СЕВЕРО-ЗАПАДНЫЙ ГОСУДАРСТВЕННЫЙ ЗАОЧНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

МАТЕМАТИКА
Часть 2
Теория вероятностей и элементы математической статистики

УЧЕБНО - МЕТОДИЧЕСКИЙ КОМПЛЕКС

Санкт-Петербург
2009

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение высшего профессионального образования

«СЕВЕРО-ЗАПАДНЫЙ ГОСУДАРСТВЕННЫЙ ЗАОЧНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»
Кафедра информатики и прикладной математики

МАТЕМАТИКА

Часть 2

Теория вероятностей и элементы математической статистики

УЧЕБНО - МЕТОДИЧЕСКИЙ КОМПЛЕКС

Институты: все
Укрупненные группы специальностей и направлений подготовки:
080000 –Экономика и управление
140000 – Энергетика, энергетическое машиностроение и электротехника
150000 – Металлургия, машиностроение и материалообработка
190000 – Транспортные средства
200000 – Приборостроение и оптотехника
210000 – Электронная техника, радиотехника и связь
220000 – Автоматика и управление
230100 – Информатика и вычислительная техника
240000 – Химия и биотехнологии
Направления подготовки высшего профессионального образования:
261000 – Технология художественной обработки материалов
280200 – Защита окружающей среды

Санкт-Петербург
Издательство СЗТУ
2009

УДК 519.2.06(07)

Математика. Часть 2. Теория вероятностей и элементы математической статистики: учебно-методический комплекс / сост. В.С. Ходоровская. - СПб.: Изд - во СЗТУ, 2009. – 191 с.

Учебно-методический комплекс разработан в соответствии с государственными образовательными стандартами высшего профессионального образования.

Дисциплина предусматривает изучение понятий и варианты расчета вероятностей случайных событий, случайных величин и их характеристик, рассматриваются элементы математической статистики и варианты проверки различных гипотез, выдвигаемых при работе со случайными величинами. Приводятся теоретические основы и разбираются варианты решения задач, формулируются тестовые вопросы к рассматриваемым темам.

Рассмотрено на заседании кафедры информатики и прикладной математики 22 декабря 2008 г., одобрено методической комиссией факультета общепрофессиональной подготовки 22 декабря 2008 г.

Рецензенты: кафедра информатики и прикладной математики СЗТУ (зав. кафедрой Г.Г.Ткаченко, канд. физ.-мат. наук, доцент); Л.В.Боброва, канд.техн.наук, проф. кафедры информатики и прикладной математики СЗТУ.

Составители: В.С. Ходоровская, доц.;
Т.Д. Бессонова, доц.;
Г.Г. Ткаченко, канд.физ.-мат. н., доц.;
М.Б.Шабаева, канд.физ.наук., доц.

1. Информация о дисциплине

Содержание курса «Теория вероятностей и элементы математической статистики» определяется государственным стандартом, рабочей программой и тематическим планом.

Данная дисциплина изучается студентами на 2-м или 3-м курсе в течение семестра. Программа курса включает изучение трех основных теоретических разделов: 1) случайные события; частота и вероятность; основные формулы для вычисления вероятностей, 2) случайные величины; числовые характеристики дискретной и непрерывной случайных величин; нормальный закон распределения, 3) элементы математической статистики, генеральная совокупность и выборка; оценки параметров; корреляция и регрессия, содержание которых представлено в виде структурно-логической схемы на рис. 2.1. Кроме того, включен раздел с указаниями по выполнению лабораторных работ и раздел с методическими указаниями к выполнению контрольной работы.

Изучение курса заканчивается выполнением контрольной работы, ответом на вопросы тестов и сдачей экзамена.

1.1. Предисловие

Учебно-методический комплекс **Математика. Часть 2 “Теория вероятностей и элементы математической статистики”** предназначен для всех форм обучения студентов всех специальностей подготовки высшего профессионального образования. Программой курса предусматривается изучение случайных событий и случайных величин, приводятся формулы, позволяющие производить расчет вероятностей различных случайных событий, характеристики случайных величин и их законы распределения, рассматриваются элементы математической статистики.

Данная дисциплина изучается студентами в течение одного семестра.

Рабочая программа включает изучение разделов по темам: «Случайные события», “ Случайные величины“, “Элементы математической статистики“.

План практических занятий предусматривает изучение тем:

1. Случайные события, частота и вероятность. Классическое определение вероятности. Условные вероятности.

2. Основные формулы для вычисления вероятностей: схема Бернулли, полной вероятности и Байеса.

3. Случайные величины. Числовые характеристики дискретной и непрерывной случайных величин.

В план выполнения лабораторных работ входят темы:

1. Генеральная совокупность и выборка. Обработка выборочных данных. Точечные и интервальные оценки параметров. Построение гистограмм

2. Моделирование дискретных случайных величин методом жребия.

3. Проверка гипотез по критерию Пирсона. Корреляция и регрессия.

Контроль знаний студента осуществляется по результатам выполнения контрольной работы, тестирования и заканчивается сдачей экзамена.
Целью изучения дисциплины является приобретение знаний и навыков решения задач, связанных с расчетом вероятностей и числовых характеристик случайных величин и случайных событий.
Задачи изучения дисциплины - правильное понимание основных положений и формул для расчета вероятностей различных событий.
Иметь представление о свойствах вероятности.
Знать основные определения и формулы для расчета.
Уметь правильно использовать расчетную формулу для конкретно решаемой задачи.
Владеть методикой расчета определения вероятности правильной работы прибора или устройства.
Место дисциплины. Математика часть 2 - в учебном процессе тесно связана с курсом Математика часть 1, курсом «Информатика» и является базовой основой для различных курсов, в которых необходимо производить обработку статистических данных и делать оценки величин при выполнении дипломных проектов.

1.2. Содержание дисциплины и виды учебной работы

Теория вероятностей и математическая статистика: случайные события, частота и вероятность, основные формулы для вычисления вероятностей; случайные величины; числовые характеристики дискретной и непрерывной случайных величин, нормальный закон распределении; генеральная совокупность и выборка; оценки параметров; корреляция и регрессия.

Объем дисциплины и виды учебной работы

Вид учебной работы	Всего часов форма обучения		
	очная	очно-заочная	заочная
Общая трудоемкость дисциплины (ОТД)	**150**		
Работа под руководством преподавателя (включая ДОТ)	**90**	**90**	**90**
В том числе аудиторные занятия: лекции практические занятия (ПЗ) лабораторные работы (ЛР)	 **28** **4** **20**	 **12** **8** **12**	 **6** **4** **10**
Самостоятельная работа студента (СР)	**60**	**60**	**60**
Промежуточный контроль, количество В том числе: курсовой проект (работа) контрольная работа	**1** **-** **-**	**1** **-** **1**	**1** **-** **1**
Вид итогового контроля (зачет, экзамен)	**Экзамен**		

Перечень видов практических занятий и контроля:

- тест (общий по дисциплине);
- одна контрольная работа (для очно-заочной и заочной форм обучения);
- практические занятия – 4 часа (для очной и заочной форм обучения),
  - 8 часов (для очно-заочной формы обучения);
- лабораторные работы – 20 часов (для очной формы), 12 часов (для очно-заочной) и 10 часов (для заочной формы обучения);
- экзамен.

2. Рабочие учебные материалы

2.1. Рабочая программа (объем 150 часов)

Введение

[1], с. 5; [2], с. 3-5

Теория вероятностей и математическая статистика – неразрывно связанные науки, изучающие закономерности случайных явлений. Математическая статистика – наука, которая разрабатывает методы обработки и анализа результатов наблюдений и опытов на основе теоретико-вероятностных понятий и методов для того, чтобы получить некоторые научные и практические выводы. Необходимость обрабатывать большие объемы опытных и статистических данных возникает в технике, экономике, медицине, финансах и т.д. Появляются задачи, для решения которых требуются вероятностно-статистические методы. К таким задачам можно отнести, например, задачи упорядочения результатов измерения, выборочного контроля качества, исследования надежности работы сложных систем.

Предметом изучения теории вероятностей служат случайные события и случайные величины, над которыми производятся многократные наблюдения, в результате чего делаются выводы и обобщения о числовых характеристиках, определяется закон распределения. Статистика достаточно древняя наука. Математической наукой она признана после того, как стала строиться на базе теории вероятностей. Одна из главных задач статистических методов заключается в рассмотрении выборки из генеральной совокупности, по которой производится оценка параметров, коэффициентов корреляции и регрессии.

Раздел 1. Случайные события (50 часов)

1.1. Понятие случайного события (10 часов)

[1], с. 6…13

Случайные события. Пространство элементарных событий. Алгебра событий. Аксиомы вероятностей и следствия из них.

1.2. Вероятности случайных событий (10 часов)

[1], с. 14…22

Частота и вероятность события и свойства. Классическое определение вероятности. Геометрическое определение вероятности. Вероятность суммы событий.

1.3. Основные формулы для вычисления вероятностей. (30 часов)

[1], с. 23…31

Независимость событий. Проведение независимых испытаний по схеме Бернулли. Условная вероятность. Вероятность произведения событий. Формула полной вероятности. Формула Байеса.

Раздел 2. Случайные величины (60 часов)

2.1. Описание случайных величин (20 часов)
[1], с. 32…41

Определение случайной величины. Дискретные и непрерывные случайные величины. Ряд распределения и функция распределения случайной величины. Плотность вероятности случайной величины. Свойства функции распределения и плотности вероятности.

2.2. Числовые характеристики случайных величин (40 часов)
[1], с. 59…83

Вычисление математического ожидания и дисперсии дискретной и непрерывной случайных величин. Законы распределения: биномиальный, пуассоновский, закон равномерной плотности, нормальный закон распределения (закон Гаусса).

Раздел 3. Элементы математической статистики (40 часов)

3.1. Основные определения (6 часов)

[2], с. 6…15

Генеральная совокупность и выборка. Оценка параметров. Получение экспериментальных характеристик случайной величины. Понятие статистической гипотезы.

3.2. Систематизация выборки (14часов)
[2], с.16…29

Вариационный ряд, статистическая вероятность события. Эмпирическая функция распределения, гистограмма частот. Получение опытных значений математического ожидания и дисперсии случайной величины.

3.3. Точечные оценки параметров распределения (10 часов).
[2], с. 24…32

Смещенная, несмещенная, состоятельная оценки параметров. Точечные оценки математического ожидания и дисперсии.

3.4. Интервальные оценки (10 часов)
[3], гл.16, с.126…138

Доверительная вероятность, доверительный интервал, построение интервальных оценок, корреляция и регрессия. Основные статистические законы распределения.

Заключение

В результате изучения указанных материалов студент может выполнить задания, включенные в контрольную работу, ответить на вопросы проверочных тестов и сдать экзамен по данному курсу.

