, поэтому формула (2.1) принимает вид

$$P(AB)=P(A)\cdot P(B)\ . \qquad (2.2)$$

Пример 1. Два стрелка независимо друг от друга стреляют в мишень. Вероятность попадания для одного стрелка равна 0,7, для второго – 0,6. Найти вероятность того, что оба попадут в мишень.

Пусть A – попадание в мишень первым стрелком, B – попадание в мишень вторым стрелком, тогда $P(A)=0{,}7$; $P(B)=0{,}6$. Попадание в мишень обоими стрелками – AB. Так как A и B независимы, то по формуле (2.2)

$$P(AB)=P(A)P(B)=0{,}7\cdot 0{,}6=0{,}42.$$

Пример 2. Из урны, в которой три синих и пять красных шаров, последовательно извлекают два шара. Какова вероятность того, что оба шара красные. Пусть событие A – извлечение красного шара первым, а B – извлечение красного шара вторым. Тогда искомым событием будет AB. По формуле (2.1) $P(AB)=P(A)P(B/A)$. Так как $P(A)=5/8$, $P(B/A)=4/7$, то $P(AB)=5/14$.

Теорема сложения вероятностей

Вероятность суммы двух событий равна сумме вероятностей этих событий без вероятности их произведения:

$$P(A+B)=P(A)+P(B)-P(AB)\,. \qquad (2.3)$$

Доказательство . Вернемся к рис. 4. Исходам, благоприятствующим событию $A+B$, соответствует область $\Omega A \cup \Omega B$. Поэтому

$$P(A+B)=\frac{S(\Omega_A \bigcup \Omega_B)}{S(\Omega)}=\frac{S(\Omega_A)+S(\Omega_B)-S(\Omega_{AB})}{S(\Omega)}=\frac{S(\Omega_A)}{S(\Omega)}+\frac{S(\Omega_B)}{S(\Omega)}-\frac{S(\Omega_{AB})}{S(\Omega)}$$

Учитывая, что

$$P(A)=\frac{S(\Omega_A)}{S(\Omega)},\ P(B)=\frac{S(\Omega_B)}{S(\Omega)},\ P(AB)=\frac{S(\Omega_{AB})}{S(\Omega)}.$$

получаем указанную ранее формулу (2.3). Теорема доказана.

Следствие 1 . Если события A и B несовместны, то $AB=V$, поэтому $P(AB)=0$ и формула (2.3) примет вид

$$P(A+B)=P(A)+P(B)\ . \qquad (2.4)$$

Следствие 2 . Для любого события A

$$P(\bar{A})=1-P(A)\ . \qquad (2.5)$$

Это следует из того, что A и $\bar{A}$ – несовместные события и $A+\bar{A}=U$, откуда $P(A+\bar{A})=P(U)=1$ и $P(A+\bar{A})=P(A)+P(\bar{A})$. Значит, $P(A)+P(\bar{A})=1$, а отсюда получаем $P(\bar{A})=1-P(A)$.

Пример 1. Два стрелка, независимо друг от друга, стреляют в мишень. Вероятность поражения мишени для первого стрелка – 0,5, для второго – 0,4. Какова вероятность поражения мишени с первого залпа?

Обозначим $A1$ и $A2$ – поражение мишени первым и вторым стрелком соответственно, тогда $P(A1)=0{,}5$, а $P(A2)=0{,}4$. Поражение мишени залпом – это сумма событий $A1$ и $A2$. Очевидно, $A1$ и $A2$ – независимые события. По формулам (2.2) и (2.3) имеем

$$P(A1+A2)=P(A1)+P(A2)-P(A1A2)=P(A1)+P(A2)-P(A1)P(A2)=$$
$$=0{,}5+0{,}4-0{,}5\bullet 0{,}4=0{,}7.$$

Замечание. Если рассматриваются три или большее число событий, то при их сложении и умножении справедлив сочетательный закон. Так, например, для трех событий A, B, C

$A+B+C=(A+B)+C=A+(B+C)$, $ABC=(AB)C=A(BC)$.

Исходя из этого, можно найти вероятность произведения трех событий:

$$P(ABC)=P(A)P(BC/A)=P(A)P(B/A)P(C/AB).$$

Для вероятности суммы трех событий формула получается существенно более сложной:

$$P(A+B+C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC).$$

Этой формулой на практике не пользуются, находят сначала вероятность события, противоположного $A+B+C$, а затем используют формулу (2.5).

Действительно, так как $A+B+C$ – появление хотя бы одного из событий A, B, C, то $\overline{A+B+C}$ – не появление ни одного из них. Не появление каждого из событий A, B, C – это события $\overline{A}$, $\overline{B}$, $\overline{C}$, поэтому $\overline{A+B+C}=\overline{A}\overline{B}\overline{C}$. Отсюда $P(A+B+C)=1-P(\overline{A+B+C})=1-P(\overline{A}\,\overline{B}\,\overline{C})$.

Очевидно, для n событий $A1$, …, An, вероятности которых равны соответственно $p1$, …, pn, вероятности $\bar{A}1$, …, $\bar{A}n$ будут $q1=1-p1$, …, $qn=1-pn$. Поэтому для попарно независимых событий $A1$, …,An справедлива формула

$$P(A1+\ldots+An)=1-q1q2\cdots qn\,. \qquad (2.6)$$

Пример. Три стрелка попадают в мишень при одном выстреле с вероятностями 0,6; 0,7; 0,4 соответственно. Какова вероятность поражения мишени при залпе этих стрелков?

Обозначим Ai – попадание в мишень i-м стрелком (i=1,2,3). Тогда событие $A1+A2+A3$ – поражение мишени при залпе. События $A1$, $A2$, $A3$ – попарно независимы, и $q1=1-0{,}6=0{,}4$; $q2=1-0{,}7=0{,}3$; $q3=1-0{,}4=0{,}6$. По формуле (2.6)

$P(A1+A2+A3)=1-0,4\cdot 0,3\cdot 0,6=0,928.$

Из сказанного выше следует, что использование теорем сложения и умножения вероятностей наиболее просто, когда события несовместны (для суммы) и независимы (для произведения). Поэтому при решении задач на отыскание вероятности события это событие стараются представить в виде суммы несовместных событий, каждое из которых равно произведению независимых событий.

Пример. Три студента сдают экзамен. Вероятность получить "отлично" для первого студента – 0,8; для второго – 0,7; для третьего – 0,9. Какова вероятность того, что не менее двух студентов получат отличные оценки?

Обозначим через Ai – получение оценки "отлично" i-м студентом, тогда $P(A1)=0,8$; $P(A2)=0,7$; $P(A3)=0,9$. Искомое событие обозначим B. Очевидно, B произойдет, когда либо все три студента получат пятерки, либо два получат, а один нет:

$$B= A1A2A3 +\bar{A}1A2A3 +A1\bar{A}2A3 +A1A2\bar{A}3.$$

Все слагаемые события здесь несовместны, а в каждом слагаемом события независимы, поэтому

$$P(B) = P(A1)P(A2)P(A3)+P(\bar{A}1)P(A2)P(A3)+P(A1)P(\bar{A}2)P(A3)+P(A1)P(A2)P(\bar{A}3).$$

Учитывая, что $P(\bar{A}1)=1-P(A1)=0,2$; $P(\bar{A}2)=1-P(A2)=0,3$; $P(\bar{A}3)=1-P(A3)=0,1$, получим

$$P(B) = 0,8\cdot 0,7\cdot 0,9+0,2\cdot 0,7\cdot 0,9+0,8\cdot 0,3\cdot 0,9=0,902.$$

Рассмотрим задачу, в которой требуется найти вероятность события A, которое наступает вместе с одним из событий $H1, \ldots, H$n, образующих полную группу несовместных событий. Очевидно,

$$A=A\cdot H1+A\cdot H2 +\ldots+ A\cdot Hn ,$$

поэтому, используя формулы (2.1) и (2.4), будем иметь

$$P(A)=P(H1)P\ (A/H1)+P(H2)P(A/H2)+\ldots+P(Hn)P(A/Hn)\ . \qquad (2.7)$$

Формула (2.7) называется формулой полной вероятности.

Пример. Два станка-автомата производят одинаковые детали, причем производительность первого в 2 раза больше, чем второго. Известно, что первый станок дает 3 % брака, а второй – 2 % брака. Для проверки наудачу берется деталь. Какова вероятность того, что деталь стандартна?

Пусть событие A – деталь стандартна. Это событие может произойти, когда деталь изготовлена либо на первом станке, либо на втором, т.е. при одной из гипотез Hi – деталь изготовлена на i-м станке (i=1, 2). Воспользуемся формулой (2.7) при n=2:

$$P(A)= P(H1)P(A/H1) + P(H2)P(A/H2).$$

Учитывая производительность станков, имеем $P(H1)$=2/3, $P(H2)$=1/3. Процент брака на каждом из станков – это информация об условных вероятностях, позволяющая найти $P(A/H1)$=0,97; $P(A/H2)$=0,98. По формуле полной вероятности получаем

$$P(A)=2/3\cdot 0{,}97+1/3\cdot 0{,}98=0{,}973.$$

Схема Бернулли

В разделе теории вероятностей, изучающем случайные события, принято отдельно рассматривать опыт, состоящий в повторении одинаковых испытаний.

Пусть событие A может произойти в некотором испытании с вероятностью p и не произойти с вероятностью q (q=1–p). Независимо друг от друга проводится n таких испытаний. Эту ситуацию называют схемой Бернулли. В схеме Бернулли ставятся два основных вопроса: какова вероятность появления события A m раз в n испытаниях $Pn(m)$ и какова вероятность появления события A не менее $m1$ раз, но не более $m2$ раз в данной схеме Бернулли? Сначала вычислим $Pn(m)$.

Формула Бернулли

Обозначим через B появление A в n испытаниях m раз. Очевидно, B наступит, если в результате n испытаний произойдет n событий, из которых m событий – события A, а (n-m) событий – события $\bar{A}$, иными словами произойдет произведение событий $A\cdots A\cdot\bar{A}\cdots\bar{A}$. Так как испытания проводятся независимо друг от друга, то

$P(A\cdots A\cdot\bar{A}\cdots\bar{A})=pm \cdot qn\text{-}m$.

В результате n испытаний такие произведения событий могут наступать разными способами, так как порядок появления A и $\bar{A}$ в испытаниях могут быть разными, например, при n=3 и m=2 событие B наступит, если появится либо $AA\ \bar{A}$, либо $A\bar{A}A$, либо $\bar{A}AA$. Очевидно, событие B равно сумме всех таких возможных произведений. Каждое произведение соответствует сочетанию порядковых номеров испытаний, в которых появилось событие A. Поэтому число слагаемых равно . Учитывая, что слагаемые – это несовместные события и их вероятности одинаковы, получаем формулу

$$P_n(m) = C_n^m p^m q^{n-m} \qquad (3.1)$$

Формула (3.1) называется формулой Бернулли.

Пример 1. Вероятность того, что станок в течение часа потребует внимания рабочего, равна 0,6. Найти вероятность того, что потребует внимания один станок из четырех обслуживаемых.

Воспользуемся формулой (3.1) при n=4, m=1, p=0,6, q=1–0,6=0,4 :

$$P_4(1) = C_4^1 0{,}6^1 0{,}4^3 = 4 \cdot 0{,}6 \cdot 0{,}4^3 = 0{,}1536.$$

Формула Пуассона

Когда число испытаний велико, вычисление затруднительно (или невозможно), поэтому формулу Бернулли заменяют приближенной формулой. Если n велико и $p \approx 0$ (p<0,1), используют формулу Пуассона

$$P_n(m) = \frac{\lambda^m e^{-\lambda}}{m!}, \qquad (3.2)$$

где $\lambda = np < 10$.

Пример 2. Завод отправил на базу 500 деталей. Вероятность повреждения детали в пути равна 0,002. Какова вероятность повреждения в пути трех деталей?

Здесь применима схема Бернулли с n=500, p=0,002. Так как $p \approx 0$ и n велико, то для отыскания $P500(3)$ применим формулу (3.2) ($\lambda = np = 500 \cdot 0{,}002 = 1$):

$$P_{500}(3) \approx \frac{1^2 e^{-1}}{3!} = \frac{1}{6e} \approx 0{,}0613.$$

Локальная формула Лапласа

Если n велико и p не близко к нулю или единице, то используют локальную формулу Лапласа (Муавра–Лапласа):

$$P_n(m) \approx \frac{1}{\sqrt{npq}} \cdot \varphi(x), \qquad (3.3)$$

где $x = \dfrac{m-np}{\sqrt{npq}}$, $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Значения затабулированы для $x \in [0,5]$. Для $x \notin [0,5]$ значения $\phi(x)$ определяют, используя ее свойства: четность и асимптотику $\lim\limits_{x\to\pm\infty} \phi(x) = 0$. При x >5 считают $\varphi(x)$ =0 (уже $\varphi(5)$ =0,0000015). График $\varphi(x)$ называют кривой Гаусса (рис. 5).

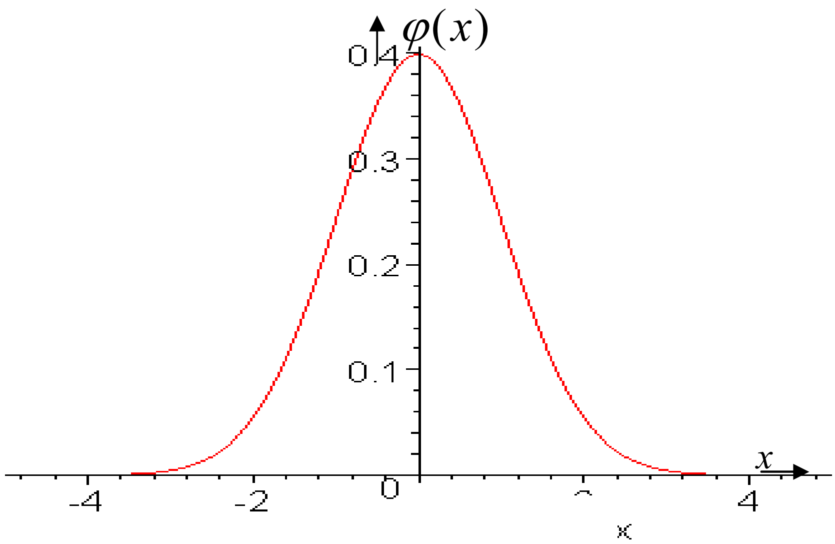


Рис. **5**

Пример 3. Найти вероятность поражения мишени 250 раз при 600 выстрелах, если вероятность попадания при одном выстреле равна 0,4.

Имеем схему Бернулли с n=600, p=0,4 и m=250. Искомая вероятность равна P600(250). Применим локальную формулу Муавра-Лапласа. Найдем сначала

$$x = \frac{250 - 600 \cdot 0{,}4}{\sqrt{600 \cdot 0{,}4 \cdot 0{,}6}} = \frac{10}{12} = 0{,}833.$$

Затем из приложения 1 можно найти $\phi(0{,}833) = 0{,}282$ и, наконец, искомую вероятность

$$P_{600}(250) \approx \frac{1}{12} \cdot 0{,}282 = 0{,}0235.$$

Интегральная формула Лапласа

Если в схеме Бернулли ставится вопрос об отыскании вероятности Pn(m1, m2) того, что событие A появится в n испытаниях не менее m1 и не более m2 раз, то при небольших n используют равенство

$Pn(m1, m2) = Pn(m1) + Pn(m1+1) + \ldots + Pn(m2)$,

где каждое слагаемое находят по формуле Бернулли.