2.2. Тематический план занятий

Тематический план дисциплины
для студентов очной формы обучения

№ п/п	Наименова-ние раздела (отдельной темы)	Кол-во часов по дневной форме обучения	Виды занятий и контроля											
			лекции		ПЗ (С)		ЛР		Самостоятель-ная работа	Тесты	Контрольны е работы	Лаб.работы		Курс.работ ы
			аудит.	ДОТ	аудит.	ДОТ	аудит.	ДОТ						
ВСЕГО		**150**	**28**	**20**	**4**	**10**	**20**	**8**	**60**		**1**	**2**		
1	**Случайные события**	**50**	**12**	**8**	**2**	**4**			**24**	**№ 1**				
1.1	Понятие случайного события	10	4	2		2			2					
1.2	Вероятности случайных событий	10	4	2					4					
1.3	Основные формулы для вычисле- ния вероят- ностей	30	4	4	2	2			18		**За-да-ча 1**			
2	**Случайные величины**	**60**	**12**	**8**	**2**	**4**	**10**	**4**	**20**	**№ 2**		**№ 1**		
2.1	Описание случайных величин	20	4	4		2	2		8					
2.2	Числовые характерис-тики случай- ных величин	40	8	4	2	2	8	4	12		**За-да-ча 2**			
3	**Элементы математи-ческой статистики**	**40**	**4**	**4**		**2**	**10**	**4**	**16**		**За-да-чи 3,4**	**№ 2**		
3.1	Основные определения	10	1	1		2			2					
3.2	Систематиза-ция выборки	10	1	1			2		2					
3.3	Точечные оценки пара-метров расп-ределения	10	1	**1**			4		6					
3.4	Интерваль-ные оценки	10	1	1			4	4	6					

Тематический план дисциплины
для студентов очно-заочной формы обучения

№ п/п	Наименование раздела (темы)	Кол-во часов по дневной форме обучения	Виды занятий и контроля											
			лекции		ПЗ (С)		ЛР		Самостоятель-ная работа	Тесты	Контрольны е работы	Лаб.работы		Курс.работ ы
			аудит.	ДОТ	аудит.	ДОТ	аудит.	ДОТ						
ВСЕГО		**150**	**12**	**24**	**8**	**20**	**12**	**14**	**60**		**1**	**2**		
1	**Случайные события**	**50**	**4**	**8**	**2**	**4**			**32**	**№ 1**				
1.1	Понятие случайного события	10	2	2					6					
1.2	Вероятности случайных событий	10		2		2			6					
1.3	Основные формулы для вычисления вероятностей	30	2	4	2	2			20		**Задача 1**			
2	**Случайные величины**	**60**	**4**	**8**	**2**	**8**	**4**	**8**	**26**	**№ 2**		**№ 1**		
2.1	Описание случайных величин	20	2	4		4	2	4	4					
2.2	Числовые характеристи-ки случайных величин	40	2	4	2	4	2	4	22		**Задача 2**			
3	**Элементы математичес-кой статистики**	**40**	**4**	**8**	**4**	**8**	**8**	**6**	**2**		**Задачи 3,4**	**№ 2**		
3.1	Основные определения	10	2	2	2		2		2					
3.2	Систематиза-ция выборки	10	2	2	2		2	2						
3.3	Точечные оценки пара-метров расп-ределения	10		2		4	2	2						
3.4	Интерваль-ные оценки	10		2		4	2	2						

Тематический план дисциплины
для студентов заочной формы обучения

№ п/п	Наименование раздела (отдельной темы)	Кол-во часов по дневной форме обучения	Виды занятий и контроля											
			лекции		ПЗ (С)		ЛР		Самостоятель-ная работа	Тесты	Контрольные работы	Лаб.работы		Курс.работы
			аудит.	ДОТ	аудит.	ДОТ	аудит.	ДОТ						
ВСЕГО		**150**	**6**	**36**	**4**	**24**	**10**	**10**	**60**		**1**	**2**		
1	**Случайные события**	**50**	**2**	**12**	**4**	**8**			**24**	**№ 1**				
1.1	Понятие случайного события	10		2		2			6					
1.2	Вероятности случайных событий	10		4		2			4					
1.3	Основные формулы для вычисления вероятностей	30	2	6	4	4			14		**Задача 1**			
2	**Случайные величины**	**60**	**4**	**12**		**8**	**4**	**4**	**28**	**№ 2**		**№ 1**		
2.1	Описание случайных величин	20	2	6		4			8					
2.2	Числовые характеристи-ки случайных величин	40	2	6		4	4	4	20		**Задача 2**			
3	**Элементы математичес-кой статисти ки**	**40**		**12**		**8**	**6**	**6**	**8**		**Задачи 3,4**	**№ 2**		
3.1	Основные определения	4	2	2		2		2	2					
3.2	Систематиза-ция выборки	4		2		2		2	2					
3.3	Точечные оценки пара-метров расп-ределения	10		4		2	2	2	2					
3.4	Интерваль-ные оценки	12		4		2	4		2					

2.3. Структурно-логическая схема дисциплины

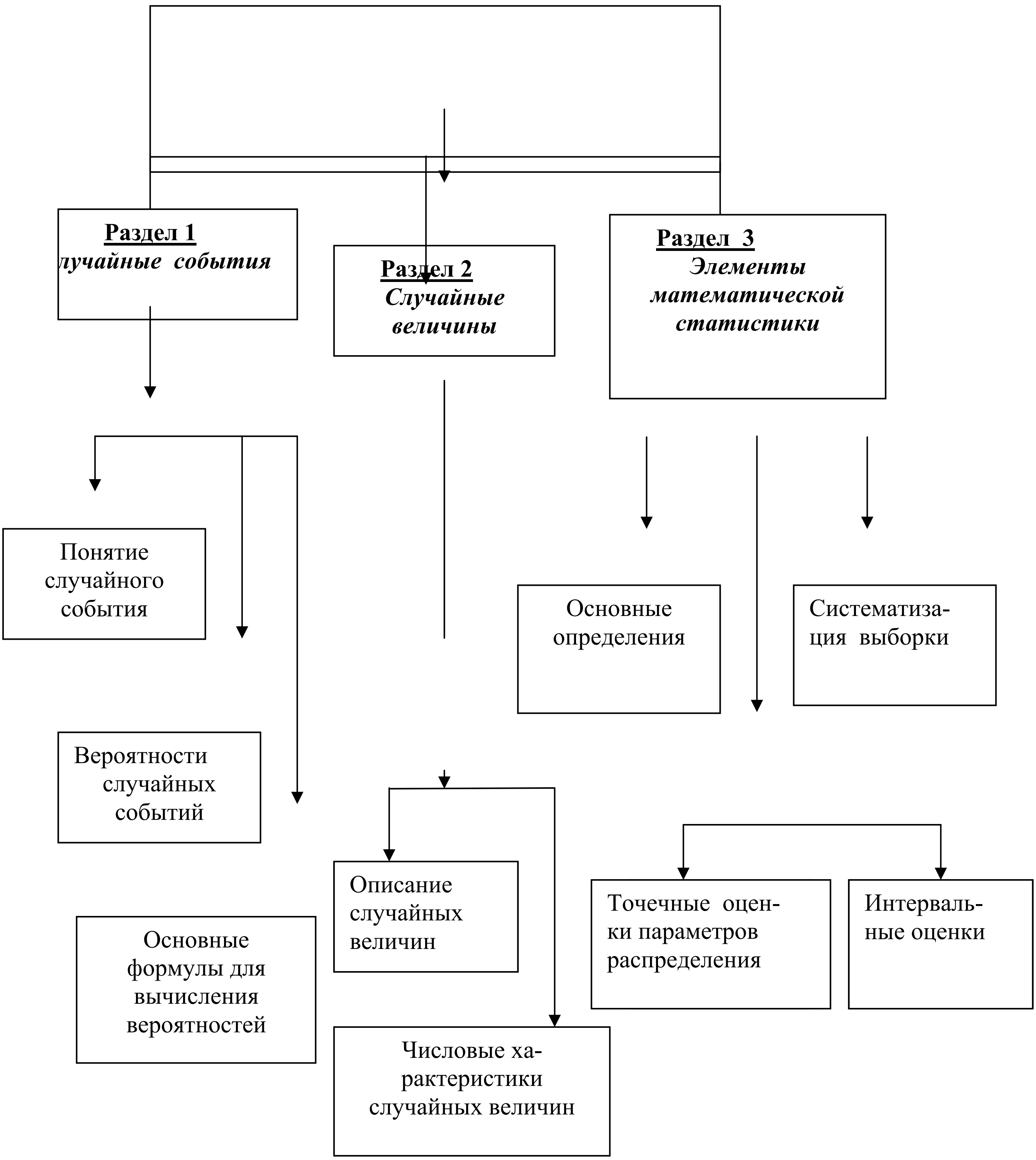


Рис. 2.1

2.4. Временной график изучения дисциплины

(для студентов заочной формы обучения с использованием ДОТ)

№	Название раздела (темы)	Продолжительность изучения раздела (темы) в днях (из расчета – 4 часа в день)
1	Раздел 1. Случайные события	15 дн.
2	Раздел 2. Случайные величины	12 дн.
3	Раздел 3. Элементы математической статистики	10,5 дн.
Итого:		37,5 дн

2.5. Практический блок

Практические занятия (очно-заочная форма обучения)

Номер и название раздела (темы)	енование практических занятий	во часов	
		Ауд.	ДОТ
Раздел 1	Случайные события	2	4
Раздел 2	Случайные величины	2	8
Раздел 3	Элементы математической статистики	4	8

Практические занятия (заочная формы обучения)

мер и название раздела (темы)	Наименование практических занятий	ол-во часов	
		Ауд.	ДОТ
Раздел 1	Случайные события	4	8
Раздел 2	Случайные величины	-	8
Раздел 3	Элементы математической статистики	-	8

Практические занятия (очная форма обучения)

Номер и название раздела (темы)	аименование практических занятий	ол-во часов	
		Ауд.	ДОТ
Раздел 1	Случайные события	2	4
Раздел 2	Случайные величины	2	4
Раздел 3	Элементы математической статистики	-	2

Лабораторные работы (очно-заочная форма обучения)

Номер и название раздела (темы)	Наименование лабораторной работы	-во часов Ауд.	ДОТ
Раздел 2	*Работа 1.* Описание случайных величин Числовые характеристики	2 2	4 4
Раздел 3	*Работа 2.* Основные определения Систематизация выборки. Точечные оценки параметров распределения Интервальные оценки	2 2 2 2	- 2 2 2

Лабораторные работы (очная форма обучения)

Номер и название раздела (темы)	Наименование лабораторной работы	-во часов Ауд.	ДОТ
Раздел 2	*Работа 1.* Описание случайных величин Числовые характеристики	2 3	4
Раздел 3	*Работа 2.* Основные определения Систематизация выборки Точечные оценки параметров распределения Интервальные оценки	2 2 4 2	2 2

Лабораторные работы (заочная форма обучения)

Номер и название раздела (темы)	Наименование лабораторной работы	-во часов уд.	ДОТ
Раздел 2	*Работа 1.* Описание случайных величин. Числовые характеристики.	4	4
Раздел 3	*Работа 2.* Основные определения Систематизация выборки Точечные оценки параметров распределения Интервальные оценки	2 4	2 2 2

2.6. Балльно-рейтинговая система

Изучение дисциплины проводится в течение одного семестра, и состоит из усвоения 3-х теоретических разделов, выполнения 2-х лабораторных работ на ПК и одной контрольной работы.

В каждом разделе приводятся вопросы тестов текущего контроля, *ответы на которые не оцениваются, но являются репетицией сдачи контрольных тестов,* для подготовки к которым предлагается пройти тренировочный тест, по репетиционным вопросам, приводимым в блоке 6.

За каждый вид самостоятельной работы начисляется определенное число баллов:

- за правильный ответ на вопрос тренировочного теста – 2 балла;
- правильно выполненную лабораторную работу – 10 баллов;
- правильно выполненную контрольную работу – 26 баллов.

При успешной работе студент может получить максимум 100 баллов. Для получения допуска к экзамену нужно набрать более двух третей от этой суммы (т.е. не менее 67 баллов).

3. Информационные ресурсы дисциплины

3.1. Библиографический список

Основной:

1. Вентцель, Е.С. Теория вероятностей / Е.С. Вентцель. - М.: Academia, 2003.
2. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. - М.: Высш. шк., 2004.
3. Ткаченко, Г.Г. Теория вероятностей: учебное пособие / Г.Г.Ткаченко. – СПб.: Изд-во СЗТУ, 2005.
4. Шабаева, М.Б. Вычислительная математика. Элементы математической статистики / М.Б. Шабаева. – СПб.: Изд-во СЗТУ, 2004.

Дополнительный:

5. Чистяков, В.П. Курс теории вероятностей / В.П. Чистяков, - СПб.: Лань, 2003.
6. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. - М.: Наука, 1988.
7. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике / В.Е. Гмурман. - М.: Высш. шк., 1979.

3.2. Опорный конспект по курсу “ Математика.Часть 2

Теория вероятностей и элементы математической статистики

ВВЕДЕНИЕ

Теория вероятностей – это математическая наука, изучающая закономерности в случайных явлениях. Случайные явления – это такие явления, которые при неоднократном воспроизведении одного и того же опыта с соблюдением некоторого комплекса условий протекают по-разному, что ведет к различным результатам опыта. Примеры случайных явлений можно наблюдать во многих областях науки и техники.

Теория вероятностей основывается на построении математических моделей случайных явлений, изучении свойств моделей и применяется в задачах, которые требуют определения и расчета вероятностей отдельных испытаний или серии опытов, а также для определения надежности или устойчивости работы устройств, приборов или систем обслуживания.

Настоящий «Опорный конспект» содержит теоретические основы дисциплины Математика. Часть 2. *Теория вероятностей и элементы математической статистики*, состоит из 3-х разделов, включает варианты решения различных задач, прилагаются вопросы для самопроверки и тестовые вопросы.

Раздел 1. СЛУЧАЙНЫЕ СОБЫТИЯ

Данный раздел курса «*Теория вероятностей и элементы математической статистики»* содержит краткое изложение теоретического материала для изучения понятия случайного события, классификации событий. Кроме того, приводится классическое и геометрическое определение вероятности, сформулированы аксиомы о вероятностях и следствия из них. Рассматриваются несовместные и независимые события и приводятся формулы, по которым можно вычислить вероятность суммы и произведения различных событий, а также формула для вычисления вероятности по схеме Бернулли, формула полной вероятности и формула Байеса.

Каждое понятие или приводимая формула обязательно поясняется примером, решение которого позволяет увидеть, в каких случаях следует использовать конкретную формулу, что в большой степени определяется формулировкой поставленной задачи.

Изучив материал раздела, студент может проверить свои знания по вопросам для самопроверки, которые даются в конце каждой темы, а также разобрать репетиционный тест № 1, приведенный в блоке контроля освоения дисциплины. После того, как эта часть работы проделана, студент может приступить к выполнению задачи № 1 из методических указаний к выполнению контрольной работы по вычислительной математике, основам теории вероятностей и элементам математической статистики [8].

1.1. Понятие случайного события

1.1.1. Сведения из теории множеств

Понятие множества относится к фундаментальным понятиям математики. Под ***множеством*** понимают некоторую совокупность объектов, называемых элементами множества. Для задания множества можно или перечислить все элементы, в него входящие, или определить свойства, которыми они обладают. Множества обозначают прописными буквами *A, B,* ..., а их элементы – строчными буквами *a, b,*… и заключают в фигурные скобки.

Пример 1.1. Обозначим A - множество положительных целых чисел, меньших 6

$$A = \{ 1,2,3,4,5\}.$$

Пример 1.2. Обозначим B – множество всех действительных чисел

$$B = \{x:\ -\infty < x < +\infty \}.$$

Пример 1.3. Обозначим C множество всех жителей некоторого города, которые старше 90 лет. Если x обозначает возраст жителя этого города, то все элементы множества C можно определить

$$C= \{x:\ x>90\}.$$

Выражение "***элемент а принадлежит множеству А***" будем символически записывать $a\in A$, а запись $a\notin A$ будет означать " ***элемент а не принадлежит множеству А".***

Множества, состоящие из конечного числа элементов, называют ***конечными***, в противном случае – ***бесконечными***. В примерах *1.1* и *1.3* определены конечные множества, а примером бесконечного множества является множество из примера *1.2*.

Символом $\varnothing$ будем обозначать множество, не содержащее элементов. Это множество называют ***пустым*** множеством. Например, для некоторого города множество C в примере *1.*3 может оказаться пустым.

Множество B называют ***подмножеством*** множества *A,* если все элементы B принадлежат множеству *A,* и символически записывают $B \subset A$ или $A \supset B$.

1.1.2. Пространство элементарных событий

Теория вероятностей изучает математические модели случайных явлений. Предположим, что производится некоторый эксперимент, исход (результат) которого непредсказуем. Множество тех исходов данного эксперимента, которые не могут происходить одновременно и появление одного и только одного из них обязательно произойдет, называют ***пространством элементарных событий***, а сами исходы называют ***элементарными событиями.*** Пространство элементарных событий обозначают Ω, а элементарное событие - ω. Пространство элементарных событий называют конечным, если множество элементарных событий конечно и - бесконечным в противном случае.

Рассмотрим некоторые примеры пространств элементарных событий.

Пример 1.4. Игральный кубик, имеющий шесть граней с изображением на каждой числа точек (1,2,3,4,5,6), подбрасывают один раз. Результатами этого эксперимента будем считать число очков, выпавшее на верхней грани кубика. Следовательно, пространство элементарных событий состоит из множества $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$, где элементарное событие ω_i обозначает число очков i, выпавшее на верхней грани кубика.

Пример 1.5. Эксперимент состоит в наблюдении числа автомобилей, обслуживаемых автозаправочной станцией с 12 до 15 часов. В этом случае элементарные события можно выразить числами 0,1,2…. Очевидно, что число обслуживаемых автомобилей в течение рассматриваемого промежутка времени конечно, но точно предсказать их число невозможно. Поэтому будем считать, что пространство элементарных событий состоит из бесконечного множества $\Omega = \{0,1,2,\ldots\}$.

Пример 1.6.Игральный кубик подбрасывают один раз. Рассмотрим следующие события: A = {выпало четное число},

B = {выпало нечетное число},

C = {выпало число ≤ 3}.

Каждое из этих событий отождествим с множеством всех исходов, при которых они наступают. Тогда события

$$A=\{\omega_2, \omega_4, \omega_6\},\ B=\{\omega_1, \omega_3, \omega_5\},\ C=\{\omega_1, \omega_2, \omega_3\}.$$

Отсюда видно, что все эти события являются подмножествами пространства элементарных событий.

1.1.3. Классификация событий

Для конечных пространств элементарных событий отождествим событие и множество всех исходов, при которых данное событие наступает. Эти исходы называют элементарными событиями, **благоприятствующими** данному событию. Для конечных пространств элементарных событий ***событие*** – это множество всех исходов ему благоприятствующих. Такой подход к определению случайного события позволяет применять теорию множеств.

Определение. Невозможным событием называется событие, которое не может наступить в условиях данного эксперимента, т.е. это событие имеет пустое множество благоприятствующих исходов.

Например, пусть событие D = {на верхней грани кубика выпало число > 7}. Это событие является невозможным и ему соответствует пустое множество $\varnothing$

благоприятствующих исходов. Будем невозможное событие обозначать символом $\varnothing$.

Определение. Достоверным называется событие Ω, которое всегда наступает в условиях данного эксперимента. Множество благоприятствующих исходов достоверного события совпадает с пространством элементарных событий Ω.

Пусть событие E = {на верхней грани кубика выпало число <= 7}. Это событие является достоверным и множество благоприятствующих ему исходов совпадает с пространством элементарных событий.

Определение. Если при каждом осуществлении события A происходит событие B**,** то говорят, что ***событие A влечет событие B***. В этом случае множество благоприятствующих исходов события A содержится в множестве благоприятствующих исходов события B, т.е. $A \subset B$.

Определение. События A и B ***называются эквивалентными***, если событие A влечет событие B, а событие B влечет A.

Определение. **Событие $\overline{A}$ =Ω-*A* называется *противоположным*** событию A. Множество благоприятствующих исходов события $\overline{A}$ является дополнением до пространства элементарных событий Ω множества благоприятствующих исходов события A (или появление события $\overline{A}$ **-** это непоявление события A).

Определение. События A и B называются несовместными**,** если они не могут произойти вместе.

1.1.4. Сумма и произведение событий

Определение. Суммой* (*объединением*) *событий A* и *B называется событие, которое наступает, когда происходит хотя бы одно из этих событий, и обозначается $A+B$. При сложении событий множества благоприятствующих исходов складываются (объединяются).

Например, для событий примера 1.6 суммой событий A и C будет событие $A+C=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_6\}$, а суммой событий A и B будет событие $A+B=\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}=\Omega$, т. е. достоверное событие.

Операцию сложения определяют и для бесконечной последовательности событий.

Определение. Суммой* (*объединением*) *последовательности событий $A_1, A_2, \dots A_n, ..$ называется событие, которое наступает, когда происходит хотя бы одно из событий последовательности и обозначается $\bigcup A_i$.

Пусть событие A состоит из благоприятствующих исходов $\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_m}$.

Тогда событие A по определению суммы можно представить в виде

$$A = \omega_{i_1} + \omega_{i_2} + \dots + \omega_{i_m}.$$

Определение*. *Произведением событий A и B называется событие, которое происходит при одновременном наступлении этих событий и обозначается AB. При умножении событий множества благоприятствующих исходов умножаются (пересекаются).

Например, для событий примера 1.6 произведением событий A и C будет событие $AC = \{\omega_1, \omega_3\}$, а произведением событий A и B будет невозможное событие $AB = \varnothing$.

Определение. Произведением последовательности событий $A_1, A_2, \ldots A_n, \ldots$ называется событие, которое происходит при одновременном наступлении всех событий последовательности и обозначается $\bigcap\limits_{i=1}^{\infty} A_i$.

Определение. Разность событий A и B происходит, когда событие A наступает, а событие B - не наступает, и обозначается $A-B$.

Используя определения действий над событиями, можно доказать следующие свойства

1) $A+B=B+A$ *2)* $AB=BA$ *3)* $A+(B+C)=(A+B)+C$

4) $A(B+C)=AB+AC$ *5)* $A+\varnothing=A$ *6)* $A\varnothing=\varnothing$

7) $A\Omega=A$ *8)* $A+A=A$ *9)* $AA=A$

10) $A+\Omega=\Omega$ *11)* $A\Omega=A$ *12)* $A+\overline{A}=\Omega$

13) $A\overline{A}=\varnothing$ *14)* $\overline{\overline{A}}=A$ *15)* $\overline{\Omega}=\varnothing$ *16)* $\overline{\varnothing}=\Omega$.

Первые семь свойств аналогичны свойствам алгебры, таким как перестановка, сочетание и распределение, при этом невозможное событие $\varnothing$ можно считать как 0, а достоверное событие Ω – как 1. Остальные свойства не имеют аналогов в алгебре.