Пример. Найти вероятность того, что при 5 бросках монеты герб выпадет не более двух раз.

Имеем схему Бернулли с n =5, $p = \frac{1}{2}$ и $q = \frac{1}{2}$. Искомая вероятность $P5(0;2)$, вычисляемая по формуле

$P5(0;2) = P5(0)+P5(1)+P5(2)$,

будет равна

$$P_5(0;2) = C_5^0\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5 + C_5^1\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^4 + C_5^2\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 = \left(\frac{1}{2}\right)^5\left(1+5+10\right) = \frac{16}{32} = \frac{1}{2}.$$

Если же в схеме Бернулли n велико и p не близко к нулю или к единице, то используют приближенную формулу, называемую интегральной формулой Лапласа:

$$P(m_1;m_2) \approx \Phi(x_2) - \Phi(x_1), \qquad (3.4)$$

где

$$x_1 = \frac{m_1 - np}{\sqrt{npq}},\; x_2 = \frac{m_2 - np}{\sqrt{npq}},$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz. \qquad (3.5)$$

Функция (3.5) называется функцией Лапласа, она нечетна и $\lim_{x\to\pm\infty} \Phi(x) = 0{,}5$. Значения ее затабулированы для $x \in [0{,}5]$ (см. приложение 2); для $x>5$ считают $\Phi(x)=0{,}5$ ($\Phi(5)=0{,}499997$); для $x<0$ пользуются нечетностью $\Phi(x)$. График $\Phi(x)$ имеет вид, представленный на рис. 6.

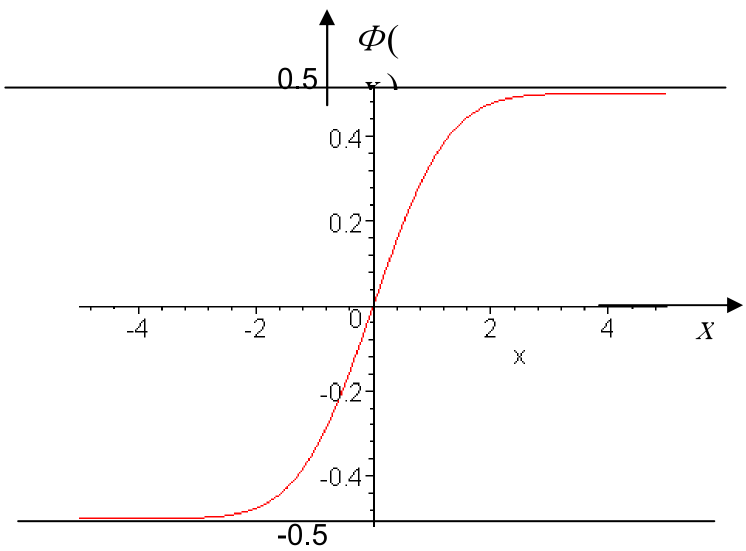


Рис. 6

Случайные величины

После случайного события следующим объектом, рассматриваемым теорией вероятностей, является случайная величина – это величина, которая в одних и тех же условиях может принимать разные значения, причем принятие каждого значения – случайное событие.

Случайные величины будем обозначать X, Y, Z, а их значения x, y, z.

Пример 1. Число очков при бросании игральной кости: $\{1,2,\ldots,6\}$.

Пример 2. Число родившихся мальчиков из 100 новорожденных: $\{0,1,2,\ldots,100\}$.

Пример 3. Число опечаток на странице рукописи: $\{0,1,2\ldots600\ldots\}$.

Пример 4. Время ожидания автобуса на остановке: $\{0 \le t \le 30\text{минут}\}$.

Пример 5. Отклонение точки попадания снаряда от цели: $\{0 \le \delta \le 10\text{ м}\}$.

Пример 6. Время безотказной работы лампы: $\{0 \le \text{T} \le 100\text{ суток}\}$.

Среди случайных величин различают *дискретные* и *непрерывные*.

Если множество значений случайной величины X является конечным или бесконечной последовательностью, то X называют дискретной случайной величиной (примеры 1 – 3). Если же множество значений X – конечный или бесконечный интервал, то X называют непрерывной случайной величиной (примеры 4 – 6).

Учитывая случайность принятых значений, при изучении случайной величины описывают не только множество ее значений, но и вероятности принятия каждого из них или вероятности принятия случайной величиной значений из любого интервала. Такое описание называют законом распределения случайной величины.

Закон распределения дискретной случайной величины

Пусть рассматривается дискретная случайная величина X, имеющая значения $x1, \ldots, x\text{n}, \ldots$. Первый способ задания закона распределения случайной величины X – задание функции

$pk = f(xk)$ $(k=1,2,\dots)$,   (4.1)

где – $pk = P(X = xk)$ вероятность того, что X примет значение xk. Функция (4.1) может быть задана таблицей:

X	x1	x2	...	xn	...
P	p1	p2	...	pn	...

графически – ломаной, соединяющей последовательно на плоскости XOP точки (xk, pk), k=1,2,…(многоугольник распределения, рис. 7),

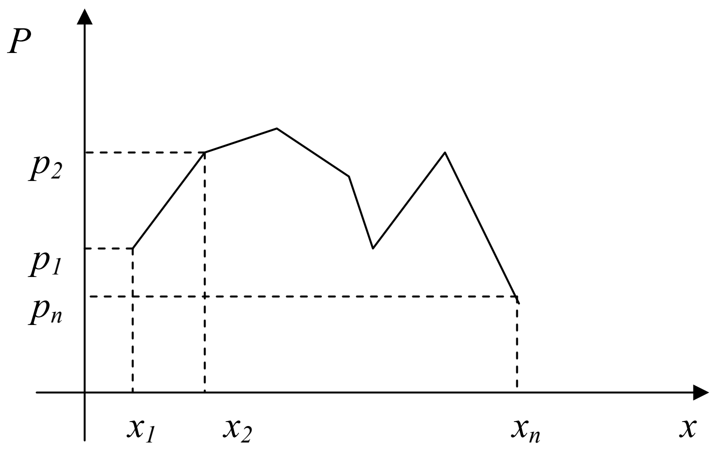


Рис. 7

и аналитически. Так как события X=xk (k=1,2,…) образуют полную группу и попарно несовместны, то, очевидно,

$\Sigma pk = 1.$ (4.2)

<u>Функция распределения дискретной случайной величины</u>

Второй способ задания закона распределения дискретной случайной величины X – задание функции, с помощью которой можно искать вероятность принятия величиной X значений из любого интервала (этот способ в дальнейшем будет распространен на непрерывные случайные величины).

Пусть x – произвольное действительное число. Функцию $F(x)$ назовем функцией распределения случайной величины X, если ее значение определяется формулой

$$F(x) = P(X < x) . \qquad (4.3)$$

Так как для каждого $x \in R$ значение $F(x)$ – это вероятность того, что X примет значение, лежащее левее точки x, то очевидна формула

$$F(x)= \Sigma pk \qquad (4.4)$$

где $pk=P(X=xk)$ и $xk< x$.

Из формул (4.1), (4.2), (4.3) следуют свойства функции распределения.

Свойство 1 . $0 \le F(x) \le 1$.

Свойство 2 . $F(x)$ – неубывающая функция.

Свойство 3 . $F(x)$ – ступенчатая функция, имеющая в точках xk (k=1, 2,…) разрывы 1-го рода со скачками, равными pk.

Пример. Два стрелка делают по одному выстрелу. Вероятности попадания в мишень соответственно равны $p1=0{,}5$, $p2=0{,}4$. Составить закон распределения числа попаданий в цель, начертить многоугольник распределения и график функции распределения.

Пусть A – событие, состоящее в том, что первый стрелок попал в цель; B – событие, состоящее в том, что второй стрелок попал в цель; X – дискретная случайная величина – число попаданий в цель. Вероятности того, что X примет значения 0, 1 и 2, будут равны:

$P(X=0)=P(\overline{A}\,\overline{B})=P(\overline{A})\,P(\overline{B}) = [1-P(B)]\cdot[1-P(A)] = (1-p1)(1-p2) = 0{,}3$;

$P(X=1)=P(A\overline{B} + \overline{A}B)=$

$= P(A)\cdot[1-P(B)]+P(B)\cdot[1-P(A)] = p1(1-p2) + p2(1-p1) = 0{,}5$;

$P(X=2)=P(AB)=P(A)P(B)= p1\,p2 = 0{,}2$.

Таким образом, таблица распределения принимает вид

X	*0*	*1*	*2*	Σ
P	*0,3*	*0,5*	*0,2*	*1*

Далее строим многоугольник распределения (рис. 8):

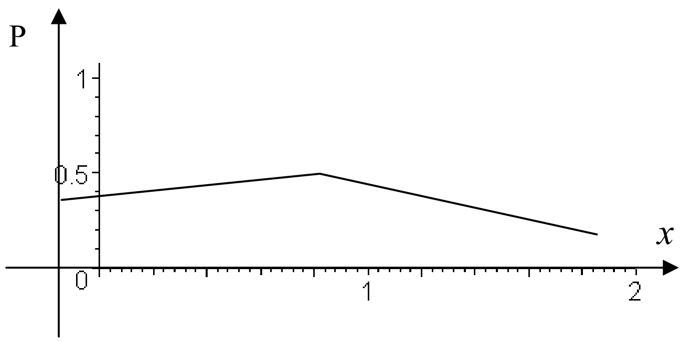


Рис. **8**

Функция распределения имеет вид

$$F(x)=\begin{cases} 0, & x \le 0; \\ 0{,}3, & 0 < x \le 1; \\ 0{,}8, & 1 < x \le 2; \\ 1, & x > 2. \end{cases}$$

ее график представлен на рис. 9.

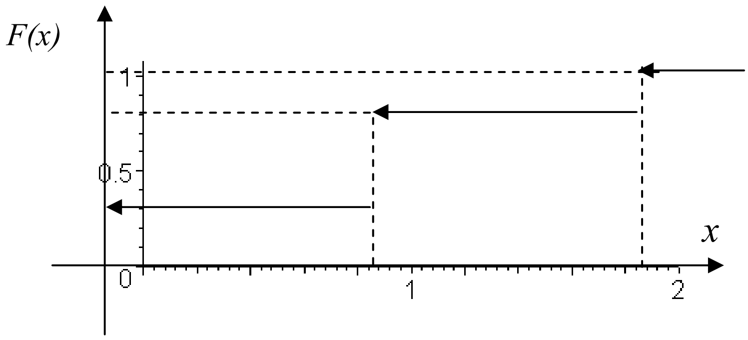


Рис. 9.

Функция распределения непрерывной случайной величины

Пусть X – непрерывная случайная величина. Как отмечалось ранее, ее закон распределения можно задать функцией распределения по формуле (21). К прежнему определению непрерывной случайной величины добавляем еще одно условие – непрерывность функции распределения. Иными словами, случайная величина непрерывна, если ее функция распределения непрерывна на R. Свойства функции распределения непрерывной случайной величины отличаются от свойств $F(x)$ дискретной случайной величины.

Свойства функции распределения непрерывной случайной величины

Свойство 1 . Функции распределения имеет смысл вероятности, поэтому $0 \le F(x) \le 1$.

Свойство 2 . Функции распределения $F(x)$ – неубывающая функция.

Для доказательства возьмем произвольные $x1$ и $x2$, где $x1 < x2$. Рассмотрим три события: $A = \{X < x2\}$, $B = \{X < x1\}$ и $C = \{x1 \le X < x2\}$. Очевидно, что $A = B + C$, а так же B и C несовместны. По теореме сложения вероятностей

$$P(A)=P(B)+P(C) \text{ , т.е. } P(A) - P(B)=P(C) \text{ .}$$

Учитывая, что $P(A) = F(x2)$, $P(B) = F(x1)$, получим $F(x2) - F(x1) = P(C) \geq 0$, что и требовалось доказать.

Свойство 3 . Для любых $\alpha \leq \beta$ справедлива формула

$$P(\alpha \leq X < \beta) = F(\beta) - F(\alpha) . \qquad (4.5)$$

Эта формула получена при доказательстве *свойства* 2 функции $F(x)$ при вычислении $P(C)$.

Свойство 4 . Для любого значения $x = x$о справедлива формула

$$P(X = x0) = 0. \qquad (4.6)$$

Для доказательства заметим, что $P(X = x\text{о}) \geq 0$ и для любого $\Delta x \geq 0$

$$P(X=x0) \leq P(x0 \leq X<x0 + \Delta x).$$

Отсюда, учитывая формулу (4.5), получим

$$0 \leq P(X = x0) \leq F(x0 + \Delta x) - F(x0). \qquad (4.7)$$

Устремляя Δx к нулю, из непрерывности $F(x)$ в точке xо имеем .

$$\lim_{\Delta x \to 0}\left(F(x_0 + \Delta x) - F(x_0)\right) = 0 .$$

По теореме о переходе к пределу в неравенствах из (4.7) получим

$$0 \leq P(X = x0) \leq 0 ,$$

откуда следует (4.6).

Замечание . Учитывая *свойство* 4, можно обобщить формулу (4.5)

$$P(\alpha < X < \beta) = P(\alpha \le X < \beta) = P(\alpha < X \le \beta) = P(\alpha \le X \le \beta) = F(\beta) - F(\alpha). \quad (4.8)$$

Свойство 5 . Если множеством значений величины X является отрезок $[a, b]$, то

$$F(x) = 0 \text{ для } x \le a \text{ и } F(x) = 1 \text{ для } x \ge b . \qquad (4.9)$$

Действительно, для любого $x_0 \le a$ событие $X < x_0$ невозможно по условию, поэтому

$P(X < x0) = F(x0) = 0$.

Аналогично для любого $x1 \ge b$ событие $X < x1$ достоверное, поэтому

$P(X < x1) = F(x1) = 1$.

Свойство 6 . Если множеством значений величины X является вся числовая ось, то аналогом *свойства 5* будут равенства

$$\lim_{x \to -\infty} F(x) = 0, \ \lim_{x \to +\infty} F(x) = 1. \qquad (4.10)$$

Из свойств функции распределения можно сделать вывод: функция распределения случайной величины – это неотрицательная неубывающая функция, удовлетворяющая условиям (4.9) или (4.10), что проиллюстрировано соответственно на рис. 10 и 11.

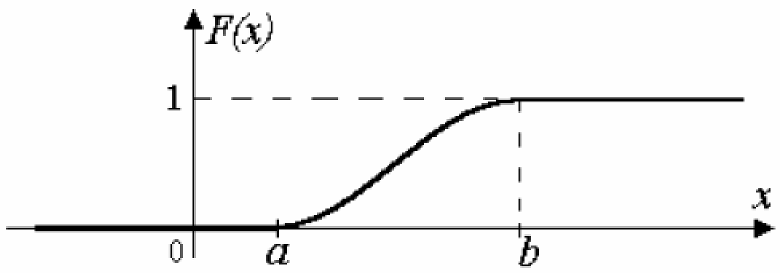


Рис. 10

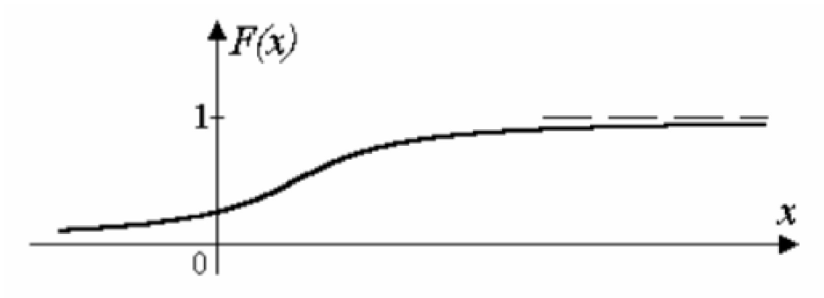


Рис. 11

<u>Плотность распределения вероятностей</u>

Пусть задана функция распределения $F(x)$ непрерывной случайной величины X. Предположим, что $F(x)$ дифференцируема в некоторой точке x. Обозначим $F/(x)$= $f(x)$ и выясним смысл $f(x)$. По определению производной

$$f(x) = F'(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}.$$

Учитывая формулу (4.8), имеем

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x \le X < x + \Delta x)}{\Delta x}$$

Отношение

$$\frac{P(x \le X < x + \Delta x)}{\Delta x}$$

можно понимать как среднюю плотность вероятности на отрезке [x, $x+\Delta x$], поэтому при $\Delta x \to 0$ предел этого отношения естественно считать плотностью вероятности в точке x. При этом величина $P(x \le X < x + \Delta x) \approx f(x)\Delta x = = f(x)dx$ называется дифференциалом вероятности (рис.12).

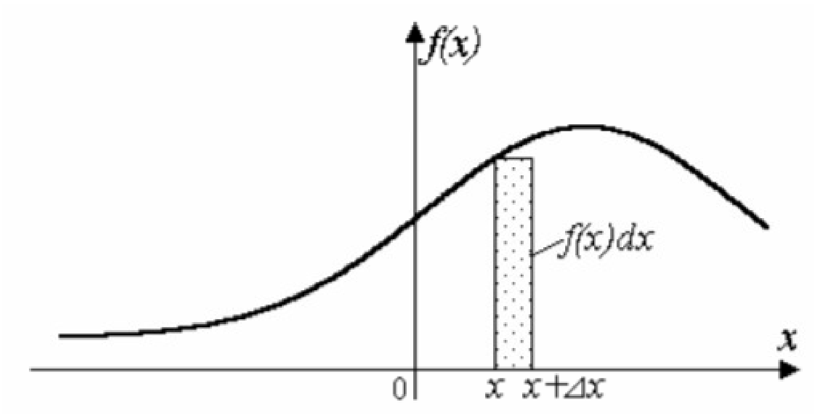


Рис. **12**

Функция $f(x)$, являющаяся производной функции распределения $F(x)$ случайной величины X, называется плотностью распределения вероятностей случайной величины X.

<ins>Свойства плотности распределения вероятностей</ins>

Свойство 1. Функция плотности распределения вероятностей неотрицательна, $f(x) \ge 0$.

Это свойство следует из того, что $f(x)$ является производной неубывающей функции $F(x)$.

Свойство 2. Если $f(x)$ непрерывна на отрезке [a, b], то

$$\int_a^b f(x)dx = P(a < X < b). \qquad (5.1)$$

Формула (5.1) следует из того, что, во-первых, $F(x)$ является первообразной функции $f(x)$, а также из равенства (4.8) и из формулы Ньютона–Лейбница.

Свойство 3.

$$\int_{-\infty}^{+\infty} f(x)dx = 1. \qquad (5.2)$$

Равенство (5.2) - это расширение *свойства* 2 на бесконечный промежуток. Действительно, считая, что

$$\int_{-\infty}^{+\infty} f(x)dx = P(-\infty < X < +\infty).$$

с учетом достоверности события $(-\infty < X < \infty)$ получим равенство (5.2).

Из *свойств* 1 и 3 следует, что плотностью распределения вероятностей случайной вели- чины может являться любая неотрицательная функция, для которой справедливо равенство (5.2), называемое *условием нормировки*.

Свойство 4. Функция распределения вероятностей может быть найдена по известной плотности распределения вероятностей с помощью формулы

$$F(x) = \int_{-\infty}^{x} f(x)dx. \qquad (5.3)$$

Для доказательства (5.3) распространим формулу (5.1) на полубесконечный промежуток $(-\infty\ , x)$

$$\int_{-\infty}^{x} f(x)dx = P(-\infty < X < x).$$

Учитывая, что событие ($-\infty < X < x$) совпадает с событием ($X < x$) и $P(X < x) = F(x)$ по определению, получим формулу (5.3).

Свойство 5. Если множеством значений величины X является отрезок $[a, b]$, то вне этого отрезка $f(x) = 0$.

Это утверждение следует из *свойства* 5 функции $F(x)$ и того, что производная постоянной равна нулю.

Замечание 1. Плотность распределения вероятностей $f(x)$ и функцию распределения вероятностей $F(x)$, учитывая формулы, связывающие эти функции, называют дифференциальной и интегральной функциями распределения случайной величины соответственно. График дифференциальной функции $f(x)$ принято называть кривой распределения, а саму функцию $f(x)$ называют законом распределения непрерывной случайной величины X.

Замечание 2. Если задана кривая распределения, то по ней можно узнать, чему равна вероятность $P(\alpha < X < \beta)$.

Действительно, из (5.1) следует, что $P(\alpha < X < \beta)$ равна площади криволинейной трапеции, ограниченной кривой распределения, осью абсцисс и прямыми $x = \alpha$ и $x = \beta$ (рис. 13). Распространив з а м е ч а н и е 2 на всю числовую прямую, с учетом условия нормировки (5.3) можно видеть, что площадь фигуры, ограниченной кривой распределения и осью абсцисс, равна единице.

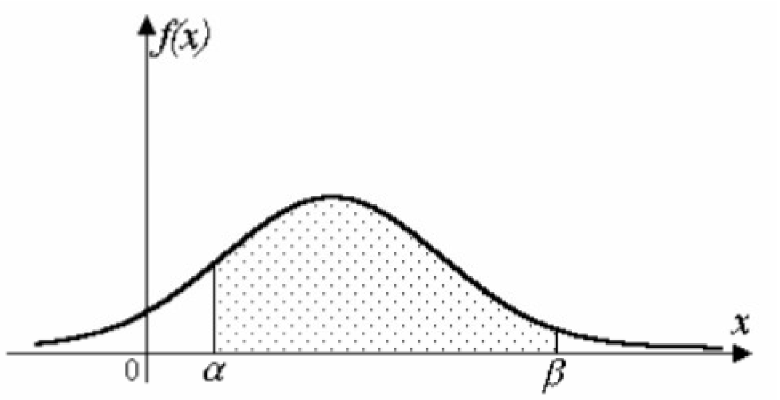


Рис. 13

Пример 1. Случайная величина X задана функцией распределения

$$F(x)=\begin{cases} 0, & x \le 0; \\ \frac{1}{4}x^2, & 0 < x \le 2; \\ 1, & x > 2. \end{cases}$$

Найти функцию плотности вероятности *f(x)*, построить графики функций *F(x)* и *f(x)* и вычислить $P(1 \le x \le 2)$.

Решение. Так как $f(x) = F'(x)$, то имеем

$$f(x)=\begin{cases} 0, & x \le 0; \\ \frac{1}{2}x, & 0 < x \le 2; \\ 0, & x > 2. \end{cases}$$

Строим графики функций $F(x)$ и $f(x)$ (рис. 14),

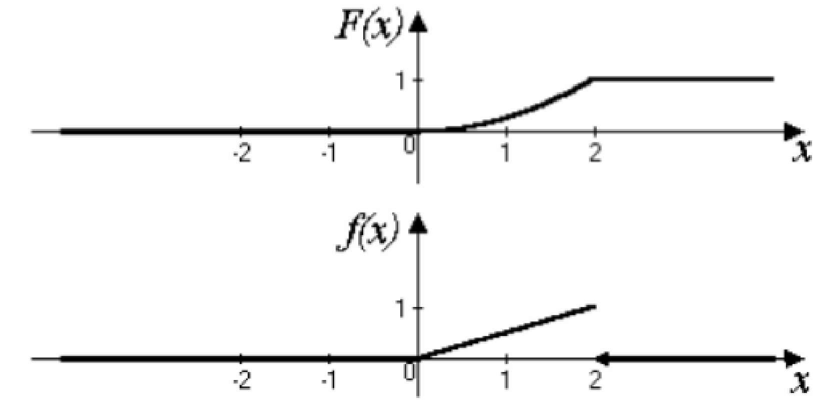


Рис. 14

а затем вычисляем искомую вероятность $P(1\leq X\leq 2)$ по формуле (4.5)

$P(1\leq X\leq 2)= F(2) - F(1)=1/4 \cdot(2)^2-1/4\cdot (1)^2=3/4.$

Числовые характеристики случайных величин

Любая случайная величина полностью описывается законом распределения. Можно показать, что для каждой случайной величины существуют числа, характеризующие эту величину в целом. Эти числа называют числовыми характеристиками случайной величины. Ниже рассмотрим некоторые числовые характеристики сначала для дискретной случайной величины, а затем распространим полученные формулы на непрерывные случайные величины.

Математическое ожидание дискретной случайной величины

Одной из основных числовых характеристик является математическое ожидание случайной величины.

Пусть закон распределения случайной величины X задан таблицей

X	x1	x2	...	xn
P	p1	p2	...	pn

Определение. Математическим ожиданием $M(X)$ случайной величины X называется число, определяемое по формуле

$$M(X) = \sum_{i=1}^{n} x_i p_i \ . \qquad (5.4)$$

Выясним вероятностный смысл математического ожидания. Пусть случайная величина наблюдается N раз и принимает при этих наблюдениях свои значения следующим образом: значение $x1 - N1$ раз, $x2 - N2$ раз, ... , $xn - Nn$ раз.

Вычислим среднее арифметическое $\bar{x}$ наблюдаемых значений по известной формуле

$$\bar{x} = \frac{x_1 N_1 + x_2 N_2 + ... + x_n N_n}{N}$$

Преобразуем $\bar{x}$ к виду

$$\bar{x} = \frac{x_1 N_1}{N} + \frac{x_2 N_2}{N} + ... + \frac{x_n N_n}{N}$$

Числа $\frac{N_i}{N}$ ($i = 1, ... , n$) равны относительным частотам случайных событий A_i , где A_i –появление значения xi (см. статистическое определение вероятности). Обозначим $\frac{N_i}{N} = W_i$. Тогда

$$x = x1W1 + x2W2 + ...+ xnWn \ . \qquad (5.5)$$

Будем неограниченно увеличивать число наблюдений N ($N\to\infty$). Тогда согласно статистическому определению вероятности случайного события

$$W_i \to P(A_i) = P(X = x_i) = p_i .$$

Переходя к пределу при $N\to\infty$ в формуле (5.5), получим

$$\lim_{N\to\infty} \overline{x} = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n = M(X). \qquad (5.6)$$

Из (5.6) можно сделать вывод, что математическое ожидание случайной величины − это число, к которому стремится среднее арифметическое наблюдаемых значений величины X, если число наблюдений неограниченно растет.

Математическое ожидание называют еще центром распределения вероятностей. Этот термин заимствован из механики. Поясним это следующим образом. Если единичная масса $m = 1$ распределена по оси OX так, что в точках x_i расположены массы $m_i = p_i$ ($i = 1,\ldots, n$), то точка x_c , где

$$x_c = \frac{\sum_{i=1}^{n} x_i p_i}{m} = \frac{\sum_{i=1}^{n} x_i p_i}{1} = M(X),$$

будет центром масс.

Прежде чем рассмотреть свойства математического ожидания, дадим несколько определений.

Определение 1. Случайные величины X и Y называются независимыми, если независимы случайные события ($X < x$) и ($Y < y$) для любых x, y.

Например, если X − число очков, выпадающих на одной игральной кости, а Y − число очков, выпадающих на другой игральной кости, то X и Y независимы.

Если события ($X < x$) и ($Y < y$) зависимы для каких-нибудь x, y, то X и Y − зависимые случайные величины.

Определение 2. Суммой $X+Y$ случайных величин X и Y называется случайная величина, значения которой равны сумме наблюдаемых значений величин X и Y.

Определение 3. Произведением X □Y случайных величин X и Y называется случайная величина, значения которой равны произведению наблюдаемых значений величин X и Y.

Например, если X и Y − число выпавших очков на двух костях соответственно, то законы распределения X и Y одинаковы и имеют вид

Найдем законы распределения величин $X+Y$ и XY. Очевидно, они таковы:

X	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

,

Y	1	2	3	4	5	6
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Найдем законы распределения величин $X+Y$ и XY. Очевидно, они таковы:

$X+Y$	2	3	4	5	6	7	8	9	10	11	12
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

,

XY	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
P	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Свойства математического ожидания

Свойство 1. Математическое ожидание постоянной равно этой постоянной, другими словами, если $X= c$, где $c=$ const, то $M(c)= c$.

Действительно, закон распределения величины $X=c$ имеет вид

X	c
P	1

Отсюда $M(c)=c\cdot 1=c$.

Свойство 2. Постоянный множитель можно выносить за знак математического ожидания, т.е. если κ= const, то $M(\kappa X)=\kappa M(X)$.

Действительно, если

X	$x1$	$x2$	…	xn
P	$p1$	$p2$	…	pn

то

κX	$\kappa x1$	$\kappa x2$	…	κxn
P	$p1$	$p2$	…	pn

Следовательно,

$$M(kX)=\sum_{i=1}^{n}(kx_i)p_i=k\sum_{i=1}^{n}x_i p_i=kM(X).$$

Свойство 3. Математическое ожидание суммы случайных величин равно сумме их математических ожиданий

$M(X+Y)=M(X)+M(Y)$.

Свойство 4. Математическое ожидание произведения независимых случайных величин равно произведению их математических ожиданий

$M(XY)=M(X)\,M(Y)$.

Замечание. *Свойства 3 и 4* можно распространить на любое конечное число слагаемых и сомножителей соответственно.

Свойство 5. Если $x1$ - наименьшее, а xn -наибольшее значения случайной величины X, то $x1 \leq M(X) \leq xn$.

Действительно, справедливы неравенства

$$M(X) = \sum_{i=1}^{n} x_i p_i \geq \sum_{i=1}^{n} x_1 p_i = x_1 \sum_{i=1}^{n} p_i = x_1 \cdot 1 = x_1,$$

$$M(X) = \sum_{i=1}^{n} x_i p_i \leq \sum_{i=1}^{n} x_n p_i = x_n \sum_{i=1}^{n} p_i = x_n \cdot 1 = x_n .$$

Из *свойств* 2 и 3 вытекает следствие.

Следствие Для любых чисел α и β и любых случайных величин X и Y

$M(\alpha X+\beta Y)= \alpha M(X)+\beta M(Y)$. (5.7)

Формулу (5.7) можно использовать при вычислении $M(X)$. Например, найдем $M(X)$ для случайной величины X

X	*1300*	*1325*	*1350*	*1375*
P	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{1}{10}$

Рассмотрим величину $\dfrac{X-1325}{25}$, ее закон распределения имеет вид**.**

$\frac{X-1325}{25}$	*-1*	*0*	*1*	*2*
P	$\frac{2}{10}$	$\frac{3}{10}$	$\frac{4}{10}$	$\frac{1}{10}$

Найдем математическое ожидание

$$M\left(\frac{X-1325}{25}\right)=-1\cdot\frac{2}{10}+0\cdot\frac{3}{10}+1\cdot\frac{4}{10}+2\cdot\frac{1}{10}=0{,}4.$$

Так как

$$M\left(\frac{X-1325}{25}\right)=\frac{M(X)-1325}{25},$$

то получим

$$0{,}4=\frac{M(X)-1325}{25}$$

и, следовательно, $M(X)$ =1325 + 0,4 □ 25 =1335.