Для событий A и B справедливы формулы, называемые соотношениями двойственности:

$$\overline{A \cup B} = \overline{A} \cap \overline{B} \qquad \overline{A \cap B} = \overline{A} \cup \overline{B} .$$

Определение. Класс событий U образует ***алгебру событий***, если

1) достоверное событие содержится в этом классе, т.е. $\Omega \in U$;

2) для любых событий $A \in U, B \in U$ из этого класса их сумма и произведение также принадлежат этому классу: $AB \in U$, $A+B \in U$;

3) если событие A из этого класса $A \in U$, то и противоположное событие также принадлежит этому классу: $\overline{A} \in U$.

Пример 1.7. Подбрасывают две монеты различного достоинства. Пространство элементарных событий Ω состоит из четырех элементов

$$\Omega = \{ГГ, ГЦ, ЦГ, ЦЦ\}.$$

Здесь *Г* означает, что монета выпала гербом вверх, а *Ц* – цифрой вверх.

Построим все подмножества пространства элементарных событий Ω:

$\varnothing$, *ГГ, ГЦ, ЦГ, ЦЦ,* { *ГГ, ГЦ* }, { *ГГ, ЦГ*}, {*ГГ, ЦЦ*}, { *ГЦ, ЦГ* }

{ *ГЦ, ЦЦ* }, { *ЦГ, ЦЦ* }, {*ГГ, ГЦ, ЦГ*}, {*ГГ, ГЦ, ЦЦ* }, {*ГГ, ЦГ, ЦЦ* },

{*ГЦ, ЦГ, ЦЦ* }, {*ГГ, ГЦ, ЦГ, ЦЦ* } $=\Omega$.

Нетрудно проверить, что все 16 событий образуют алгебру событий. Для точного определения события в произвольном пространстве элементарных событий рассмотрим следующее определение.

Определение. Алгебра событий U образует **σ-алгебру событий**, если для бесконечной последовательности событий A_i из σ-алгебры событий их объединение и пересечение принадлежат σ-алгебре

$$\bigcup_{i=1}^{\infty} A_i \in U, \qquad \bigcap_{i=1}^{\infty} A_i \in U.$$

Если задано пространство элементарных событий Ω и σ-алгебра событий U, то говорят, что задано ***измеримое пространство*** **{ Ω, *U* }.**

В случае произвольного пространства элементарных событий Ω, ***событиями называют*** только такие подмножества пространства элементарных событий Ω, которые образуют σ-алгебру событий U. Все остальные подмножества Ω, не входящие в σ-алгебру событий U, событиями не являются.

Вопросы для самопроверки

1. При подбрасывании монеты выпала сторона с изображением герба (условно обозначим это событие буквой A). Какое событие будет являться противоположным событию A?
2. Подбрасываются две монеты, в результате чего видим изображение двух гербов. Что будет являться противоположным событием в этом случае?
3. Написать действие, соответствующее тому факту, что при подбрасывании двух монет на одной будет изображен герб (событие A), а на другой монете – цифра (событие B).

1.2. Вероятности случайных событий

1.2.1. Относительная частота события, аксиомы теории вероятностей. Классическое определение вероятности

Заметим, что всякое событие есть некоторое высказывание о результатах рассматриваемого эксперимента. При многократном проведении опыта с соблюдением некоторого комплекса условий S возможны различные исходы:

1) в каждом из опытов можно наблюдать один и тот же результат;

2) ни в одном из опытов интересующий результат не появился;

3) в ряде опытов интересующий результат можно было наблюдать, а в оставшихся опытах этого не происходило.

В первом случае говорят, что происходит ***достоверное*** событие.

Во втором случае речь идет о ***невозможном*** событии.

В третьем случае говорят, что происходит ***случайное*** событие.

Однако случайность события по отношению к комплексу условий S не означает отсутствия всякой закономерной связи между ними.

Будем обозначать случайные события большими латинскими буквами: A, B, C, D, F и т. д.

Допустим, что производится серия из **n** одинаковых опытов. В каждом из опытов интересующее событие, назовем его A, может произойти или не произойти. В результате наблюдений за опытами замечено, что событие A появилось **m** раз, обозначим как $\boldsymbol{m(A)}$.

<u>Определение.</u> **Частотой** события A (относительной частотой) называется величина, равная отношению числа опытов, в которых событие A произошло, ко всем проводимым опытам **n** , т.е. $\mu_n(A) = \dfrac{m_n(A)}{n}$.

Экспериментально установлено, что для многих опытов, в которых рассматривается появление случайного события A, имеет место ***закон устойчивости относительных частот***: если при неограниченном увеличении числа опытов n относительная частота события A $\mu_n(A) = \dfrac{m_n(A)}{n}$ колеблется около некоторого числа $\boldsymbol{p(A)}$, то число $\boldsymbol{p(A)}$ называют **вероятностью** события A, а само это свойство называют законом устойчивости относительных частот.

Из определения видно, что частота имеет свойства:

1) $\mu_n(\Omega)=1$;

2) $\mu_n(\varnothing)=0$;

3) $0 =< \mu_n(A) <= 1$;

4) $\mu_n(A+B)= \mu_n(A)+\mu_n(B)$, если события несовместны, т.е. такие, которые не могут происходить вместе.

Из закона устойчивости относительных частот следует, что **вероятность** события является в некотором смысле пределом относительной частоты этого события. Поэтому свойства относительной частоты событий можно перенести в качестве **аксиом** на вероятность этих событий.

1) $p(\Omega) = 1$ - *вероятность достоверного* события равна 1;

2) $p(\varnothing) = 0$ - *вероятность невозможного* события равна **0;**

3) $0 =< p(\mathrm{A}) = < 1$ - это означает, что ***вероятность*** любого события не может быть **меньше 0** и **больше 1.**

4**)** ***аксиома конечной аддитивности:*** если события *A* и *B несовместны,* то

$$p(A+B) = p(A) + p(B) .$$

На основании этих аксиом можно получить формулу для ***классического определения вероятности***:

$$\boldsymbol{P(A) = m / n,} \qquad (1.1)$$

т.е. ***вероятность любого события*** есть величина, равная ***отношению числа m*** опытов (исходов), в которых событие *A* произошло, *к* ***общему числу*** проводимых опытов ***n***, причем возможность появления каждого из элементов одинакова.

Пример 1.8. В ящике находятся 15 хороших деталей и 10 бракованных. Найти вероятность того, что при изъятии одной детали из ящика, она окажется бракованной (обозначим как событие *A*).

Решение. Общее число деталей в ящике $n = 15+10=25$. Число исходов, соответствующих тому, что выбрана бракованная деталь, $m = 10$. Следовательно, вероятность того, что будет вынута бракованная деталь, равна

$$P(A) = m / n = 10/25= 2/3.$$

1.2.2. Геометрическое определение вероятности

Если пространство элементарных событий содержит бесконечное число элементарных событий, то классическое определение вероятности неприменимо. В тех случаях, когда пространство элементарных событий может быть представлено некоторой областью на прямой, плоскости или в пространстве, то, учитывая равную возможность исходов эксперимента, можно построить геометрическое определение вероятности события. Допустим, что пространство элементарных событий можно геометрически представить на плоскости некоторой областью Ω, а любое событие A – подмножеством этой области Ω. Обозначим $S(\Omega)$ *меру* области Ω, $S(A)$ – *мера* области A. Тогда вероятность события A можно определить как отношение соответствующей меры $S(A)$ к мере всей области Ω:

$$p(A) = \frac{S(A)}{S(\Omega)}. \tag{1.2}$$

В этом случае выполняются все аксиомы теории вероятности. Следует заметить, что событиями в этом примере считаются множества, для которых может быть определена их площадь.

Пример 1.9. Производится один выстрел по круглой мишени радиуса R. Предполагается, что каждая точка мишени может быть поражена с одинаковой вероятностью. Найти вероятность того, что расстояние от точки попадания до центра мишени меньше r ($r<R$).

Решение. Обозначим событие A ={точка попадания лежит в заданном круге радиуса r}.Тогда вероятность этого события по формуле (1.2) будет равна

$$p(A) = \frac{\pi r^2}{\pi R^2} = \frac{r^2}{R^2} .$$

1.2.3. Вычисление вероятности события через элементы комбинаторного анализа

Использование методов комбинаторного анализа широко описано в учебном пособии по теории вероятностей. В данной работе рассмотрим ***гипергеометрический способ*** вычисления вероятности события.

Пример 1.10. В урне имеется N шаров, из них M белых и N-M черных. Наудачу из урны извлекают n шаров. Найти вероятность того, что среди них будет ровно m белых шаров.

Решение. Обозначим искомое событие A: A= {выбрано ровно m белых шаров}. Из генеральной совокупности N шаров выбирают без учета порядка следования n шаров. Следовательно, число различных выборок будет равно числу сочетаний из N по n -- C_N^n . Теперь найдем число выборок объема n, в которых ровно m белых шаров. Для этого будем выбирать из всех M белых шаров ровно m шаров. Всего число таких различных выборок объема m будет равно числу сочетаний из M по m - C_M^m . Аналогично, число различных выборок из всех черных шаров по n-m шаров, будет равно числу сочетаний из ***N-M*** по ($\boldsymbol{n - m}$), т.е. C_{N-M}^{n-m} . Объединяя каждую выборку из $\boldsymbol{m}$ белых шаров с каждой выборкой из $(\boldsymbol{n - m})$ черных шаров, получаем искомое число выборок объема $\boldsymbol{n}$, в которых будет ровно m белых, - $C_M^m C_{N-M}^{n-m}$. Из классического определения вероятности следует, что

$$P(A)=\frac{C_M^m C_{N-M}^{n-m}}{C_N^n} ,$$

где C_r^k - число сочетаний из $\boldsymbol{r}$ по $\boldsymbol{k}$, вычисляемое по формуле

$$C_r^k = \frac{A_r^k}{k!} = \frac{r!}{k!(r-k)!} , \qquad k! = 1\cdot 2\cdot 3\cdot \ldots k \quad \text{и} \quad 0! = 1.$$

Пример 1.11. Среди десяти изделий находится три бракованных. Выбирают наугад четыре изделия. Определить вероятности следующих событий:

A = {среди выбранных изделий нет бракованных };

B = {из выбранных изделий ровно два бракованных}.

Решение. Определим общее число способов выбора четырех деталей из десяти.

$$C_{10}^4 = \frac{10!}{4!\cdot 6!} = \frac{10\cdot 9\cdot 8\cdot 7}{4\cdot 3\cdot 2} = 10\cdot 3\cdot 7 = 210 .$$

Число выборок, благоприятствующих событию A, будет равно

$$C_7^4 = \frac{7!}{4!\cdot 3!} = \frac{7\cdot 6\cdot 5}{3\cdot 2} = 7\cdot 5 = 35 .$$

Тогда вероятность события A будет равна

$$p(A)= \frac{C_7^4}{C_{10}^4} = \frac{35}{210} = \frac{1}{6}.$$

Для вычисления вероятности события *B* применим гипергеометрическое определение. Будем считать изделия шарами, бракованные изделия - белыми шарами, а небракованные изделия - черными шарами. Тогда найти вероятность события *B* означает найти вероятность того, что среди 4 наугад выбранных шаров будет два белых:

$$p(B)= \frac{C_3^2 C_7^2}{C_{10}^4} = \frac{3 \cdot 21}{210} = \frac{3}{10}.$$

1.2.4. Свойства вероятностей событий

Свойство 1. Если для события *A* известна его вероятность *p(A),* то вероятность противоположного события $\overline{A}$ можно найти по формуле

$$p(\overline{A}) = 1 - p(A). \qquad (1.3)$$

Свойство 2 . Вероятность невозможного события равна 0, т.е.

$$p(\varnothing) = 0.$$

Свойство 3. Если справедливо соотношение $A \subset B$, т.е. событие *A* влечет событие *B*, то $p(A) \le p(B)$.