Дисперсия дискретной случайной величины

Рассмотрим две случайные величины X и Y, где

X	-100	100
P	0,5	0,5

Y	-0,01	-0,01

P	0,5	0,5

Очевидно, математические ожидания обеих случайных величин являются одинаковыми и равными нулю, $M(X)=M(Y)=0$. В то же время сами случайные величины отличаются на четыре порядка. Поэтому наряду с математическим ожиданием вводят характеристики, показывающие, сколь далеко рассеяны значения случайной величины вокруг математического ожидания. Рассмотрим одну из числовых характеристик рассеяния, называемую дисперсией случайной величины.

Определение. Дисперсией случайной величины X называется число, обозначаемое $D(X)$ и вычисляемое по формуле

$$D(X)= M((X - M(X))^2). \qquad (6.1)$$

Случайная величина $X - M(X)$ − это отклонение X от математического ожидания $M(X)$. Очевидно,

$$M(X-M(X)) = M(X) - M(M(X)) = M(X) - M(X) = 0.$$

для любой случайной величины X (отклонения являются числами разных знаков и их среднее значение равно нулю). Это говорит о том, что отклонение нельзя принять за меру рассеяния. Поэтому для характеристики рассеяния на практике рассматривают модули отклонений или, чаще всего, квадраты отклонений.

Итак, дисперсия случайной величины X, определенная формулой (6.1) − это математическое ожидание квадрата ее отклонения от $M(X)$. Чтобы вычислить $D(X)$, составим закон распределения величины $(X-M(X))^2$. Пусть закон распределения дискретной случайной величины X имеет вид

X	$x1$	$x2$	...	xn
P	$p1$	$p2$	...	pn

(6.2)

Очевидно, случайная величина $(X - M(X))2$ принимает значения $(xi - M(X))^2$ с вероятностями pi (i=1,...,n). Поэтому

$$D(X) = M((X - M(X))^2) = \sum_{i=1}^{n} (x_i - M(x))^2 \cdot p_i \,. \qquad (6.3)$$

Если учесть свойства математического ожидания, то можно преобразовать формулу (6.1) к следующему виду:

$$D(X)=M(X^2) - M^2(X). \qquad (6.4)$$

Действительно,
$D(X)=M((X-M(X))^2)=M(X^2-2XM(X)+(M(X))^2)=M(X^2)+M(-2X\,M(X))+$
$+M(M(X)^2)=M(X^2)-2M(X)\,M(X)+(M(X))^2=M(X^2)-(M(X))^2$.

Для дискретной случайной величины X, определенной законом распределения (6.2), распределение величины X^2 задается таблицей

X^2	$x1^2$	$x2^2$	...	xn^2
P	$p1$	$p2$	...	pn

поэтому $M(X^2) = \sum_{i=1}^{n} x_i^{\,2} p_i$. Подставляя $M(X2)$ и $M(X)$ в (6.4), окончательно получаем формулу

$$D(X) = \sum_{i=1}^{n} x_i^{\,2} p_i - \left(\sum_{i=1}^{n} x_i p_i \right)^2 . \qquad (6.5)$$

<u>Свойства дисперсии</u>

Свойство 1. Дисперсия постоянной величины равна нулю, т.е. $D(c)$=0.

Действительно, так как

$M(c)= c$; $M^2(c)= c2$; $M(c^2)= c^2$;

то

$D(c) = M(c^2) – M^2(c) = c^2 − c^2 = 0.$

Свойство 2. Постоянный множитель можно выносить за знак дисперсии, возводя его в квадрат, т.е.

$D(cX)= c^2D(X).$

Покажем это, используя свойства математического ожидания:

$D(cX)=M[(cX − M(cX))^2]=M[c^2 (X − M(X))^2]=c^2M[(X − M(X))^2]=c^2D(X).$

Свойство 3. Дисперсия суммы двух независимых случайных величин равна сумме их дисперсий

$D(X + Y)= D(X) + D(Y).$

Доказательство.

$D(X +Y)=M[(X +Y)^2] − [M(X +Y)]^2 = M(X^2+ 2XY +Y^2) −[M(X) + M(Y)]^2=$

$= M(X^2) + 2M(X)M(Y) + M(Y^2) − M^2 (X) − 2M(X)M(Y) − M^2 (Y) =$

$= M(X^2) − M^2 (X)+ M(Y^2) − M^2 (Y)= D(X) + D(Y).$

В частности,

$D(X + c)= D(X).$

Свойство 4. Дисперсия разности двух независимых случайных величин равна сумме их дисперсий

$D(X-Y) = D(X+(-Y)) = D(X) + D(-Y) = D(X) + (-1)^2 D(Y) = D(X) + D(Y).$

Замечание. Если $X1, X2, \ldots, Xn$ − конечное число попарно независимых случайных величин, то нетрудно доказать, что

$$D(X1 + X2 + \ldots + Xn) = D(X1) + D(X2) + \ldots + D(Xn).$$

Так же, как и в случае математического ожидания, для дисперсии есть аналог в механике. Если в точках находятся массы pi (причем $p = \sum_{i=1}^{n} p_i = 1$) и xс – центр масс, то степень разброса такой системы материальных точек вокруг центра масс характеризуется моментом инерции системы относительно центра масс

$$J = \frac{\sum_{i=1}^{n} (x_i - x_c)^2 \cdot p_i}{p}.$$

Так как xс $= M(X)$, а $p=1$, то $J = D(X)$.

Среднее квадратическое отклонение

Дисперсия случайной величины имеет размерность квадрата случайной величины. Для наглядности удобнее пользоваться величиной, размерность которой совпадает с размерностью случайной величины. Для этого из дисперсии извлекают квадратный корень. Полученная таким образом величина называется средним квадратическим отклонением случайной величины X и обозначается через $\sigma(X)$:

$$\sigma(X) = \sqrt{D(X)}. \qquad (6.6)$$

Пример. Дан закон распределения дискретной случайной величины X

X	3	5	10	12
P	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{6}$

Найти $M(X)$, $D(X)$ и $\sigma(X)$.

Решение. Вычисляем $M(X)$ по формуле (5.4):

$$M(X) = 3\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 10\cdot\frac{1}{2} + 12\cdot\frac{1}{6} = \frac{25}{3}.$$

Дисперсию $D(X)$ и среднее квадратическое отклонение $\sigma(X)$ вычисляем по формулам (6.5) и (6.6):

$$D(X) = 3^2\cdot\frac{1}{6} + 5^2\cdot\frac{1}{6} + 10^2\cdot\frac{1}{2} + 12^2\cdot\frac{1}{6} - \left(\frac{25}{3}\right)^2 = \frac{92}{9},$$

$$\sigma(X) = \sqrt{D(X)} = \sqrt{\frac{92}{9}} \approx 3{,}2.$$

Числовые характеристики непрерывной случайной величины

Введем понятие математического ожидания для непрерывной случайной величины X, $a \le X \le b$. Пусть $f(x)$ - плотность вероятности величины X. Отрезок $[a, b]$ разобьем на n элементарных отрезков $[a, x1]$, $[x1, x2]$,...,$[xn\text{-}1, b]$ с длинами $\Delta xк = xк$ - $xк$-1. В каждом из элементарных участков произвольным образом выберем точку $\xi\in[xк\text{-}1, xк]$ и составим произведение $f(\xi к)\Delta xк$. Полученная величина приближенно выражает вероятность $pк$ попадания случайной величины в $к$-й отрезок разбиения отрезка [a, b]. Образуем произведения ξ_k на соответствующие вероятности $p_k = f(\xi_k)\Delta x_k$ ($k=1,\ldots, n$), а затем просуммируем эти произведения для всех участков

$$\sum_{k=1}^{n} \xi_k p_k = \sum_{k=1}^{n} \xi_k f(\xi_k)\Delta x_k .$$

Эта сумма, с одной стороны, является математическим ожиданием дискретной случайной величины X_n

X_n	ξ_1	ξ_2	...	ξ_n
P	p_1	p_2	...	p_n

“близкой” к данной случайной величине X, а с другой – будет интегральной для функции $xf(x)$ на отрезке [a, b]. Неограниченно измельчая разбиение [a, b], в пределе получим число, которое называется математическим ожиданием непрерывной случайной величины X (если этот предел существует). Но с другой стороны, полученный предел будет равен определенному интегралу

$$\int_a^b x \cdot f(x)dx .$$

Таким образом, получаем следующее определение.

Определение. Математическое ожидание непрерывной случайной величины X, распределенной на отрезке [a, b], определяется по формуле

$$M(X) = \int_a^b x \cdot f(x)dx . \qquad (6.11)$$

Замечание 1. Если значения случайной величины X заполняют всю числовую ось и $f(x)$– плотность вероятности величины X, то математическое ожидание величины X определяется по формуле

$$M(X)=\int\limits_{-\infty}^{+\infty} x\cdot f(x)dx. \qquad (6.12)$$

Замечание 2. Можно доказать, что математическое ожидание непрерывной случайной величины обладает теми же свойствами, что и математическое ожидание дискретной случайной величины.

Замечание 3. Дисперсия непрерывной случайной величины также определяется по формуле (6.1). Поэтому для дисперсии $D(X)$ непрерывной случайной величины остаются справедливыми формула (6.3) и все свойства дисперсии дискретной случайной величины.

Найдем формулу для вычисления дисперсии непрерывной случайной величины X, имеющей плотность вероятности $f(x)$. Учтем, что случайная величина $(X - M(X))2$ принимает значения $(x - M(X))2$ и плотность вероятностей этих значений $f(x)$. Используя формулу (6.12), получаем

$$D(X)=\int\limits_{-\infty}^{+\infty}(x-M(X))^2\cdot f(x)dx. \qquad (6.13)$$

Аналогично можно записать $D(X)$ по формуле (6.3)

$$D(X)=\int\limits_{-\infty}^{+\infty}x^2\cdot f(x)dx-\left(\int\limits_{-\infty}^{+\infty}x\cdot f(x)dx\right)^2. \qquad (6.14)$$

Замечание 4. Если X непрерывно распределена на отрезке $[a, b]$, то $D(X)$ вычисляется по формулам (6.13) или (6.14), в которых несобственный интеграл превращается в определенный интеграл по отрезку $[a, b]$.

Замечание 5. Следует отметить, что при нахождении числовых характеристик непрерывных случайных величин приходится иметь дело с несобственными интегралами, которые не всегда сходятся. Иными словами, некоторые случайные величины не имеют, например, математического ожидания или дисперсии.

Пример. Непрерывная случайная величина *X* имеет функцию плотности распределения

$$f(x)=\begin{cases} 0, & x\le 0; \\ Ax^2, & 0<x\le 5; \\ 0, & x>5. \end{cases}$$

Найти значение параметра *A*, функцию распределения величины *X*, *M*(*X*), $\sigma(X)$ и $P(0\le X\le 2)$.

Решение. Для отыскания значения *A* воспользуемся свойством дифференциальной функции $f(x)$

$$\int_{-\infty}^{+\infty} f(x)dx=1.$$

Получим

$$1=\int_{-\infty}^{0} f(x)dx+\int_{0}^{5} f(x)dx+\int_{5}^{+\infty} f(x)dx=0+\int_{0}^{5} Ax^2dx+0=A\frac{x^3}{3}\bigg|_0^5=\frac{125}{5}A,$$

Откуда

$$A=\frac{3}{125},$$

и, следовательно,

$$f(x)=\begin{cases} 0, & x\le 0; \\ \frac{3}{125}x^2, & 0<x\le 5; \\ 0, & x>5. \end{cases}$$

Интегральную функцию $F(x)$ найдем по формуле

$$F(x)=\int_{-\infty}^{x} f(x)dx.$$

Для $x \le 0$ $\qquad F(x)=\int_{-\infty}^{x} 0\cdot dx=0.$

Для $0< x \le 5$ $\qquad F(x)=\int_{-\infty}^{0} 0\cdot dx+\int_{0}^{x}\frac{3}{125}\cdot x^2dx=\frac{x^3}{125}\Big|_0^5=\frac{x^3}{125}.$

Для $x>5$ $\qquad F(x)=\int_{-\infty}^{0} 0\cdot dx+\int_{0}^{5}\frac{3}{125}\cdot x^2dx+\int_{5}^{+\infty} 0\cdot dx=1.$

Функция распределения будет иметь вид

$$F(x)=\begin{cases} 0, & x\le 0; \\ \dfrac{x^3}{125}, & 0<x\le 5; \\ 1, & x>5. \end{cases}$$

Учитывая, что $0\le X\le 5$, вычислим $M(X)$ по формуле (6.11):

$$M(X)=\int_{0}^{5} x\frac{3x^2}{125}dx=\frac{3x^4}{500}\Big|_0^5=\frac{15}{4}.$$

Для нахождения дисперсии $D(X)$ воспользуемся формулой (6.14):

$$D(X)=\int_{0}^{5} x^2\frac{3x^2}{125}dx-\left(\frac{15}{4}\right)^2=\frac{3x^5}{125\cdot 5}\Big|_0^5-\frac{225}{16}=\frac{15}{16}.$$

Следовательно, среднее квадратическое отклонение будет равно

$$\sigma(X)=\sqrt{D(X)}=\sqrt{\frac{15}{16}}\approx 0{,}97.$$

Теперь, используя формулу $P(\alpha\leq X\leq\beta)=F(\beta)-F(\alpha)$, вычислим $P(0\leq X\leq 2)$:

$$P(0\leq X\leq 2)=F(2)-F(0)=\frac{2^3}{125}-0=\frac{8}{125}.$$

Замечание. Кроме $M(X)$, $D(X)$, $\sigma(X)$ существуют другие характеристики случайных величин, например, начальные и центральные моменты различных порядков и др. [1]. В этом курсе лекций они не приводятся. Ниже мы рассмотрим примеры законов распределения наиболее часто встречающихся случайных величин.

Биномиальное распределение

Пусть проводится n одинаковых независимых испытаний, в каждом из которых событие A появляется с вероятностью p и не появляется с вероятностью $q=1-p$ (схема Бернулли). Рассмотрим дискретную случайную величину X, равную числу появлений события A в этих n испытаниях. Очевидно, закон распределения величины X имеет вид

X	0	1	…	n
P	$Pn(0)$	$Pn(1)$		$Pn(n)$

$\boxed{(7.1)}$

Такой закон распределения называют биномиальным распределением. Здесь вероятности $P\mathrm{n}(k)$ вычисляются по формуле Бернулли.

Найдем математическое ожидание и дисперсию дискретной случайной величины X, распределенной по биномиальному закону. Случайную величину представим в виде

$X = X1 + X2 + \ldots + Xn$,

где Xk – число появления события A в κ-м испытании. По свойству математического ожидания

$M(X)=M(X1) + M(X2) + \ldots +M(Xn)$.

Посчитаем отдельно $M(X\text{k})$. Очевидно, для любого κ имеем

$X\text{k}$	0	1
P	q	p

Поэтому

$M(X\text{к})= x1p1+ x2p2= 1\cdot p+ 0\cdot q= p$, ($\kappa$=1,2, ... n).

Таким образом, математическое ожидание дискретной случайной величины, распределенной по биномиальному закону, равно
$M(X)=n\cdot M(X\text{k})= n\cdot p$.

Так же, как и при вычислении $M(X)$, для дисперсии будем иметь

$D(X)= D(X1) + D(X2) + \ldots + D(Xn)$.

Так как $D(X\kappa)= M(X\kappa^2) - (M(X\kappa))^2$, то найдем сначала $M(X^2\kappa)$. Закон распределения величины $X^2\kappa$ имеет вид

X^2k	0^2	1^2
P	q	p

поэтому $M(X^2k) = 0 \cdot q + 1 \cdot p = p$, откуда

$D(X\kappa) = p - (p)^2 = p(1 - p) = pq$, ($\kappa$=1, ... , n).

Следовательно, дисперсия $D(X) = n\ D(X\text{к}) = npq$; среднее квадратическое отклонение $\sigma(X) = \sqrt{npq}$.

Вывод. Случайная величина X, распределенная по биномиальному закону, характеризуется двумя параметрами n и p, а ее числовые характеристики равны

$$M(X) = np,\ D(X) = npq,\ \sigma(X) = \sqrt{npq}\ . \qquad (7.2)$$

Пример. Вероятность попадания в цель при одном выстреле p=0,6. Найти математическое ожидание и дисперсию числа попаданий, если произведено 10 выстрелов.

Очевидно, что число попаданий − это случайная величина, имеющая биномиальное распределение с параметрами n=10 и p=0,6. Используя формулы (7.2), получаем $M(X)$=10$\cdot$0,6=6, $D(X)$=10 $\cdot$0,6 $\cdot$0,4 = 2,4 и σ(X)= 2,4 ≈1,55. Значения $M(X)$=6 и $\sigma(X)$≈1,55 говорят о том, что, производя много серий из 10 выстрелов, среднее число попаданий близко к шести, а число попаданий отклоняется от шести в среднем не более чем на два попадания.

Равномерное распределение

Существуют случайные величины, распределенные на отрезке [a, b], причем каждое значение $x \in [a,\ b]$ является равновозможным. Такие величины называются равномерно распределенными на [a, b].

Пример 1. Известно, что на электрических часах минутная стрелка передвигается раз в минуту. Если часы не отстают и не спешат, то разность реального времени и времени показанного на часах, для случайно взглянувшего на часы человека − это случайная величина, имеющая равномерное распределение на отрезке [0,1] (здесь единица измерения − минута).

Пример 2. Если значение некоторой величины округлено до десятых, то для человека, пользующегося затем этим значением, погрешность округления – это случайная величина, равномерно распределенная на отрезке [− 0,05; 0,05]. Плотность вероятностей равномерно распределенной случайной величины X постоянна на [a, b]. Ее значение на [a, b] легко определяется с помощью условия нормировки (5.2).

Определение. Непрерывная случайная величина X называется равномерно распределенной на отрезке [a, b], если ее плотность вероятности задается формулой

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b; \\ 0, & x \notin [a,b]. \end{cases} \qquad (7.3)$$

Вычислим последовательно математическое ожидание и дисперсию равномерно распределенной случайной величины X, используя формулы (6.11) и (6.13):

$$M(X) = \int_a^b x \cdot f(x)dx = \frac{1}{b-a}\int_a^b xdx = \frac{b^2-a^2}{b-a} = \frac{a+b}{2},$$

$$D(X) = \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b xdx = \frac{(b-a)^2}{12}.$$

Найдем функцию распределения случайной величины X:

$$F(x) = \int_{-\infty}^{x} f(x)dx,$$

которая для $x \le a$ будет равна нулю:

$$F(x) = \int_{-\infty}^{x} 0dx = 0,$$

для $x \in [a,b]$

$$F(x) = \int_{-\infty}^{a} 0dx + \int_{a}^{x} \frac{1}{b-a}dx = \frac{1}{b-a}x\Big|_{a}^{x} = \frac{x-a}{b-a},$$

а для $x > b$

$$F(x) = \int_{-\infty}^{a} 0dx + \int_{a}^{b} \frac{1}{b-a}dx + \int_{b}^{x} 0dx = \int_{a}^{b} \frac{1}{b-a}dx = 1,$$

Таким образом, функция распределения равномерно распределенной случайной величины имеет вид

$$F(x) = \begin{cases} 0, & x \le a; \\ \dfrac{x-a}{b-a}, & a < x \le b; \\ 1, & x > b. \end{cases} \qquad (7.4)$$

Вычислим $P(\alpha \le X \le \beta)$ для $(\alpha, \beta) \in [a, b]$.

Так как $P(\alpha \le X \le \beta) = F(\beta) - F(\alpha)$, а $F(\beta) = \dfrac{\beta - a}{b-a}$ и $F(\alpha) = \dfrac{\alpha - a}{b-a}$, то

$$P(\alpha \le X \le \beta) = \frac{\beta - \alpha}{b-a}. \qquad (7.5)$$

Вывод. Равномерно распределенная на отрезке [*a, b*] случайная величина характеризуется двумя параметрами *a* и *b*, графики ее функций распределения и плотности распределения имеют вид, представленный на рис. 15 ,

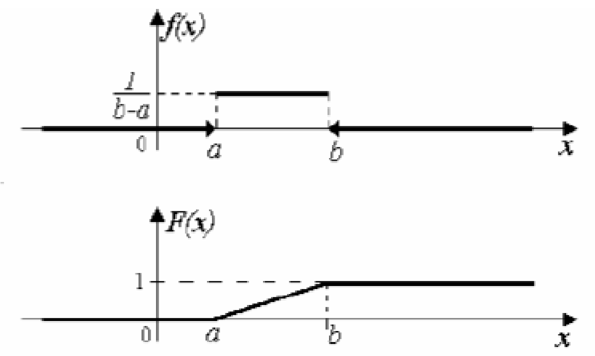


Рис. 15

а числовые характеристики равны

$$M(X) = \frac{a+b}{2},\ D(X) = \frac{(b-a)^2}{12},\ \sigma(X) = \frac{(b-a)}{2\sqrt{3}}.$$

Нормальное распределение

Нормальным распределением (распределением Гаусса) называется распределение непрерывной случайной величины, для которой функция плотности имеет вид

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \qquad (8.6)$$

где *a*, σ - параметры распределения (константы); $-\infty < a < \infty$, $\sigma > 0$. Влияние параметров *a* и σ на вид кривой распределения иллюстрируют рис. 16–18.

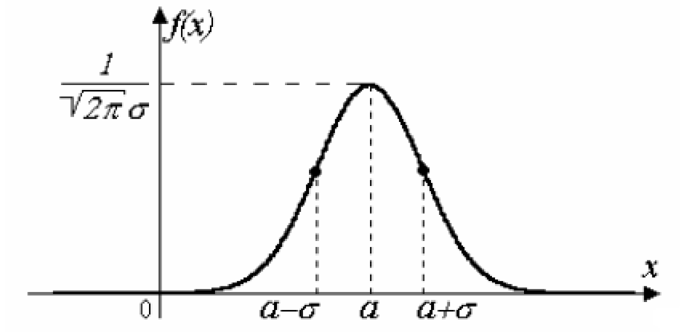


Рис.16

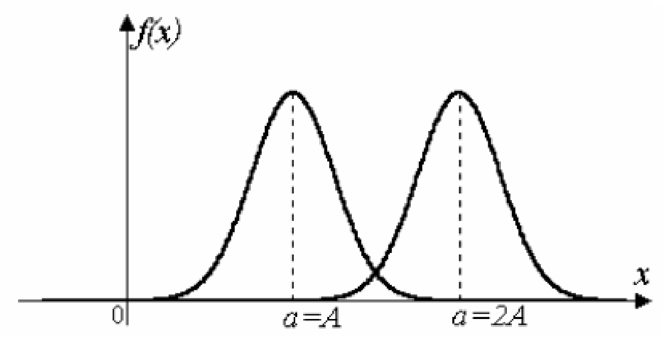


Рис. 17

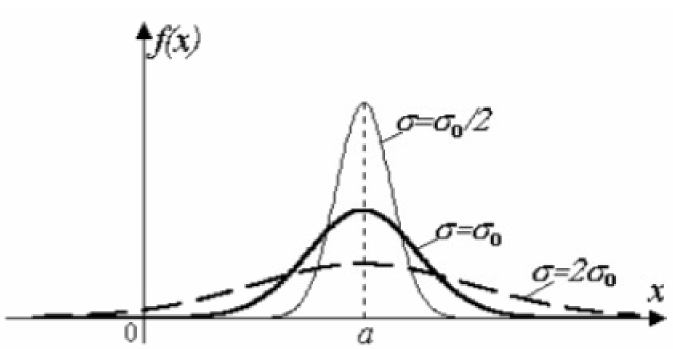


Рис. 18

Нормально распределенную величину с параметрами a и σ обозначают N(a; σ). Величину N(0; 1) называют нормированной нормально распределенной величиной (a=0, σ=1). Для нее законом распределения будет изученная ранее функция φ (x) (см. схему Бернулли):

$$f_0(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad (8.7)$$

Функция распределения Fo(x) величины N(0, 1) в свою очередь определится формулой

$$F_0(x) = \int_{-\infty}^{x} f_0(x)dx = \int_{-\infty}^{0} \varphi(x)dx + \int_{0}^{x} \varphi(x)dx.$$

Преобразуем полученное выражение, используя функцию Лапласа

$$\Phi(x) = \int_{0}^{x} \varphi(x)dx = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-\frac{x^2}{2}} dx,$$

рассмотренную ранее в интегральной формуле Лапласа (3.4). Из свойств функции $\Phi(x)$ следует, что

$$\int_{-\infty}^{0} \varphi(x)dx = \Phi(0) - \Phi(-\infty) = \Phi(0) + \Phi(+\infty).$$

Таким образом, получаем связь функций Fo(x) и $\Phi(x)$:

$$F\text{o}(x) = 0{,}5 + \Phi(x). \qquad (8.8)$$

График функции Fo(x) представлен на рис. 19.

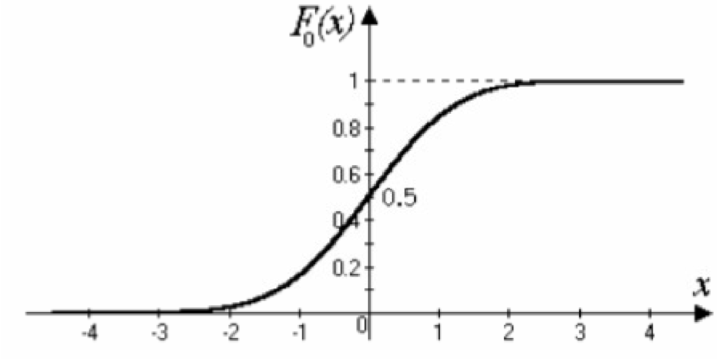


Рис. 19

Нетрудно увидеть, что для любой нормально распределенной величины N(a; σ) будут справедливы равенства

$$f(x) = \frac{1}{\sigma} f_0\left(\frac{x-a}{\sigma}\right), \qquad F(x) = F_0\left(\frac{x-a}{\sigma}\right).$$

Поэтому из (8.8) имеем

$$F(x) = 0{,}5 + \Phi\left(\frac{x-a}{\sigma}\right). \qquad (8.9)$$

Из равенства

$$f(x) = \frac{1}{\sigma} f_0\left(\frac{x-a}{\sigma}\right) = \frac{1}{\sigma} \varphi\left(\frac{x-a}{\sigma}\right)$$

следует, что график $f(x)$ можно получить из графика $\phi(x)$ растяжением последнего по оси Ox в σ раз, затем сжатием по оси Oy в σ раз и, наконец, сдвигом по оси Ox на a.

Для случайной величины X = N(a; σ) нетрудно получить формулу вычисления $P(\alpha \le X \le \beta)$. Действительно, используя формулу (8.9), имеем

$$P(\alpha \le X \le \beta) = F(\beta) - F(\alpha) = \left(0{,}5 + \Phi\left(\frac{\beta - a}{\sigma}\right)\right) - \left(0{,}5 + \Phi\left(\frac{\alpha - a}{\sigma}\right)\right).$$

т.е.

$$P(\alpha \le X \le \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right). \qquad (8.10)$$

Замечание. Отметим, что плотность распределения вероятностей удовлетворяет условию нормировки для любых a и σ. Этот факт легко проверить, если воспользоваться интегралом Пуассона

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}\ . \qquad (8.11)$$

Из (8.11) следует, что

$$\int_{-\infty}^{+\infty} \varphi(x) dx = 1$$

и тогда для любой случайной величины N(a, σ) имеем

$$\int_{-\infty}^{+\infty} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \left| \begin{matrix} t = \dfrac{(x-a)}{\sigma} \\ x = a + t \cdot \sigma; dx = \sigma \cdot dt \\ x \to -\infty; t \to -\infty \\ x \to +\infty; t \to +\infty \end{matrix} \right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{t^2}{2}} dt = 1$$

Найдем числовые характеристики нормально распределенной случайной величины N(a,σ).

Сначала вычислим для N(a, σ) математическое ожидание учитываем, что

$$M(X)=\int\limits_{-\infty}^{+\infty} x\cdot f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} x\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left|\begin{array}{c} t=\dfrac{(x-a)}{\sigma} \\ x=a+t\cdot\sigma; dx=\sigma\cdot dt \\ x\to-\infty; t\to-\infty \\ x\to+\infty; t\to+\infty \end{array}\right|=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}(a+t\cdot\sigma)\cdot e^{-\frac{t^2}{2}}dt=\frac{a}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} a\cdot e^{-\frac{t^2}{2}}dt+\frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} t\cdot e^{-\frac{t^2}{2}}dt=\left|\textit{используем}(8.11)\right|=$$

$$=a\cdot 1+\frac{\sigma}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}d\frac{t^2}{2}=a-\frac{\sigma}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{+\infty}=\left|\textit{учитывая, что}\lim_{t\to\pm\infty}e^{-\frac{t^2}{2}}=0\right|=a$$

Теперь вычислим дисперсию

$$D(X)=\int\limits_{-\infty}^{+\infty}\left(x-M(X)\right)^2\cdot f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty}\left(x-a\right)^2\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx=\left|\begin{array}{c} t=\dfrac{(x-a)}{\sigma} \\ x=a+t\cdot\sigma; dx=\sigma\cdot dt \\ x\to-\infty; t\to-\infty \\ x\to+\infty; t\to+\infty \end{array}\right|=$$

$$=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} t^2\cdot\sigma^2\cdot e^{-\frac{t^2}{2}}dt=\frac{\sigma^2}{\sqrt{2\pi}}\int\limits_{-\infty}^{+\infty} t\cdot t\cdot e^{-\frac{t^2}{2}}dt=\left|\begin{array}{c} u=t, du=dt \\ dv=t\cdot e^{-\frac{t^2}{2}}dt \\ v=\int e^{-\frac{t^2}{2}}d\dfrac{t^2}{2}=-e^{-\frac{t^2}{2}} \end{array}\right|=$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left(-te^{-\frac{t^2}{2}}\Bigg|_{-\infty}^{+\infty}+\int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt\right)=\left|\textit{учитывая, что}\lim_{t\to\pm\infty}te^{-\frac{t^2}{2}}=\lim_{t\to\pm\infty}\frac{t}{e^{\frac{t^2}{2}}}=0\right|=$$

$$=\frac{\sigma^2}{\sqrt{2\pi}}\left(-\int\limits_{-\infty}^{+\infty} e^{-\frac{t^2}{2}}dt\right)=\left|\textit{используем}(8.11)\right|=\sigma^2\cdot 1=\sigma^2$$

Теперь можно определить среднее квадратическое отклонение

$$\sigma(X)=\sqrt{D(X)}=\sigma.$$

Вывод. Параметры a и σ нормального распределения являются его математическим ожиданием и средним квадратическим отклонением соответственно.