Свойство 4 ***(*** **теорема сложения).**

Для любых событий из алгебры событий $A, B \in U$ справедливо равенство

$$\mathbf{p(A+B) = p(A) + p(B) - p(AB)}, \qquad (*)$$

вероятность суммы *двух совместных событий равна* ***сумме вероятностей*** *этих событий* ***минус вероятность произведения*** *данных событий.*

Свойство 5. Для *несовместных событий A и B* **вероятность суммы** этих событий равна **сумме вероятностей** события *A* и события *B*:

$$\mathbf{P(A + B) = P(A) + P(B).} \qquad (**)$$

1.2.5. Независимые события

Рассмотрим определение независимости событий, которое отражает понятие реальной независимости несвязанных событий.

Определение. События называются ***независимыми,*** если появление одного из них не влияет на вероятность появления другого.

Пример 1.12. Предположим, что подбрасывают два игральных кубика независимо друг от друга. Пространство элементарных событий состоит из упорядоченных пар чисел (n = 36):

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)

(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6)

Рассмотрим два события.

Событие A = { число очков на первом кубике > 4 } состоит из всех пар пятой и шестой строк (m = 12) и имеет вероятность $p(A) = \frac{m}{n} = \frac{12}{36} = = \frac{1}{3}$.

Событие B = { число очков на втором кубике < 4 } состоит из всех пар первых трех столбцов (n = 18) и имеет вероятность $p(B)$ = $\frac{18}{36} = \frac{1}{2}$. Очевидно, что эти события причинно не связаны друг с другом и *независимы* в этом смысле.

Найдем ***вероятность произведения*** этих событий.

Событие AB состоит из шести пар выпадающих цифр (m = 6)

{ (5,1) (5,2) (5,3) (6,1) (6,2) (6,3) }

и имеет вероятность p (AB)= $\frac{6}{36} = \frac{1}{6}$. Очевидно, что выполняется равенство ***p(AB) = p(A) p(B).*** Оно отражает независимость событий A и B.

Определение. События A и B называются ***независимыми,*** если выполняется равенство

$$\boldsymbol{p(AB) = p(A)\, p(B)}, \qquad (1.4)$$

т.е. ***вероятность произведения*** двух ***независимых*** событий равна ***произведению*** вероятностей этих событий. В противном случае события считают зависимыми.

Пример 1.13. (продолжение примера 1.12).

Рассмотрим событие C = {сумма очков равна 8}. Оно состоит из 5 пар

$$\{(2,6)\ (3,5)\ (4,4)\ (5,3)\ (6,2)\}$$

и имеет вероятность $p(C)=\frac{5}{36}$.

Событие, которое получается как *произведение* события A на событие C, состоит из пар {(5,3) (6,2)} и имеет вероятность

$$p(AC)=\frac{2}{36}=\frac{1}{18}.$$

Так как $p(AC) \neq p(A)\ p(C) = \frac{1}{3}\cdot\frac{1}{18} = \frac{1}{54}$–, то события A и C следует считать *зависимыми.*

Свойство 6 . Вероятность *суммы двух* ***независимых*** *событий* A и B равна **сумме** вероятностей этих событий **минус** произведение вероятности события A на вероятность события B, т.е.

$$\boldsymbol{p(A+B) = p(A) + p(B) - p(A)\,p(\mathrm{B}).} \qquad (***)$$

Примечания:

а) Формулы (*), (**), (***) позволяют вычислить вероятность суммы двух любых событий.

б) Если требуется вычислить вероятность **суммы трех и более** событий, то вычисление надо производить, введя замену переменных таким образом, чтобы поэтапно свести расчет к вычислению вероятности суммы двух событий.

Например, найти $\boldsymbol{p(A+B+C+D) = p(R+K) = p(R) + p(K) - p(RK)}$,

где $R=A+B$, $K=C+D$. Тогда по формуле 1.3 находим

p(R) =p(A +B) = p(A) +P(B) – P(AB), $p(K) = p(C+D) = p(C) + p(D) - p(CD)$, значения которых надо подставить в исходную формулу.

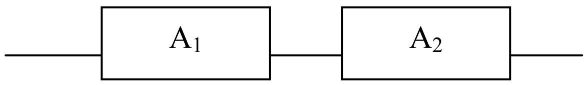
1.2.6. Расчет вероятности безотказной работы прибора

Рассмотрим примеры, в которых требуется вычислить вероятность безотказной работы и вероятность отказа работы прибора, в состав которого входят несколько элементов и используются различные способы их соединения между собой.

Пример 1.14. Прибор состоит из двух независимо работающих элементов. Вероятность отказа первого элемента равна $p_1 = 0{,}1$, а второго – $p_2 = 0{,}2$.

Рассмотрим событие A = {прибор откажет работать}.

***1*)** Вычислим вероятность события A, если элементы соединены **<u>последовательно</u>**,



Решение: Обозначим через A_1 событие, которое заключается в том, что откажет элемент A_1 = {откажет первый элемент}, и через A_2 -

A_2 = {откажет второй элемент}.

Тогда данный прибор не будет работать (событие А), если выйдет из строя **хотя бы** один из элементов (или первый, или второй, или оба не будут работать). Такое состояние прибора можно описать, используя **определение суммы событий**, т.е. $\boldsymbol{A=A_1+A_2}$. Из теоремы о вероятности суммы двух независимых событий [формула (***)] получаем

$$p(A) = p(A_1+A_2) = p(A_1) + p(A_2) - p(A_1 A_2) = p(A_1) + p(A_2) - p(A_1)\ p(A_2) =$$
$$= p_1+p_2 - p_1 p_2 = 0{,}1 + 0{,}2 - 0{,}1 * 0{,}2 = 0{,}28.$$

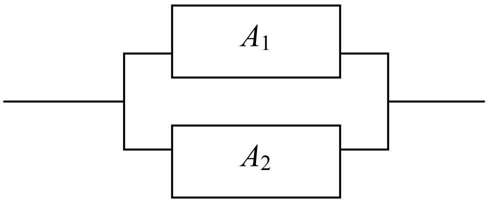
Итак, вероятность того, что данный прибор **откажет** $p(A) = 0{,}28$.

Состояние прибора, когда он работает правильно, есть событие $\overline{A}$ - противоположное событию A, когда прибор откажет.

Тогда, используя свойства вероятности, можно найти **вероятность правильной работы** $\overline{A}$ данного прибора по формуле:

$$\mathrm{p}(\overline{A}) = 1 - \mathrm{p}(A) = 1 - 0{,}28 = 0{,}72.$$

***2*)** Вычислим вероятность отказа прибора (событие A), если элементы соединены параллельно:



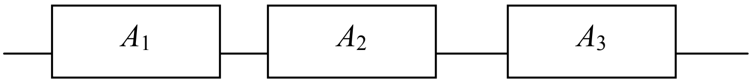
Решение. Данный прибор откажет в том случае, если **откажут оба элемента одновременно**. Следовательно, отказ прибора в этом случае может быть представлен как **произведение** событий A_1 и A_2 , т.е. $\boldsymbol{A=A_1A_2}$. Так как элементы перестают работать **независимо** друг от друга, то из независимости событий A_1 и A_2 получаем $p(A) = p(A_1)\, p(A_2) = p_1\, p_2 = 0{,}1 * 0{,}2 = 0{,}02$.

Определение. События $A_1\ A_2\ \dots\ A_n$ называют ***взаимно независимыми***, если для любой их части $A_{i_1} A_{i_2} \cdots A_{i_m}$ выполняется равенство

$$p(A_{i_1} A_{i_2} \cdots A_{i_m}) = p(A_{i_1})\, p(A_{i_2}) \dots p(A_{i_m}), \qquad (1.5)$$

$$1 <= i_1 < i_2 \dots < i_m <= n\ ,\ m = 2,\ \dots, n.$$

Пример 1.15. Прибор состоит из трех последовательно соединенных и независимо работающих друг от друга элементов. Каждый из элементов может быть признан бракованным или стандартным:



Обозначим вероятность того, что первый элемент оказался бракованным, равной p_1, второй элемент бракованный - p_2, третий элемент бракованный - p_3. Прибор будем считать *бракованным*, если хотя бы один из его элементов бракованный. Найти вероятность того, что прибор *стандартный.*

***Решение**:* Обозначим события

A_1 = {первый элемент - стандартный},

A_2 = {второй элемент - стандартный },

A_3 = {третий элемент - стандартный }

$\boldsymbol{A}$ = {прибор стандартный }.

В данном случае прибор нормально работает в том случае, если все три элемента одновременно работают, т.е. все три элемента, входящие в прибор, стандартные. Тогда работу прибора можно описать как событие A, состоящее из **произведения** трех **независимых** событий $\boldsymbol{A=A_1*A_2*A_3}$, вероятность которого можно вычислить по формуле вероятности произведения независимых событий

$$p(A) = p(A_1)p(A_2)p(A_3) = (1 - p_1)\ (1 - p_2)\ (1 - p_3).$$

Вероятность отказа прибора (событие $\overline{A}$) в данном случае есть величина, равная вероятности события, противоположного событию A.

$$p(\overline{A}) = 1 - p(A).$$

Примечание. Рассмотренные примеры 1.13, 1.14 и 1.15 являются аналогом решения контрольной задачи №3 (первого пункта задания) из методических указаний для выполнения контрольных работ.

Рассмотрим некоторые свойства независимых событий.

Свойство7. Если A и B независимы, то $\overline{A}$ и B независимы.

Свойство 8. Если событие A не зависит от событий B_1 и B_2, а события B_1 и B_2 несовместны, тогда события A и B_1+B_2 независимы.

Свойство 9. Если события A, A_1 и A_2 взаимно независимы, тогда события A и A_1+A_2 независимы.

Вопросы для самопроверки

1. В чем заключается геометрический подход к вычислению вероятности?
2. Чему равна вероятность суммы двух противоположных событий?
3. Перечислите основные свойства вероятности события.
4. Что такое независимые события?

1.3. Формулы для вычисления вероятности событий

1.3.1. Последовательность независимых испытаний

(схема Бернулли)

Предположим, что некоторый эксперимент можно проводить неоднократно при одних и тех же условиях. Пусть этот опыт производится n раз, т. е. проводится последовательность из n испытаний.

Определение. Последовательность n испытаний называют ***взаимно независимой***, если любое событие, связанное с данным испытанием, не зависит от любых событий, относящихся к остальным испытаниям.

Допустим, что некоторое событие A может произойти с вероятностью p в результате одного испытания или не произойти с вероятностью $q=1-p$.

Определение. Последовательность из n испытаний образует схему Бернулли, если выполняются следующие условия:

1) последовательность n испытаний взаимно независима;

2) вероятность события A не изменяется от испытания к испытанию и не зависит от результата в других испытаниях.

Событие A называют “ успехом” испытания, а противоположное событие $\overline{A}$ - “неудачей”. Рассмотрим событие

B_n^m ={ в n испытаниях произошло ровно m “успехов”}.