Так как величина N(a, σ) распределена симметрично относительно своего математического ожидания, то часто вычисляют вероятность попадания значений такой величины в симметричный интервал (a−δ, a+δ) некоторого радиуса δ с центром в точке a . Для вычисления этой вероятности воспользуемся формулой (8.10):

$$P(|X-a|<\delta)=P(a-\delta\le X\le a+\delta)=\Phi\left(\frac{a+\delta-a}{\sigma}\right)-\Phi\left(\frac{a-\delta-a}{\sigma}\right)=\Phi\left(\frac{\delta}{\sigma}\right)-\Phi\left(\frac{-\delta}{\sigma}\right)$$

Учитывая нечетность функции Лапласа, получим формулу

$$P(|X-a|<\delta)=2\Phi\left(\frac{\delta}{\sigma}\right). \qquad (8.12)$$

Пример. Станок-автомат штампует детали, у которых контролируется один размер. Деталь считается годной, если отклонение этого размера от проектного не более 0,1 ед. Найти, какой процент годных изделий можно гарантировать, если известно, что при штамповке деталей контролируемый размер является нормально распределенной случайной величиной с σ= 0,15 ед.

Решение. Процент годных деталей − это вероятность того, что случайно взятая деталь небракованная в %. Обозначим контролируемый размер детали через X, X=N(a; 0,1), a − проектный размер. Искомая вероятность − это вероятность того, что $|X-a|<0{,}1$. По формуле (8.12)

$$P(|X-a|<0{,}1)=2\Phi\left(\frac{0{,}1}{0{,}15}\right)=2\Phi(0{,}67)=2\cdot 0{,}2486=0{,}4972\,.$$

Умножая 0,4972 на 100 %, получаем, что годных деталей можно гарантировать почти в 50% от всех штампуемых.

Так как Ф(x) для $x \ge 5$ можно считать равной единице (см. свойства), то уже при δ=3σ получим

$$P(|X-a|<3\sigma)=2\Phi\left(\frac{3\sigma}{\sigma}\right)=2\Phi(3)=0{,}9973\approx 1. \qquad (8.13)$$

Равенство (8.13) называют *правилом трех сигм*. Его можно сформулировать так: для случайной величины, распределенной по нормальному закону, можно практически достоверно утверждать, что ее значения отклоняются по модулю от математического ожидания не более чем на три средних квадратических отклонения (рис. 20).

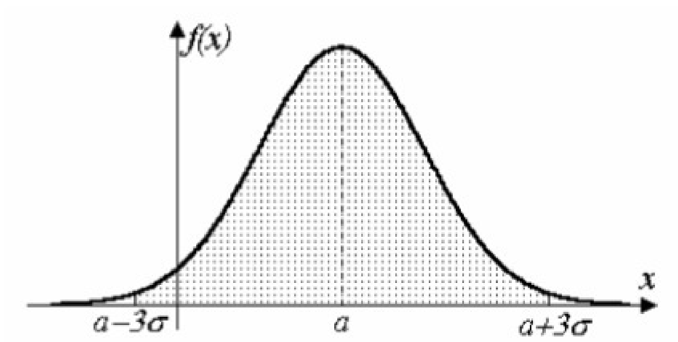


Рис. 20

Нормально распределенные случайные величины наиболее распространены на практике. Объяснение этому факту дал русский математик А.М. Ляпунов, доказав центральную предельную теорему.

Теорема. Если случайная величина X равна сумме очень большого числа попарно независимых случайных величин, влияние каждой из которых на всю сумму ничтожно мало, то X имеет распределение, близкое к нормальному.

Действительно, случайность многих величин порождена большим количеством случайных факторов, влияние каждого из которых в отдельности ничтожно.

Замечание. Рассмотрим биномиальное распределение X с параметрами n, p, $q = 1 - p$ и величину N(np; $\sqrt{npq}$). Пусть n велико, тогда в биномиальном распределении (8.1) значения вероятностей $pm = P(X = m)$ вычисляются приближенно по локальной формуле Лапласа:

$$p_m \approx \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(m-np)^2}{2npq}}$$

С другой стороны, плотность вероятности величины N(np; $\sqrt{npq}$) имеет вид

$$f(x) = \frac{1}{\sqrt{npq}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-np)^2}{2npq}}$$

Очевидно, для x=m получаем приближенное равенство $pm \approx f(m)$. Это означает, что при больших n многоугольник распределения биномиального закона близок к кривой Гаусса.

Закон больших чисел

Как известно, нельзя заранее уверенно предвидеть, какое из возможных значений примет случайная величина в итоге испытания; это зависит от многих случайных причин, учесть которые невозможно. Однако при некоторых сравнительно широких условиях суммарное поведение достаточно большого числа случайных величин почти утрачивает случайный характер и становится закономерным.

Для практики очень важно знание условий, при выполнении которых совокупное действие очень многих случайных причин приводит к результату, почти не зависящему от случая, так как позволяет предвидеть ход явлений. Эти условия и указываются в теоремах, носящих название закона больших чисел. К ним относятся теоремы Чебышева и Бернулли.

Теорема Бернулли

Пусть проводится n независимых испытаний в каждом из которых вероятность появления события A равна p. Можно ли предвидеть, какова примерно будет относительная частота появлений события? Положительный ответ на этот вопрос дает теорема, доказанная Якобом Бернулли, которая получали название «закона больших чисел» и положила начало теории вероятностей как науке.

Теорема Бернулли. Если в каждом из n независимых испытаний в вероятность p появления события A постоянна, то как угодно близка к единице вероятность того, что отклонение относительной частоты от вероятности p по абсолютной величине будет сколь угодно малым, если число испытаний достаточно велико., т.е.:

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n} - p\right| < \varepsilon\right) = 1.$$

<u>Предмет математической статистики</u>

Математическая статистика – раздел математики, занимающийся обработкой статистических данных с целью установления закономерностей, присущих массовым случайным явлениям. Статистические данные представляют собой сведения о том, какие значения принял в результате наблюдений интересующий нас признак (случайная величина). Методы математической статистики разработаны на основе методов теории вероятностей. Основной метод математической статистики – *выборочный*. Суть его в том, что по сравнительно небольшому количеству статистических данных делаются выводы о рассматриваемом явлении, процессе и т. п. Разумеется, эти выводы – лишь приблизительные оценки вероятностного характера для изучаемого явления или процесса. Математическая статистика разработала методы сбора выборочных данных и их описание, позволяющее получать по возможности более точные и надежные оценки, указывая при этом степень их надежности.

Математическая статистика возникла в XVI в. и развивалась параллельно с теорией вероятностей. В XIX–XX вв. большой вклад в развитие математической статистики внесли П.Л. Чебышев, А.А. Марков, А.Н. Ляпунов, К. Гаусс, К. Пирсон, А.Н. Колмогоров, Р. Фишер, Ю. Нейман и другие известные ученые-математики.

Выборочный метод

Пусть требуется изучить признак, присущий некоторой совокупности однородных объектов, называемой генеральной совокупностью, или некоторому процессу, протекающему в постоянных условиях. Пусть этот признак характеризуется случайной величиной X. Например, X – средний балл аттестата выпускников школ Воронежской области за последние десять лет или X – прочность кирпичей, изготавливаемых по некоторой технологии. Для изучения величины X из генеральной совокупности выбирают сравнительно немного объектов. Это последнее множество объектов называют выборкой, а количество выбранных объектов – объемом выборки. Чтобы по выборке можно было делать выводы о величине X, она должны обладать свойством репрезентативности (представительности). Выполнения этого свойства можно добиться, во-первых, методами отбора объектов в выборку, во-вторых, достаточно большим объемом выборки. Отобрать объекты в выборку необходимо случайным образом, что, как правило, является достаточно сложным. Здесь мы не останавливаемся на этом вопросе, но рекомендуем отнестись к нему серьезно и прежде чем собирать выборочные данные, изучить дополнительную литературу.

Основные понятия математической статистики

Пусть имеется выборка объема n и для каждого объекта выборки найдено значение x изучаемой случайной величины X (найденные значения x называют вариантами). Варианты $x1, \ldots, xn$, расположенные в порядке возрастания, называют вариационным рядом. В вариационном ряду варианты могут повторяться. Количество повторений варианты xi называют ее частотой и обозначают ni , а отношение $\frac{n_i}{n} = w_i$ называют относительной частотой варианты xi. Отметим, что определение относительной частоты варианты xi совпадает с определением относительной частоты случайного события $X=xi$ в n испытаниях при статистическом определении вероятности ($wi=W(X=xi)$). Такое совпадение используется в статистическом описании случайной величины X. Если в вариационном ряду различных вариантов немного, то это говорит о дискретности случайной величины X. В этом случае рассматривают соответствие между различными вариантами и их частотами (относительными частотами) в виде таблицы

xi	$x1$	$x2$	…	xk
ni	$n1$	$n2$	…	nk

или

xi	$x1$	$x2$	…	xk
wi	$w1$	$w2$	…	wk

Очевидно, здесь $\sum_{i=1}^{k} n_i = n$, $\sum_{i=1}^{k} w_i = 1$.

Указанное соответствие называют статистическим распределением частот (относительных частот) выборки, оно является аналогом закона распределения дискретной случайной величины (стоит лишь заменить $ni(wi)$ на $pi=P(X=xi)$). Аналогом многоугольника распределения является ломаная на координатной плоскости, последовательно соединяющая точки (xi, ni) (или (xi, wi)). Эта ломаная называется полигоном частот (относительных частот).

Пример. Построить полигон относительных частот по данному статистическому распределению относительных частот выборки (табл. 2).

Таблица 2

x_i	4	10	16	20	24	30	$\sum$
n_i	15	18	6	4	5	12	60
w_i	0,25	0,3	0,1	0,067	0,083	0,2	1

Строим на координатной плоскости точки (x_i, w_i) и соединяем их (рис. 21).

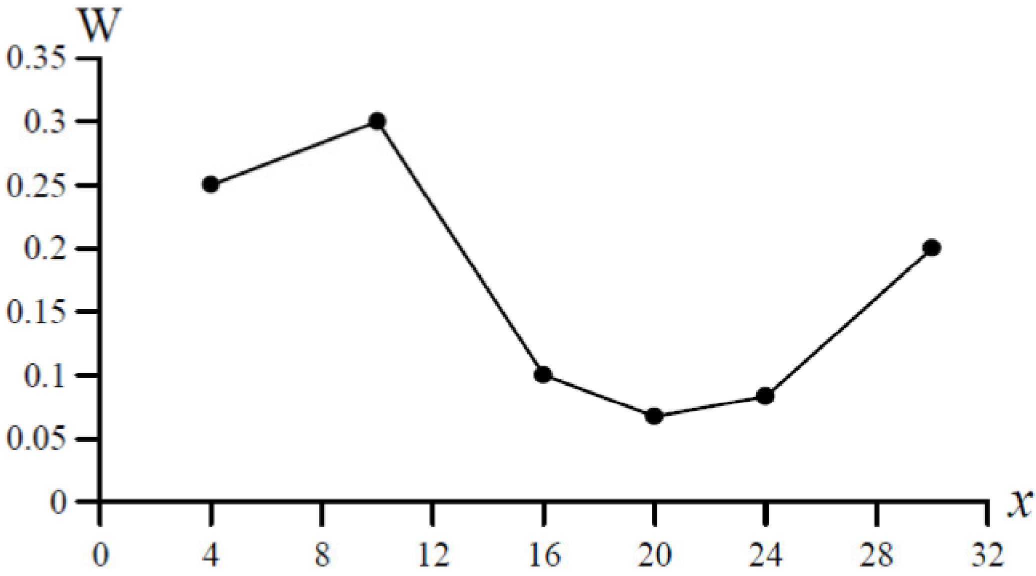


Рис. 21

Статистическим аналогом функции распределения случайной величины X ($F(x)=P(X<x)$) является эмпирическая функция распределения $F^*(x)=W(X<x)$, задающая для каждого $x\in R$ относительную частоту события ($X<x$). При этом считается, что $W(X=x_i)=w_i$, где w_i определены по выборке. В рассмотренном выше примере $F^*(x)$ задается формулой

$$F^*(x)=\begin{cases} 0, & x\le 4; \\ 0{,}25, & 4<x\le 10; \\ 0{,}55, & 10<x\le 16; \\ 0{,}65, & 16<x\le 20; \\ 0{,}717, & 20<x\le 24; \\ 0{,}8, & 24<x\le 30; \\ 1, & x>30. \end{cases}$$

и имеет график, представленный на рис. 22.

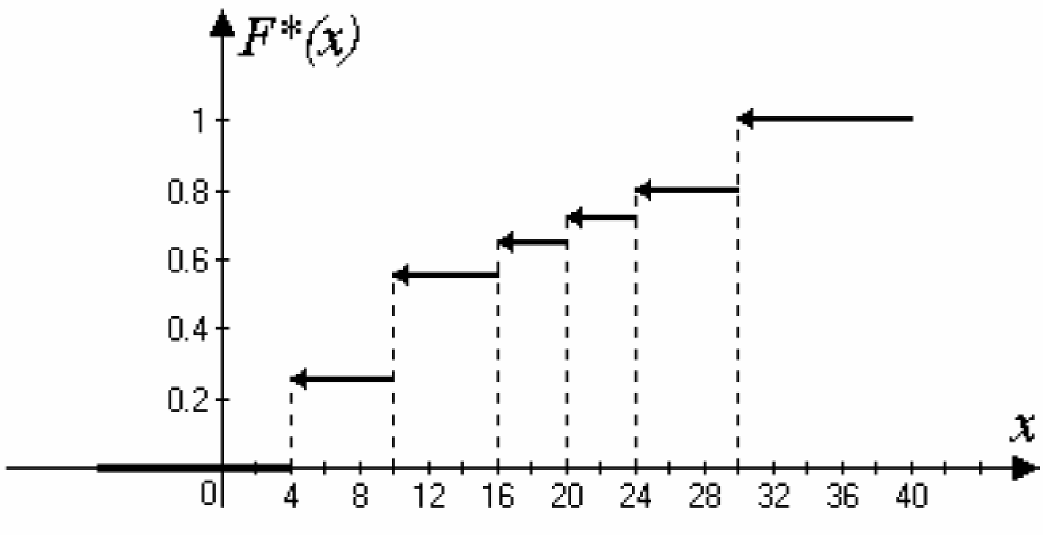


Рис. 22

$F^*(x)$ обладает такими же свойствами, как и функция распределения $F(x)$ дискретной случайной величины.

Если вариационный ряд содержит достаточно много различных вариант, то их разбивают на группы, принадлежащие отдельным интервалам. Число интервалов m должно быть не очень большим, но и не очень малым, чтобы не потерять особенности распределения величины X. Число интервалов для выборки объема n рекомендуется вычислять согласно формуле Стерджеса:

$$m \approx 1+3{,}322 \lg n .$$

При найденном таким образом числе m величина интервала h определяется по формуле

$$h = \frac{x_{\max} - x_{\min}}{1 + 3{,}322 \lg n},$$

где $x$max – наибольшая варианта, а $x$min – наименьшая. Здесь под частотой ni понимается число вариант, принадлежащих i-му интервалу, а под wi=ni/n – относительная частота. Иногда ni и wi называют весами. Соответствие между построенными интервалами и частотами (относительными частотами) называют интервальным распределением частот (относительных частот). Очевидно, для непрерывной случайной величины X всегда составляют по выборке интервальное распределение (практически все варианты различны). По интервальному распределению строят гистограмму относительных частот, задающую на каждом интервале плотность относительной частоты. Гистограмма состоит из прямоугольников, основаниями которых являются построенные нами интервалы, а высоты равны отношениям соответствующих относительных частот к ширине интервала, т. е. для i-го интервала высота равна wi/h. Нетрудно видеть, что гистограмма представляет собой статистический аналог плотности распределения вероятностей.