Для вычисления вероятности этого события справедлива формула Бернулли

$$p(B_n^m) = C_n^m p^m q^{n-m}, \quad m = 1, 2, \dots, n, \qquad (1.6)$$

где C_n^m - число сочетаний из n элементов по m :

$$C_n^m = \frac{n!}{m!\cdot(n-m)!} = \frac{1\cdot 2\cdot \ldots . n}{1\cdot 2\cdot \ldots \cdot m\cdot 1\cdot 2\cdot \ldots \cdot (n-m)}.$$

Пример 1.16. Три раза подбрасывают кубик. Найти:

а) вероятность того, что 6 очков выпадет два раза;

б) вероятность того, что число шестерок не появится более двух раз.

Решение. “Успехом” испытания будем считать выпадение на кубике грани с изображением 6 очков.

а) Общее число испытаний – $n=3$, число "успехов" – $m=2$. Вероятность "успеха" - $p=\frac{1}{6}$, а вероятность "неудачи" - $q=1-\frac{1}{6}=\frac{5}{6}$. Тогда по формуле Бернулли вероятность того, что результате трехразового бросания кубика два раза выпадет сторона с шестью очками, будет равна

$$p\left(\left(B_3^2\right)=C_3^2\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^1\right)=3\cdot\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)=\frac{5}{72}.$$

б) Обозначим через A событие, которое заключается в том, что грань с числом очков 6 появится не более двух раз. Тогда событие можно представить в виде *суммы трех несовместных* событий $A=B_3^0+B_3^1+B_3^2$,

где $B_3{}^0$ – событие, когда интересующая грань ни разу не появится,

$B_3{}^1$ - событие, когда интересующая грань появится один раз,

$B_3{}^2$ - событие, когда интересующая грань появится два раза.

По формуле Бернулли (1.6) найдем

$$p(A)=p\left(B_3^0+B_3^1+B_3^2\right)=p(B_3^0)+p(B_3^1)+p(B_3^2))=C_3^0p^0q^3+C_3^1pq^2+C_3^2p^2q=$$

$$=\left(\frac{5}{6}\right)^3+3\,\frac{1}{6}\left(\frac{5}{6}\right)^2+3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)=\frac{215}{216}.$$

1.3.2. Условная вероятность события

Условная вероятность отражает влияние одного события на вероятность другого. Изменение условий, в которых проводится эксперимент, также влияет на вероятность появления интересующего события.

Определение. Пусть A и B – некоторые события, и вероятность $p(B)>0$.

Условной вероятностью события A при условии, что "событие B *уже* произошло" называется отношение вероятности произведения данных событий к вероятности события, которое произошло раньше, чем событие, вероятность которого требуется найти. Условная вероятность обозначается как $p(A/B)$. Тогда по определению

$$p(A/B)=\frac{p(AB)}{p(B)}. \qquad (1.7)$$

Пример 1.17. Подбрасывают два кубика. Пространство элементарных событий состоит из упорядоченных пар чисел

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,1) (2,2) (2,3) (2,4) (2,5) (2,6)

(3,1) (3,2) (3,3) (3,4) (3,5) (3,6)

(4,1) (4,2) (4,3) (4,4) (4,5) (4,6)

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6).

В примере 1.16 было установлено, что событие A ={число очков на первом кубике > 4} и событие C ={сумма очков равна 8} зависимы. Составим отношение

$$\frac{p(AC)}{p(A)} = \frac{2/36}{12/36} = \frac{2}{12} = \frac{1}{6}.$$

Это отношение можно интерпретировать следующим образом. Допустим, что о результате первого бросания известно, что число очков на первом кубике > 4. Отсюда следует, что бросание второго кубика может привести к одному из 12 исходов, составляющих событие A:

(5,1) (5,2) (5,3) (5,4) (5,5) (5,6)

(6,1) (6,2) (6,3) (6,4) (6,5) (6,6) .

При этом событию C могут соответствовать только два из них (5,3) (6,2). В этом случае вероятность события C будет равна $\frac{2}{12} = \frac{1}{6}$. Таким образом, информация о наступлении события A оказала влияние на вероятность события C.

1.3.3. Вероятность произведения событий

Теорема умножения

Вероятность произведения событий $A_1 A_2 \ldots A_n$ определяется формулой

$$p(A_1 A_2 \ldots A_n) = p(A_1)\, p(A_2 / A_1))\ldots p(A_n / A_1 A_2 \ldots A_{n-1}). \qquad (1.8)$$

Для произведения двух событий отсюда следует, что

$$p(AB) = p(A/B)\,p\{B\} = p(B/A)\,p\{A\}. \qquad (1.9)$$

Пример 1.18. В партии из 25 изделий 5 изделий бракованных. Последовательно наугад выбирают 3 изделия. Определить вероятность того, что все выбранные изделия бракованные.

Решение. Обозначим события:

$\boldsymbol{A_1}$ = {первое изделие бракованное},

$\boldsymbol{A_2}$ = {второе изделие бракованное},

$\boldsymbol{A_3}$ = {третье изделие бракованное},

$\boldsymbol{A}$ = {все изделия бракованные}.

Событие $\boldsymbol{A}$ есть произведение трех событий $\boldsymbol{A = A_1 A_2 A_3}$.

Из теоремы умножения (1.6) получим

$$p(A) = p(A_1 A_2 A_3) = p(A_1)\,p(A_2/A_1))p(A_3/A_1A_2).$$

Классическое определение вероятности позволяет найти $p(A_1)$ – это отношение числа бракованных изделий к общему количеству изделий:

$$p(A_1) = \frac{5}{25};$$

$p(A_2)$ – это отношение числа бракованных изделий, оставшихся после изъятия одного, к общему числу оставшихся изделий:

$$p(A_2/A_1)) = \frac{4}{24};$$

$p(A_3)$ – это отношение числа бракованных изделий, оставшихся после изъятия двух бракованных, к общему числу оставшихся изделий:

$$p(A_3/A_1A_2) = \frac{3}{23}.$$

Тогда вероятность события A будет равна

$$p(A) = \frac{5}{25}\,\frac{4}{24}\,\frac{3}{23} = \frac{1}{5\cdot 2\cdot 23} = \frac{1}{230}.$$

1.3.4. Формула полной вероятности и формула Байеса

Теорема. Пусть A - некоторое событие, и события $H_1, H_2, \ldots, H_n$, , попарно несовместные, т.е. $H_i H_j = \varnothing$, $i \neq j$, образующие полную группу, и появление одного из них и только одного есть событие достоверное.

Допустим, что событие A может произойти вместе с одним из событий H_i Тогда имеет место формула полной вероятности

$$p(A) = \sum_{i=1}^{n} p(H_i)\, p(A \mid H_i) \tag{1.10}$$

и формула Байеса

$$p(H_i \mid A) = \frac{p(H_i)p(A/H_i)}{\sum_{j=1}^{n} p(H_j)p(A/H_j)} . \tag{1.11}$$

Замечание. События H_i называют ***гипотезами***, вероятности $p(H_i)$ - ***априорными вероятностями гипотез*** $\boldsymbol{H_i}$, а вероятности $p(H_i/A)$ - ***апостериорными вероятностями гипотез*** $\boldsymbol{H_i}$.

Пример 1.19. В первой урне находится 4 белых и 6 черных шаров, а во второй – 2 белых и 7 черных шаров. Из первой урны во вторую наудачу перекладывают один шар. После этого из второй урны наудачу извлекают один шар.

1) Найти вероятность, что шар, извлеченный из второй урны, белый.

2) Известно, что из второй урны извлечен белый шар. Найти вероятность того, что из первой урны во вторую был переложен черный шар.

Решение. Обозначим события

H_1 = {из первой урны во вторую был переложен белый шар},

H_2 = {из первой урны во вторую был переложен черный шар},

A = {шар, извлеченный из второй урны, белый}.

а) По классическому определению вероятности находим вероятности гипотез H_1 и H_2 :

$p(H_1)$ - вероятность того, что из всех находящихся в первой урне шаров во вторую был переложен *белый шар* - $p(H_1) = 4 / 10 = 0{,}4$,

$p(H_2)$ - - вероятность того, что из всех находящихся в первой урне шаров во вторую был переложен *черный шар* - $p(H_2) = 6/10 = 0{,}6$,

и условные вероятности:

$p(A/H_1)$ – вероятность того, что из второй урны *будет вынут белый шар*, при условии, что в нее был переложен белый шар - $p(A/H_1) = 3/10 = 0{,}3$.

$p(A/H_2)$ - вероятность того, что из второй урны *будет вынут белый шар*, при условии, что в нее был переложен черный шар - $p(A/H_2) = 2/10 = 0{,}2$.

Тогда по формуле полной вероятности находим вероятность искомого события A

$$p(A) = p(H_1)\,p(A/H_1)+p(A/H_2)p(H_2) = 0{,}4\cdot0{,}3 +0{,}6\cdot0{,}2 = 0{,}24.$$

б) Во второй части задачи требуется найти условную вероятность $p(H_2/A)$. Для этого можно использовать формулу Байеса (1.11):

$$p(H_2/A) = \frac{p(H_2)p(A/H_2)}{\mathrm{p(H_1)\,p(A/H_1) + p(A/H_2)p(H_2)}} = \frac{0{,}6\cdot0{,}2}{0{,}4\cdot0{,}3+0{,}6\cdot0{,}2} = \frac{0{,}12}{0{,}24} = \frac{1}{2}\ .$$

Вопросы для самопроверки

1. Какие условия должны быть выполнены для проведения опытов по схеме Бернулли?
2. По какой формуле вычисляется условная вероятность?
3. Как вычисляется вероятность произведения:
   а) для независимых событий?
   б) для несовместных событий?
   в) для любых других типов событий?
4. В чем смысл формулы полной вероятности?

В результате изучения материала **раздела 1** студент может выполнить задание № 1 из методических указаний по выполнению контрольной работы по вычислительной математике, основам теории вероятностей и элементам математической статистики [8]

Раздел 2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

В этом разделе содержится материал о случайных величинах: дискретных и непрерывных. Приводятся определения и описываются способы задания случайных величин с помощью ряда распределения и функции распределения. Способы определения плотности вероятности зависят от того, какое распределение имеет случайная величина. Рассматривается решение задач, в которых необходимо вычислить числовые характеристики случайных величин, а также написать выражение для плотности вероятности случайной величины, если она имеет биномиальное распределение, равномерное распределение или нормальное распределение.

Изучение материала раздела заканчивается ответами на вопросы для самопроверки и рассмотрением репетиционного теста №2, приведенного в блоке контроля и освоения дисциплины. После того, как эта часть работы проделана, студент может приступить к выполнению задачи № 2 из методических указаний к выполнению контрольной работы по вычислительной математике, основам теории вероятностей и элементам математической статистики [8].

2.1. Описание случайных величин

2.1.1. Определение и способы задания случайной величины

Одним из основных понятий теории вероятности является понятие случайной величины. Рассмотрим некоторые примеры случайных величин:

1) число попаданий в цель при трех выстрелах.

Возможные результаты таковы: 0,1,2 или 3 раза попадания.

2) число вызовов, поступивших на телефонную станцию за сутки. Значениями может быть любое число от 1, 2, 3,….

Случайной величиной называется такая величина, которая в результате опыта может принимать различные значения, причем заранее неизвестно, какое именно, и известны вероятности, с которыми случайная величина принимает каждое конкретное значение.

Определение. Любое соотношение, устанавливающее связь между всеми возможными значениями случайной величины и их вероятностями, называется ***законом распределения*** случайной величины.

Чтобы задать случайную величину, надо указать ее закон распределения. Случайные величины принято обозначать греческими буквами ξ, η, θ, а их возможные значения – латинскими буквами с индексами x_i, y_i, z_i.

Пример 2.1. Обозначим буквой ξ число гербов, выпавших при подбрасывании монеты три раза. Это число зависит от случайных результатов подбрасывания и поэтому будет случайной величиной. В этом примере случайная величина ξ может принять четыре значения 0,1,2,3, но невозможно предсказать какое из них. Найдем вероятности этих значений. Пространство элементарных событий Ω в этом примере состоит из восьми упорядоченных троек

$$\Omega=\{\omega_1=ГГГ,\ \omega_2=ГГЦ,\ \omega_3=ГЦГ,\ \omega_4=ЦГГ,\ \omega_5=ГЦЦ,\ \omega_6=ЦГЦ,\ \omega_7=ЦЦГ,\ \omega_8=ЦЦЦ\},$$

где *Г* обозначает выпадение герба при одном подбрасывании, а *Ц* – выпадение цифры.

Обозначим через A_i событие, в котором при подбрасывании монеты появились i гербов (i=0,1,2,3). Каждое событие A_i является составным событием и содержит все элементарные события ω_i, которые привели к появлению i гербов:

$$A_i=\{\omega:\xi=i\}.$$

Следовательно,

$$A_0=\{\omega:\xi=0\}=\{ЦЦЦ\},\quad A_1=\{\omega:\xi=1\}=\{ГЦЦ,\ ЦГЦ,\ ЦЦГ\},$$

$$A_2=\{\omega:\xi=2\}=\{ГГЦ,\ ГЦГ,\ ЦГГ\},\quad A_3=\{\omega:\xi=3\}=\{ЦЦЦ\}.$$

Дополнительно предположим, что подбрасывают правильную монету. Тогда из независимости испытаний следует, что вероятность каждого элементарного события ω_i равна $\frac{1}{2}*\frac{1}{2}*\frac{1}{2}=\frac{1}{8}$. Из классического определения вероятности события A_i имеют вероятности, равные

$$p_0=P(A_0)=P\{ЦЦЦ\}=\frac{1}{8},\qquad p_1=P(A_1)=P\{ГЦЦ,ЦГЦ,\ ЦЦГ\}=\frac{3}{8},$$

$$p_2=P(A_2)=P\{ГГЦ,ГЦГ,ЦГГ\}=\frac{3}{8},\qquad p_3=P(A_3)=P\{ЦЦЦ\}=\frac{1}{8}.$$

Отметим, что все события A_i несовместны и составляют пространство элементарных несовместных событий Ω, т.е.

$$\Omega=A_0+A_1+A_2+A_3.$$

Из аксиом вероятности следует равенство

$$p(\Omega)=\mathrm{p}(A_0+A_1+A_2+A_3)=p(A_0)+p(A_1)+p(A_2)+p(A_3)=1.$$

Составим таблицу из полученных возможных значений этой случайной величины и соответствующих вероятностей :

ξ	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Пример 2.2. Производится один выстрел по плоской круглой мишени радиуса R. Учитываются только выстрелы, которые приводят к попаданию в мишень. В качестве случайной величины рассмотрим расстояние ξ от точки попадания до центра мишени. Тогда множество возможных значений случайной величины образует числовой интервал $[0, R]$. Предположим, что любая точка мишени может быть поражена с одинаковой вероятностью. Отсюда следует, что с одинаковой вероятностью случайная величина ξ принимает любое значение из интервала $[0, R]$. В этом случае вероятность того, что расстояние ξ не превзойдет числа x ($0\leq x \leq R$), можно найти из геометрического определения вероятности по формуле $P(\xi \leq x)=x/R$. Очевидно, что при $x>R$ вероятность $P(\xi \leq x)=1$, а при $x<0$ вероятность $P(\xi \leq x)=0$.

Из приведенных примеров видно, что ***случайной величиной является функция f, которая каждому элементарному событию ω ставит в соответствие число*** $f(\omega)$. Эти числа $f(\omega)$ ***называют возможными значениями случайной величины***.

В зависимости от множества возможных значений случайной величины выделяют два типа случайных величин:

а) дискретная случайная величина – это величина, значения которой можно (пересчитать) перенумеровать;

б) непрерывная случайная величина – это такая величина, значения которой заполняют целиком некоторый промежуток числовой оси или всю числовую ось.

Определение. Функция распределения F(x) случайной величины ξ определяется равенством

$$F(x) = P(\omega : f(\omega) \leq x), \qquad (2.1)$$

для всех действительных чисел x.

В примере 2.2 функция распределения определяется формулой

$$F(x) = \begin{cases} 0 & \text{при} \quad x < 0 \\ \dfrac{x}{R} & \text{при} \quad 0 \leq x < R \\ 1 & \text{при} \quad x \geq R \end{cases}.$$

2.1.2. Дискретные случайные величины

Определение. Случайную величину ξ называют ***дискретной***, если множество ее возможных значений образует конечную или бесконечную последовательность чисел, т.е. конечно или счетно.

Пусть возможные значения дискретной случайной величины ξ упорядочены по возрастанию

$$x_1 \le x_2 \le \ldots \le x_n \le \ldots .$$

Рассмотрим события A_i, содержащие все элементарные события ω, приводящие к значению x_i:

$$A_i = \{\omega:\ \xi = x_i\},\ i = 1, 2, \ldots$$

Пусть p_i *обозначает вероятность события* A_i*:*

$$p_i = P(A_i) = P(\omega:\ \xi = x_i),\ i = 1, 2, \ldots .$$

События A_i - *несовместные* события, которые составляют разбиение пространства элементарных событий Ω, т.е. $\Omega = \sum\limits_i A_i$.

Тогда для вероятностей p_i выполняются свойства

$$p_i \ge 0,\ i = 1, 2, \ldots \qquad \sum_i p_i = 1 . \qquad (2.2)$$

Закон распределения дискретной случайной величины задается **рядом распределения.**

Ряд распределения дискретной случайной величины ξ может быть представлен таблицей, в первой строке которой помещают возможные значения x_i, а во второй - вероятности p_i, соответствующие этим значениям.

ξ	x_1	x_2	$\ldots x_n \ldots$
Pi	p_1	p_2	$\ldots p_n \ldots$

Кроме ряда распределения, дискретная случайная величина может быть задана с помощью *функции распределения*.

Определение. Функция распределения F(x) случайной величины ξ это такая функция переменной *x,* которая равна вероятности того, что случайная величина примет значение, меньшее, чем заданное *x,*

$$F(x) = P(\omega: f(\omega) \le x) \qquad (2.3)$$

для всех действительных чисел *x*.

Для дискретной случайной величины функция распределения определяется как сумма вероятностей для тех значений случайной величины, которые меньше заданного *x.* Обозначим через $B(x)$ множество возможных значений случайной величины ξ, предшествующих числу *x:*

$$B(x) = \{x_i:\ x_i \le x\}. \qquad (2.4)$$

Тогда формулу (2.3) можно записать в виде

$$F(x) = P(B(x)) = \sum_{x_i \in B(x)} p_i \, . \qquad (2.5)$$

Приведем несколько примеров функций распределения дискретных случайных величин.

Пример 2.3. Правильный кубик подбрасывают один раз, и величина ξ обозначает число очков, выпавшее на его верхней грани. Построим функцию распределения этой случайной величины.

Решение. Обозначим через X возможные значения случайной величины ξ. В данном примере $X=\{1,2,3,4,5,6\}$, и вероятность появления грани с любым количеством очков равна $p_i = \frac{1}{6}$.

Напишем ряд распределения этой дискретной случайной величины

x	1	2	3	4	5	6
p	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Построим функцию распределения по формуле (2.5). Для этого на числовой оси отметим точки из множества X. Они разбивают числовую ось Оx на интервалы $(-\infty,1)$, $[1,2)$ $[2,3)$ $[3,4)$ $[4,5)$ $[5,6)$ $[6,+\infty)$.

Последовательно будем вычислять функцию распределения на каждом из указанных выше интервалов. При любом $x \in (-\infty,1)$ множество $B(x)=\{x_i : x_i \le x\}$ не содержит возможных значений случайной величины, т.е. является пустым множеством. Тогда по формуле (2.5)

$$F(x) = P\{B(x)\} = 0.$$

При любом $x \in [1,2)$ множество будет состоять из одного значения - 1: $B(x)=\{x_i : x_i \le x\}=\{1\}$. Тогда по формуле (2.5)

$$F(x) = P\{B(x)\} = p_1 = \frac{1}{6}.$$

При любом $x \in [2,3)$ множество $B(x)=\{x_i : x_i \le x\}=\{1,2\}$. Тогда по формуле

$$F(x)=P\{B(x)\}=p_1+p_2=\frac{2}{6}=\frac{1}{3}.$$

При любом $x\in[3,4)$ множество $B(x)=\{x_i : x_i\le x\}=\{1,2,3\}$. Тогда

$$F(x)=P\{B(x)\}=p_1+p_2+p_3=\frac{1}{2}.$$

При любом $x\in[4,5)$ множество $B(x)=\{x_i : x_i\le x\}=\{1,2,3,4\}$. Тогда

$$F(x)=P\{B(x)\}=p_1+p_2+p_3+p_4=\frac{2}{3}.$$

При любом $x\in[5,6)$ множество $B(x)=\{x_i : x_i\le x\}=\{1,2,3,4,5\}$.Тогда

$$F(x)=P\{B(x)\}=p_1+p_2+p_3+p_4+p_5=\frac{5}{6}.$$

При любом $x\in[6,+\infty)$ множество $B(x)=\{x_i : x_i\le x\}=\{1,2,3,4,5,6\}=X$. Тогда

$$F(x)=P\{B(x)\}=p_1+p_2+p_3+p_4+p_5+p_6=1.$$

Заметим, что при переходе от одного интервала к другому множество $B(x)$ расширяется на одно значение и от пустого множества переходит к множеству всех возможных значений $X=\{1,2,3,4,5,6\}$.

Все вычисления можно объединить в формулу

$$F(x)=\begin{cases}0 & \text{при} \quad x<1,\\ \dfrac{i}{6} & \text{при} \quad x\in[i,i+1),\ i=1,2,...,5,.\\ 1 & \text{при} \quad x\in[6,+\infty).\end{cases} \qquad (2.6)$$

Пример 2.4. Построим функцию распределения для появления числа гербов при трех подбрасываниях монеты (*пример 2.1*).

Решение. Ряд распределения был найден в *примере 2.1*.

ξ	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Обозначим через X множество всех возможных значений этой случайной величины $X=\{0, 1, 2, 3\}$. Заметим, что множество $B(x)$ при любом x является подмножеством X. Числа из множества X разбивают числовую ось на интервалы $(-\infty,0)$, $[0,1)$, $[1,2)$, $[2,3)$, $[3,+\infty)$.

Пусть x любое число из интервала $(-\infty,0)$. Тогда множество $B(x)$ не содержит значений случайной величины ξ, т.е. $B(x)=\varnothing$, следовательно, $F(x)=0$ при всех x из $(-\infty,0)$.

Возьмем любое $x\in[0,1)$. Множество $B(x)$ содержит значение 0:

$B(x)=\{0\}$ и $F(x)=p_0=\dfrac{1}{8}$.

Возьмем $x\in[1,2)$. Множество $B(x)=\{0,1\}$, и $F(x)=p_0+p_1=\dfrac{1}{8}+\dfrac{3}{8}=\dfrac{1}{2}$.

Для всех $x\in[2,3)$ множество $B(x)=\{0,1,2\}$, и $F(x)=p_0+p_1+p_2=\dfrac{1}{8}+\dfrac{3}{8}+\dfrac{3}{8}=\dfrac{7}{8}$.