Пример. Изучается изменение выработки на одного рабочего цеха в отчетном году по сравнению с предыдущим (в %). Для выборки, состоящей из 100 рабочих, получены варианты, по которым составлено интервальное распределение. Так как n=100, то надо взять восемь интервалов (m≈1+3,322 lg100=7,644). В вариационном ряду $x$min=97 %, $x$max=141 %, поэтому ширина интервала
h=(141–97)/8=5,74 %. Возьмем h=6,0 %. Округление шага приводит к тому, что mh=8·6=48% > $x$max–$x$min = 141-97=44 %. Отрезок длины mh=48 накроет все варианты выборки, если взять, например, за начало первого интервала
xнач=xmin–h/2=97–3=94 %. Теперь интервальное распределение примет вид, представленный в табл. 3.

Таблица 3

Номер интервала i	Границы i-го интервала	Частота n_i	Относительная частота	Накопленная частота	Накопленная относительная частота

	(в %)				
1	94-100	3	0,03	3	0,03
2	100-106	7	0,07	10	0,10
3	106-112	11	0,11	21	0,21
4	112-118	20	0,20	41	0,41
5	118-124	28	0,28	69	0,69
6	124-130	19	0,19	88	0,88
7	130-136	10	0,10	98	0,32
8	136-142	2	0,02	100	1,00
	∑	100	1,00	-	-

По данному интервальному распределению построим гистограмму (рис. 23). Известно, что по графику функции плотности вероятности можно найти вероятность попадания значений случайной величины X в заданный интервал (a, b) как площадь криволинейной трапеции. Аналогично по гистограмме можно приближенно найти такую вероятность.

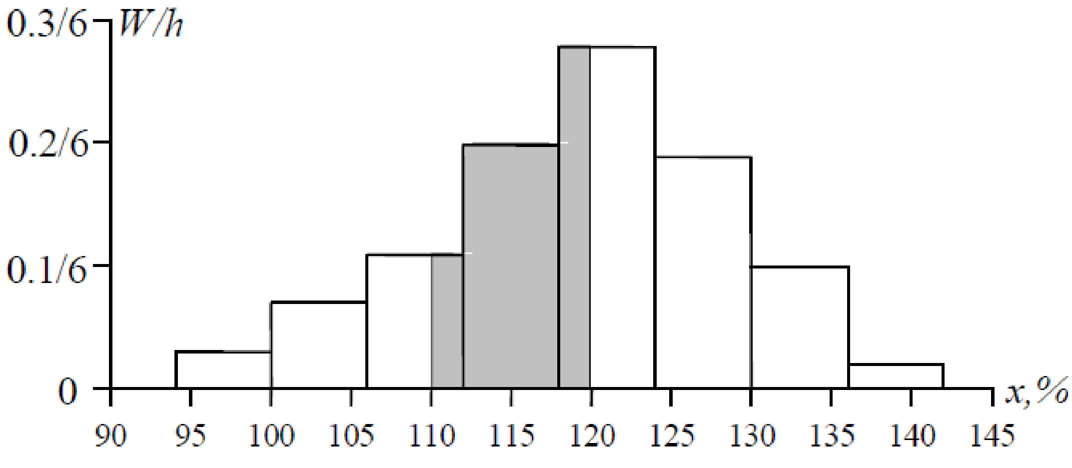


Рис. 23

Например, вероятность того, что изменение выработки X не более 120 % и не менее 110 %, приближенно равно площади заштрихованной фигуры на рис. 23, т.е.

p=(2□0,11+6□0,20+2□0,28):6=1,98:6=0,33.

По интервальному распределению вместо эмпирической функции распределения $F^*(x)$ строят непрерывную функцию, график которой называется кумулятой (кумулятивной кривой), или кривой накопленных частот (относительных частот). Например, кумулятивная кривая относительных частот представляет собой ломаную, соединяющую последовательно точки с координатами (xi, wiнак), где xi – правый конец i-го интервала (i=0,1,…,m), а $w\mathrm{i}^{\text{нак}}$ – накопленная относительная частота ($w\mathrm{i}^{\text{нак}}$= $\sum_{i=1}^{k} w_i$), причем xo – начало первого интервала и $w\mathrm{o}^{\text{нак}}$=0. Для нашего примера кумулята имеет вид, представленный на рис. 24.

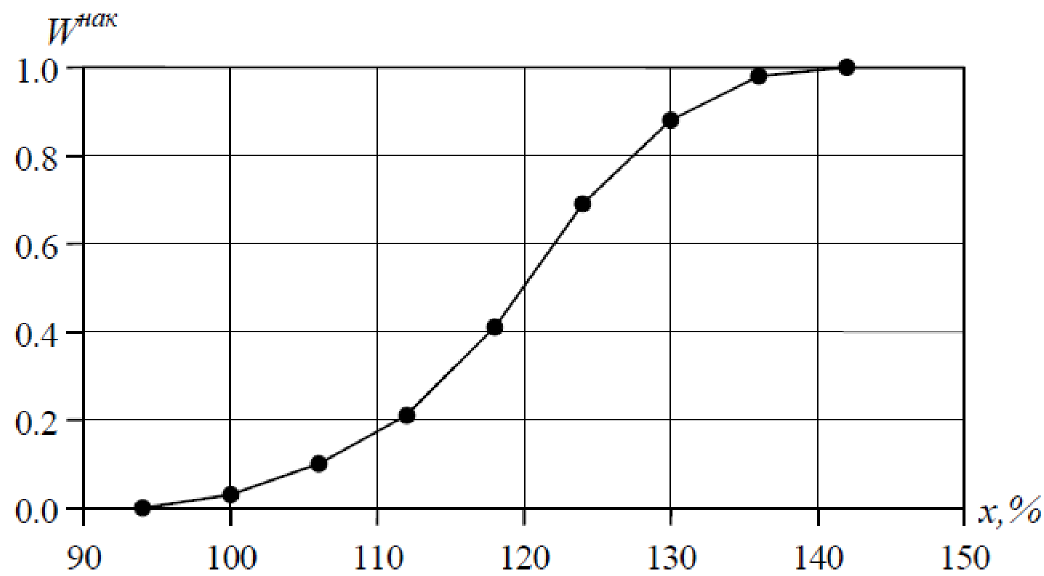


Рис. 24

От интервального распределения иногда переходят к статистическому распределению частот, взяв в качестве вариант середины интервалов (xi – середина i-го интервала). Для рассмотренного примера статистическое распределение представлено в табл. 4.

Таблица 4

xi	97	103	109	115	121	127	133	139	$\sum$
n_i	3	7	11	20	28	19	10	2	100
w_i	0,03	0,07	0,11	0,20	0,28	0,19	0,10	0,03	1

Числовые характеристики выборки

Пусть имеется выборка объема n и найдены варианты xi (i=1,…, n). Аналогами числовых характеристик случайной величины являются выборочные числовые характеристики.

Определение 1. Выборочной средней x называется среднее арифметическое всех вариант

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i . \qquad (11.1)$$

Если варианты сгруппированы, то $\bar{x}$ вычисляется по формуле

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i n_i . \qquad (11.2)$$

где xi – варианты или середины интервалов в интервальном распределении, а ni – соответствующие частоты ($\sum_{i=1}^{k} n_i = n$).

Из (11.2) следует, что $\bar{x} = \sum_{i=1}^{k} x_i w_i$ т. е. $\bar{x}$ является аналогом математического ожидания случайной величины ($M(X)=\Sigma xipi$). Свойства выборочной средней полностью соответствуют свойствам математического ожидания.

Определение 2. Медианой m вариационного ряда называют значение x, слева и справа от которого находится одинаковое число вариант. Если n=2k+1, то m=xk+1; если n=2k, то m=1/2(xk+xk+1).

Например, для статистического распределения выборки, заданного табл. 4, n=100, x50=121 и x51=121, поэтому m=121.

Определение 3. Модой M вариационного ряда называют варианту, которой соответствует наибольшая частота.

Например, для вариационного ряда, заданного табл. 2, M=10, а для вариационного ряда, заданного табл. 3, M=121.

Разброс вариант отражают показатели вариации, простейшим из которых является вариационный размах R=$x$max–$x$min.

Определение 4. Выборочной дисперсией называется среднее арифметическое квадратов отклонений вариант от выборочной средней

$$D_в = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2. \qquad (11.3)$$

Для сгруппированных данных

$$D_в = \frac{1}{n}\sum_{i=1}^{k}\left(x_i - \overline{x}\right)^2 n_i. \qquad (11.4)$$

Выборочным средним квадратическим отклонением называется

$$\sigma_x = \sqrt{D_в}. \qquad (11.5)$$

Видно, что выборочная дисперсия Dв и выборочное среднее квадратическое отклонение σx являются аналогами дисперсии D(X) и среднего квадратического отклонения σ(X) случайной величины X.

Рассматривают также коэффициент вариации ν – безразмерную величину, равную процентному отношению выборочного среднего квадратического отклонения к выборочной средней

$$\nu = \frac{\sigma_x}{\bar{x}} \cdot 100\% . \qquad (11.6)$$

Отметим, что свойства выборочной дисперсии *Dв* аналогичны свойствам дисперсии $D(X)$ случайной величины. В частности, формула для вычисления *Dв* может быть записана в виде

$$D_в = \overline{x^2} - (\bar{x})^2 . \qquad (11.7)$$

где $\bar{x}$ – выборочная средняя, а $\overline{x^2}$ – выборочная средняя квадратов вариант, т.е.

$$D_в = \frac{1}{n}\sum_{i=1}^{k}(x_i)^2 n_i - \left(\frac{1}{n}\sum_{i=1}^{k} x_i n_i\right)^2 . \qquad (11.8)$$

Статистические оценки параметров распределения

Пусть выборочным методом изучается случайная величина *X*, характеризуемая параметром Θ (Θ может быть и числовой характеристикой величины *X*). Тогда по выборке с помощью некоторой формулы вычисляется значение $\widetilde{\Theta}_n$ (*n* – объем выборки), называемое статистической оценкой параметра Θ. Так как выборка проводится случайно, то каждое значение (в зависимости от выборки) будет случайной величиной. Иными словами, оценка параметра Θ – это случайная величина $\widetilde{\Theta}_n$. Например, в качестве оценки математического ожидания $M(X)$можно взять выборочную среднюю x , или медиану m , или моду *M*. Какая же из них "лучше"? Будем считать оценку $\widetilde{\Theta}_n$ наилучшей для параметра Θ, если она обладает следующими качествами.

Определение 1. Оценка параметра Θ называется несмещенной, если выполняется условие $M(\tilde{\Theta}_n)=\Theta$. В противном случае оценку называют смещенной. Смещенная оценка в среднем дает значение либо больше Θ, либо меньше. Несмещенность оценки гарантирует отсутствие систематических ошибок при оценивании.

Определение 2. Оценка $\tilde{\Theta}_n$ параметра Θ называется состоятельной, если при $n\to\infty$ $\tilde{\Theta}_n$ по вероятности стремится к оцениваемому параметру Θ, т.е. для любого $\varepsilon>0$ $\lim\limits_{n\to\infty} P\left(\left|\tilde{\Theta}_n-\Theta\right|<\varepsilon\right)=1$.

Состоятельность оценки позволяет при оценивании сделать маловероятными значительные ошибки путем увеличения объема выборки.

Определение 3. Несмещенная оценка $\tilde{\Theta}_n$ параметра Θ является эффективной, если ее дисперсия $D(\tilde{\Theta}_n)$ будет наименьшей по сравнению с дисперсиями всех возможных несмещенных оценок параметра Θ, вычисляемых по выборкам одного и того же объема *n*.

Существует много методов нахождения оценок, например, метод моментов (предложенный К. Пирсоном), метод максимального правдоподобия (предложенный Р. Фишером), метод наименьших квадратов [1,7,9]. Здесь мы не будем останавливаться на сути этих методов.

Для математического ожидания $M(X)$ несмещенной и состоятельной оценкой является средняя выборочная $\bar{x}$. Для дисперсии $D(X)$ состоятельной оценкой является выборочная дисперсия $D_в$. В то же время известно, что $D_в$ – смещенная оценка $D(X)$, а именно $M(D_в)=\frac{n-1}{n}D(X)$. Чтобы получить несмещенную оценку, выборочную дисперсию исправляют, рассматривая так называемую исправленную выборочную дисперсию S^2:

$$S^2=\frac{n}{n-1}D_в.$$

Очевидно, $M(S^2)=\dfrac{n}{n-1}M(D_в)=\dfrac{n}{n-1}\cdot\dfrac{n-1}{n}D(X)=D(X)$, т.е. S^2 – несмещенная и состоятельная оценка дисперсии $D(X)$.

Оценки, о которых шла речь, называют точечными, так как для каждой выборки значение оценки является одним числом. Точечная оценка представляет собой как бы приближенное значение оцениваемого параметра и не дает информации о точности этого приближенного значения.

Чтобы оценить параметр с некоторой точностью, используют интервальную оценку параметра.

Определение 4. Интервальной оценкой параметра Θ называется интервал $\left(\widetilde{\Theta}_n^1, \widetilde{\Theta}_n^2\right)$, который с некоторой вероятностью γ накрывает оцениваемый параметр Θ, т.е. $P\left(\widetilde{\Theta}_n^1 < \Theta < \widetilde{\Theta}_n^2\right) = \gamma$.

Границы $\widetilde{\Theta}_n^1$ и $\widetilde{\Theta}_n^2$ интервала находятся по выборке, поэтому являются случайными величинами, т.е. интервал $\left(\widetilde{\Theta}_n^1, \widetilde{\Theta}_n^2\right)$ накрывает параметр Θ случайно. Указанный интервал называют доверительным, а вероятность γ называют доверительной вероятностью или надежностью оценки. Обычно построение доверительного интервала для параметра Θ начинается с отыскания его точечной оценки $\widetilde{\Theta}_n$. После этого доверительный интервал строится симметричным относительно найденной оценки, т.е. в виде

$$\left(\widetilde{\Theta}_n - \Delta, \widetilde{\Theta}_n + \Delta\right).$$

Радиус искомого доверительного интервала определяется по заданной надежности γ так, чтобы выполнялось условие

$$P\left(\left|\widetilde{\Theta}_n - \Theta\right| < \Delta\right) = \gamma. \qquad (11.9)$$

Например, если известно, что исследуемая случайная величина нормально распределена, то ее параметры a и σ оцениваются, исходя из их точечных оценок: для параметра a – это выборочная средняя $\bar{x}$, для параметра σ – это квадратный корень из исправленной выборочной дисперсии $\sqrt{S^2} = S$. Для выборки объема n и заданной надежности γ доверительные интервалы для параметров a и σ имеют вид

$$\bar{x} - t(\gamma, n)\frac{S}{\sqrt{n}} < a < \bar{x} + t(\gamma, n)\frac{S}{\sqrt{n}}, \qquad (11.10)$$

$$S(1 - q(\gamma, n)) < \sigma < S(1 + q(\gamma, n)), \qquad (11.11)$$

Вывод формул (10.10), (10.11) здесь не приводится (см., например, [1, 3, 6]). Значение $t(\gamma,n)$ и $q(\gamma,n)$ находят по специальным таблицам (см. приложения 3 и 4). Если доверительный интервал найден из условия (10.9), то радиус Δ этого интервала будет точностью оценки $\tilde{\Theta}_n$ параметра Θ с надежностью γ.

Понятие о корреляционном и регрессионном анализе

Во многих задачах нас интересуют случайные величины в их взаимосвязи друг с другом. Ранее определялось понятие зависимых и независимых случайных величин. Уточним понятие зависимости между случайными величинами. Пусть рассматриваются две случайные величины X и Y. Если каждому значению x величины X соответствует одно значение y величины Y, то $Y = g(X)$ – функциональная зависимость между величинами. Например, Y (кг) – масса кирпича, X(см3) – объем кирпича, тогда $Y=d\times X$, где d (*кг/см*3) – плотность материала, из которого изготовлен кирпич. Если каждому значению x величины X соответствует конкретный закон распределения величины Y (условное распределение), зависящий от x, то зависимость между величинами X и Y называется стохастической. Для каждого значения x математическое ожидание величины Y зависит от x, его называют условным математическим ожиданием и обозначают $MX(Y)$ или $M(Y/X=x)$. Функциональную зависимость условного математического ожидания одной случайной величины от значений другой называют корреляционной зависимостью между этими величинами, т.е.

$$MX(Y)=g(x) \text{ или } MY(X)=h(y), \qquad (12.1)$$

($g(x) \neq$const и $h(y) \neq$ const). Уравнения (12.1) называют уравнениями регрессии Y на X и X на Y соответственно. Функции g(x) и $h(y)$ называют функциями регрессии, а их графики – линиями регрессии. Например, между производительностью труда и энерговооруженностью на предприятии существует корреляционная зависимость. Далее рассмотрим числовые характеристики пары случайных величин, по значениям которых можно судить о виде зависимости между этими величинами.

Коэффициент корреляции

Пусть рассматриваются две случайные величины X и Y. Коэффициентом корреляции этих величин называется безразмерная величина, обозначаемая $r=r(X,Y)$ и определяемая по формуле

$$r=\frac{M\big((X-M(X))\cdot(Y-M(Y))\big)}{\sigma(X)\cdot\sigma(Y)}. \qquad (12.2)$$

Коэффициент корреляции обладает следующими свойствами:

Свойство 1. Коэффициент корреляции по модулю не превышает единицы, т.е. $r \le 1$.

Свойство 2. Если X и Y независимы (в частности, если X или Y постоянны), то $r=0$.

Свойство 3. Если X и Y связаны линейной функциональной зависимостью, $Y=aX+b(a\neq 0)$, то $|r|=1$, причем знак r совпадает со знаком a.

Таким образом, коэффициент корреляции представляет собой число, по величине которого судят, сколь «тесна» линейная корреляционная зависимость между X и Y. Если $|r|\neq 1$, то линейной функциональной зависимости между X и Y нет. Чем ближе $|r|$ к единице, тем теснее линейная корреляционная зависимость между X и Y (функции регрессии близки к линейным); чем ближе $|r|$ к 0, тем слабее линейная корреляционная зависимость между X и Y.

Существуют и другие числовые характеристики пары случайных величин, например, корреляционный момент [1].

На практике зависимость между величинами изучается статистическими методами. Установлением вида зависимости и ее тесноты занимается корреляционный анализ. Задачами регрессионного анализа является отыскание функций регрессии, оценка параметров этих функций и т.д.

Выборочное уравнение прямой регрессии

Пусть по случайной выборке объема n изучается зависимость между величинами X и Y, характеризующими объекты некоторой генеральной совокупности. Полученные пары вариант (xi, yi) $(i = 1, \dots, n)$ удобно сводить в корреляционную таблицу. Если различных пар вариантов немного, то корреляционная таблица выглядит так:

X_i \ Y_j	y_1	...	y_m	n_i^x
x_1	n_{11}	...	n_{1m}	n_1^x
.	.	...	.	.
.	.	...	.	

.	.	...	.	.
x_k	n_{k1}	...	n_{km}	. n_k^x
n_j^y	n_1^y	...	n_m^y	$\sum = n$

Здесь

xi	$x1$	...	xk
n_i^x	n_1^x	...	n_k^x

и

y_j	$y1$	...	y_m
n_j^y	n_1^y	...	n_m^y

– статистические распределения частот для величин X и Y соответственно, а nij – частоты пар (xi, yj). Если пары практически не повторяются, то для X и Y составляют интервальные распределения частот, строят такую же корреляционную таблицу, беря в качестве (xi, yj) середины соответствующих интервалов, а в качестве nij – частоты пар вариантов, компоненты которых попадают в эти интервалы. Пары вариантов, найденные по выборке, изображают точками на координатной плоскости. Полученное множество точек называется полем корреляции.

Пусть нас интересует вопрос о том, сколь тесна линейная корреляционная зависимость между величинами X и Y. Естественно попытаться найти коэффициент корреляции r для этих величин. Но по выборке можно лишь найти оценку для r. Для нахождения оценки r удобно переписать формулу для вычисления r, воспользовавшись свойствами математического ожидания:

$$r = \frac{M\left(\left(X - M(X)\right)\cdot\left(Y - M(Y)\right)\right)}{\sigma(X)\cdot\sigma(Y)} = \frac{M\left(X\cdot Y\right) - M(X)\cdot M(Y)}{\sigma(X)\cdot\sigma(Y)}. \qquad (12.3)$$

Подставляя в формулу (11.3) вместо математических ожиданий их оценки (выборочные средние и выборочные средние квадратические отклонения) получим оценку коэффициента корреляции

$$r_в = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\sigma_x \cdot \sigma_y}. \qquad (12.4)$$

Здесь $\overline{x}$, $\overline{y}$ – выборочные средние величин X и Y; σ x , σ y – выборочные средние квадратические отклонения этих величин; $\overline{xy}$ – выборочная средняя произведения XY: $\overline{xy} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij} x_i y_j$. Оценка *rв* называется выборочным коэффициентом корреляции. Свойства *rв* аналогичны свойствам коэффициента корреляции r, поэтому по его величине судят о тесноте линейной корреляционной зависимости между X и Y (чем ближе $|r_в|$ к единице, тем теснее эта зависимость). Если оказалось, что линейная корреляционная связь достаточна тесная (т.е. функции регрессии близки к линейным), то по корреляционной таблице находят оценки этих функций, называемые выборочными функциями регрессии.

Линейные уравнения регрессии имеют вид

$$Mx\,(Y) = ax + b \text{ и } My\,(X) = cy + d. \qquad (12.5)$$

Чтобы получить оценки этих уравнений, заменим условные математические ожидания $Mx\,(Y)$ и $My\,(X)$ их оценками, которые называются условными средними и обозначаются $\overline{y}_x$ и $\overline{x}_y$. Тогда выборочные уравнения прямых регрессии будут иметь вид

$$\overline{y}_x = ax + b \quad \text{и} \quad \overline{x}_y = cy + d. \qquad (12.6)$$

Неизвестные коэффициенты a, b, c, d уравнений (12.6) находятся методом наименьших квадратов. По корреляционной таблице для каждого xi ищут среднее значение $\overline{y}_i$ условной средней $\overline{y}_x$ по формуле

$$\overline{y_i} = \frac{1}{n_i^x}\sum_{j=1}^{m} n_{ij} y_j \,. \qquad (12.7)$$

Аналогично, для каждого yj находят среднее значение соответствующих значений $\overline{x}_j$ условной средней $\overline{x}_y$:

$$\overline{x_j} = \frac{1}{n_j^y}\sum_{i=1}^{k} n_{ij} x_i \,. \qquad (12.8)$$

Таким образом, получают две таблицы

xi	$x1$	…	xk
$\overline{y}_i$	$\overline{y}_1$	…	$\overline{y}_k$

и (12.9)

y_j	$y1$	…	y_m
$\overline{x}_j$	$\overline{x}_1$	…	$\overline{x}_m$

По данным (12.9) ищут коэффициенты a, b и c, d функций (12.6), наилучшие в смысле метода наименьших квадратов. Не приводя здесь вычислений этих коэффициентов, выпишем полученные уравнения выборочных прямых регрессии в виде, отличном от (12.6):

$$\overline{y}_x - \overline{y} = r_в \frac{\sigma_y}{\sigma_x}\left(x - \overline{x}\right), \qquad (12.10)$$

$$\bar{x}_y - \bar{x} = r_в \frac{\sigma_x}{\sigma_y}(y - \bar{y}), \qquad (12.11)$$

Значения x, y, $\sigma\ x$, $\sigma\ y$, $rв$ были найдены ранее при вычислении выборочного коэффициента корреляции $rв$. Если полученные прямые построить в плоскости, на которой изображено поле корреляции, то при тесной линейной корреляционной зависимости угол между прямыми будет мал и облако точек (xi, yi) будет плотно примыкать к этим прямым.

Если выборочный коэффициент корреляции $rв$ близок к нулю, то линейной корреляционной зависимости нет. В этом случае пытаются построить нелинейные выборочные уравнения регрессии или установить, что величины независимы.

Пример. Изучается зависимость между суточной выработкой продукции Y (т) и величиной основных производственных фондов X (млн р.) для совокупности 50 однотипных предприятий. Выборочные данные изображены полем корреляции на рис. 25 и сведены в табл. 5.

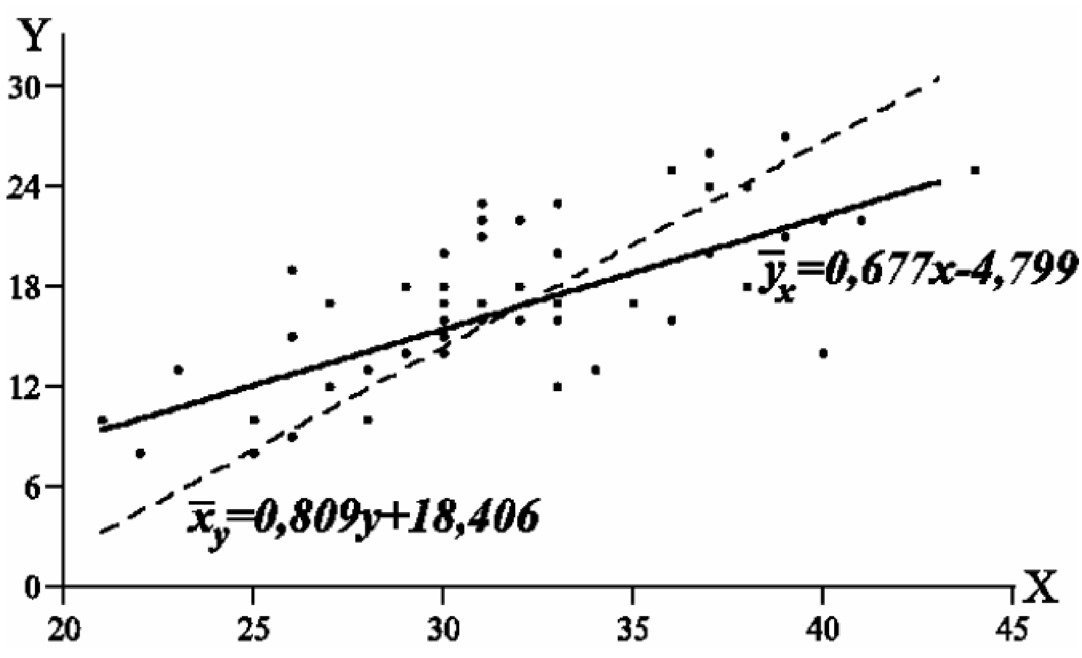


Рис. 25

Таблица 5

Величина ОПФ млн р.(X)	Суточная выработка продукции, т (Y) 7-11	11-15	15-19	19-23	23-27
20-25	2	1	-	-	-
25-30	3	6	4	-	-
30-35	-	3	11	7	-
35-40	-	1	2	6	2
40-45	-	-	-	1	1

Требуется оценить тесноту линейной корреляционной связи между величинами X и Y и составить выборочные уравнения прямых регрессии Y на X и X на Y . Составим корреляционную табл. 6, выбирая в качестве (xi, yj) середины заданных в таблице интервалов. Имеем

Таблица 6

x_i \ y_j	9	13	17	21	25	n_i^x
22,5	2	1	-	-	-	3
27,5	3	6	4	-	-	13
32,5	-	3	11	7	-	21
37,5	-	1	2	6	2	11
42,5	-	-	-	1	1	2
n_j^y	5	11	17	14	3	n=50

Для оценки тесноты линейной корреляционной связи между величинами X и Y найдем выборочный коэффициент корреляции . Для этого вычислим

$$\bar{x}=\frac{1}{50}(22{,}5\cdot 3+27{,}5\cdot 13+32{,}5\cdot 21+37{,}5\cdot 11+42{,}5\cdot 2)=32{,}1;$$

$$\bar{y}=\frac{1}{50}(9\cdot 5+13\cdot 11+17\cdot 17+21\cdot 14+25\cdot 3)=16{,}92;$$

$$\overline{x^2}=\frac{1}{50}(22{,}5^2\cdot 3+27{,}5^2\cdot 13+32{,}5^2\cdot 21+37{,}5^2\cdot 11+42{,}5^2\cdot 2)=1052{,}25;$$

$$\overline{y^2}=\frac{1}{50}(9^2\cdot 5+13^2\cdot 11+17^2\cdot 17+21^2\cdot 14+25^2\cdot 3)=304{,}52;$$

$$\sigma_x=\sqrt{\overline{x^2}-(\bar{x})^2}=\sqrt{1052{,}25-32{,}1^2}\approx 4{,}67;$$

$$\sigma_y=\sqrt{\overline{y^2}-(\bar{y})^2}=\sqrt{304{,}52-16{,}92^2}\approx 4{,}27;$$

$$\overline{xy}=\frac{1}{50}(9\cdot 22{,}5\cdot 2+13\cdot 22{,}5\cdot 1+9\cdot 27{,}5\cdot 3+13\cdot 27{,}5\cdot 6+17\cdot 27{,}5\cdot 4+$$
$$+13\cdot 32{,}5\cdot 3+17\cdot 32{,}5\cdot 11+21\cdot 32{,}5\cdot 7+13\cdot 37{,}5\cdot 1+17\cdot 37{,}5\cdot 2+$$
$$+21\cdot 37{,}5\cdot 6+25\cdot 37{,}5\cdot 2+21\cdot 42{,}5\cdot 1+25\cdot 42{,}5\cdot 1)=557{,}9.$$

Подставив найденные значения в формулу (12.4), получим

$$r_в=\frac{557{,}9-32{,}1\cdot 16{,}92}{4{,}67\cdot 4{,}27}\approx 0{,}74$$

Так как коэффициент $|r_в|$ близок к единице, то можно сделать вывод, что линейная корреляционная зависимость между X и Y достаточно тесная.

Теперь запишем выборочные уравнения прямых регрессии, используя формулы (12.10) и (12.11):

$$\bar{y}_x-16{,}92=0{,}74\cdot\frac{4{,}27}{4{,}67}\left(x-32{,}1\right),$$

$$\bar{x}_y-32{,}1=0{,}74\cdot\frac{4{,}67}{4{,}27}\left(y-16{,}92\right),$$

или

$y_x = 0{,}677x - 4{,}799$, $x_y = 0{,}809y + 18{,}406$.

Найденные прямые регрессии изображены в поле корреляции на рис. 25. Из полученных уравнений видно, что увеличение основных производственных фондов на 1 млн р. приводит к среднему увеличению суточной выработки продукции на 0,68 т, а для увеличения суточной выработки продукции на 1 т нужно в среднем увеличить производственные фонды на 0,81 млн р.

Билеты к Экзамену (Учебным планом предусмотрен ТОЛЬКО **зачет**)

РГР(тексты) Расчетно-графическое задание по теории вероятностей и математической статистике выполняется по материалам «Методических указаний» N 674, (авторы Акчурина Л.В., Кущев А.Б., Ханкин Е.И.) имеющихся в достаточном количестве в библиотеке ВГАСУ.