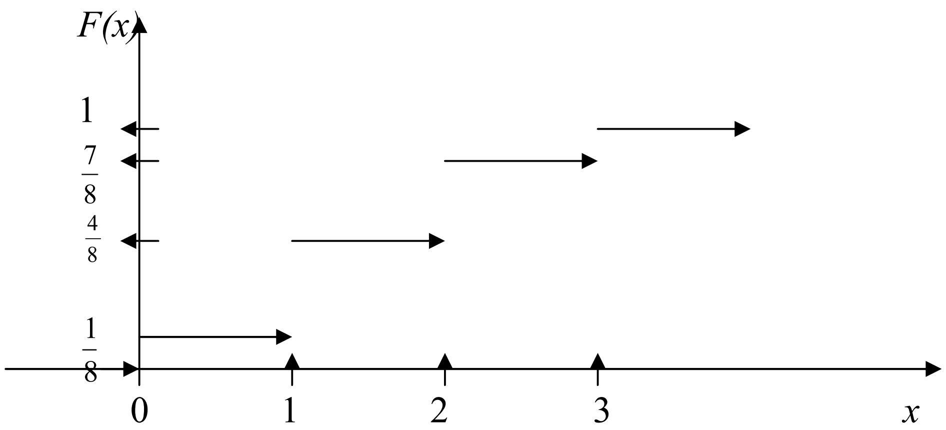
Для всех $x\in[3,\infty)$ множество $B(x)=\{0,1,2,3\}=X$. Отсюда следует

$$F(x)=p_0+p_1+p_2+p_3=\frac{1}{8}+\frac{3}{8}+\frac{3}{8}+\frac{1}{8}=1.$$

Запишем полученные значения функции распределения на отдельных интервалах в виде формулы

$$F(x)=\begin{cases}0 & \text{при всех } x\in(-\infty,0),\\ \dfrac{1}{8} & \text{при всех } x\in[0,1),\\ \dfrac{4}{8} & \text{при всех } x\in[1,2),\\ \dfrac{7}{8} & \text{при всех } x\in[2,3),\\ 1 & \text{при всех } x\in[3,+\infty)\end{cases}.$$

Построим график функции распределения $F(x)$ дискретной случайной величины



2.1.3. Непрерывные случайные величины

Кроме дискретных существуют непрерывные случайные величины. Примером непрерывной случайной величины является расстояние от точки попадания до центра мишени.

Определение. Случайная величина ξ называют ***непрерывной***, если ее значения целиком заполняют конечный или бесконечный интервал числовой оси.

Определение. Непрерывная случайная величина ξ называется ***абсолютно непрерывной,*** если существует такая функция $\boldsymbol{f(x)}$, что для любых действительных чисел x функция распределения этой случайной величины $\boldsymbol{F(x)}$ представима в виде

$$F(x) = \int_{-\infty}^{x} f(t)dt \,, \qquad (2.7)$$

т. е. функция распределения $\boldsymbol{F(x)}$ этой случайной величины является интегралом с переменным пределом от $\boldsymbol{f(x)}$. ***Функция f(x) называется плотностью распределения вероятностей*** случайной величины ξ такой, что

$$f(x) = F^{I}(x).$$

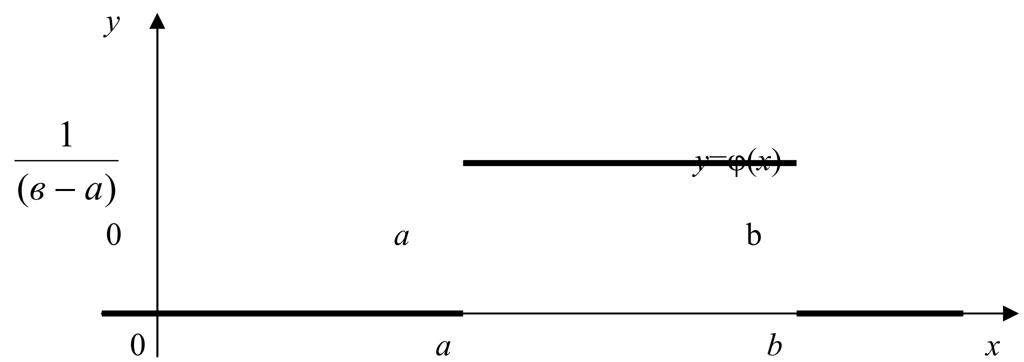
Рассмотрим пример абсолютно непрерывной случайной величины.

Пример 2.5. Равномерное распределение

Случайная величина ξ распределена ***равномерно*** на промежутке **[*a,b*],** если ее плотность распределения вероятностей задается равенством

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{при} \quad x \in [a,b] \\ 0, & \text{при} \quad x \notin [a,b] \end{cases}, \qquad (2.8)$$

и имеет график



Построим функцию распределения **$F(x)$** этой случайной величины, используя равенство (2.7). При всех $x<a$ по определению $f(x)$ подынтегральная функция $f(t)=0$. Следовательно, функция распределения $F(x)=0$.

При всех $a\le x<b$ по определению $f(x)$ подынтегральная функция $f(t)=\frac{1}{b-a}$. Тогда из равенства (2.7) получаем

$$F(x)=\int_{-\infty}^{x} f(t)dt=\int_{-\infty}^{a} f(t)dt+\int_{a}^{x} f(t)dt=\int_{a}^{x}\frac{1}{b-a}dt=\frac{x-a}{b-a}\ .$$

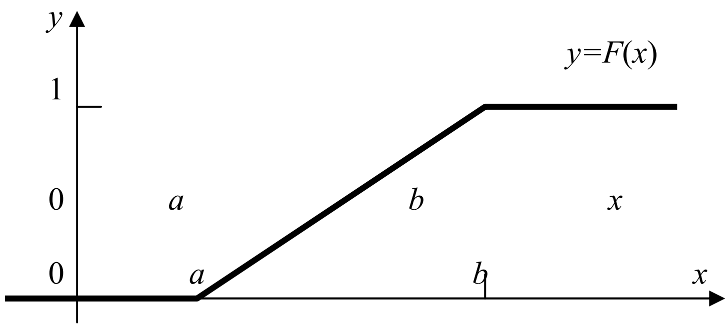
При $x\ge b$ по свойству определенного интеграла и определению подынтегральной функции получаем следующее равенство

$$F(x)=\int_{a}^{x} f(t)dt=\int_{-\infty}^{a} f(t)dt+\int_{a}^{b} f(t)dt+\int_{b}^{x} f(t)dt=\int_{a}^{b}\frac{1}{b-a}dt=1\ .$$

Следовательно, функция распределения равномерного закона имеет вид:

$$F(x)=\begin{cases}0, & \text{если } x<a,\\ \dfrac{x-a}{b-a}, & \text{если } a\le x<b,\\ 1, & \text{если } x\ge b\end{cases}\qquad (2.9)$$

График функции распределения имеет вид:



2.1.4. Функция распределения и ее свойства

Для функции распределения $F(x)$ случайной величины ξ имеют место следующие **свойства:**

1)* Для всех *x выполняются неравенства $0\le F(x)\le 1$.

2) **Функция распределения** случайной величины есть **неубывающая функция** аргумента ***x****:* $\boldsymbol{F(x_1) \le F(x_2)}$, если $x_1 \le x_2$. (2.10)

3) $F(-\infty) = 0$, $F(+\infty) = 1$.

4) Функция распределения непрерывна справа $F(x+0)=F(x)$.

5) **Вероятность попадания** случайной величины ξ **в интервал** $\boldsymbol{(a,b]}$ можно вычислить по формуле

$$P\{a<\xi\le b\}=F(b)-F(a). \quad (2.11)$$

2.1.5. Плотность распределения вероятности и ее свойства

Плотность распределения вероятностей $f(x)$ удовлетворяет следующим свойствам.

1) В точках x непрерывности $f(x)$ имеет место равенство

$$F'(x) = f(x). \quad (2.12)$$

.

2) Для всех x плотность распределения вероятностей **неотрицательна,** *т.е.* $f(x) \ge 0$.

3) Вероятность попадания случайной величины ***в любой интервал*** **[a,b]** ***равна***

$$P(a < x < b) = P(a \le x < b) = P(a < x \le b) = = P(a \le x \le b) = \int_a^b f(x)dx. \quad (2.13)$$

Это свойство означает, что удаление конечного числа точек из промежутка $[a,b]$ не влияет на величину вероятности попадания в этот интервал.

4) Условие нормировки $\int_{-\infty}^{+\infty} f(x)dx = 1$. (2.14)

График $y=f(x)$ называют ***кривой распределения***. Графически свойство 3**)** означает, что вероятность попадания случайной величины в интервал $[a,b]$ численно равна площади S криволинейной трапеции, ограниченной сверху кривой распределения $y=f(x)$, снизу - осью Оx, вертикальными прямыми $x=a$ и

$x=b$. Графически свойство 4) означает, что площадь криволинейной трапеции под всей кривой распределения $y=f(x)$ равна 1.

Вопросы для самопроверки

1. Какая величина называется дискретной случайной величиной?
2. Какую случайную величину называют непрерывной?
3. Что такое ряд распределения случайной величины?
4. Сформулируйте определение функции распределения случайной величины и ее свойства.
5. Что такое плотность вероятности случайной величины и ее свойства?

2.2. Числовые характеристики случайных величин

2.2.1. Математическое ожидание случайной величины

Закон распределения случайной величины является наиболее полной ее характеристикой. Он одновременно указывает и на значение случайной величины и на его вероятность. Однако часто в теории вероятностей и в ее приложениях большую роль играют постоянные числа, которые можно получить, используя законы распределения. В этом параграфе рассмотрим математическое ожидание или среднее значение случайной величины.

Пусть ξ обозначает дискретную случайную величину с рядом распределения

x_i	$x_1\ \ x_2\ ,\ldots,x_n,\ldots$
p_i	$p_1\ \ p_2\ ,\ldots,\ p_n,\ldots$

Определение. ***Математическим ожиданием*** дискретной случайной величины ξ называется сумма произведений возможных значений x_i на соответствующие им вероятности p_i. Будем обозначать математическое ожидание как $M(\xi)$. Тогда можно написать, что математическое ожидание вычисляется по формуле

$$M(\xi)=x_1 p_1 + x_2 p_2 + \ldots x_n p_n + \ldots = \sum_{i=1}^{\infty} x_i p_i \ , \qquad (2.15)$$

если числовой ряд сходится абсолютно. Если числовой ряд (2.15) расходится или сходится условно, то в этом случае математическое ожидание случайной величины ξ не существует.

Пример 2.6. Найдем среднее число очков при одном подбрасывании правильного кубика. Используя ряд распределения из примера 2.3

x_i	1	2	3	4	5	6
p_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

по формуле (2.15) находим математическое ожидание

$$M(\xi)=1*\frac{1}{6}+2*\frac{1}{6}+3*\frac{1}{6}+4*\frac{1}{6}+5*\frac{1}{6}+6*\frac{1}{6}=3{,}5.$$

Пример 2.7. Найдем среднее число суммы очков на двух кубиках.

Используя ряд распределения

η	2	3	4	5	6	7	8	9	10	11	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

,

получаем

$$\mathrm{M}(\xi)=2\cdot\frac{1}{36}+3\cdot\frac{2}{36}+4\cdot\frac{3}{36}+5\cdot\frac{4}{36}+6\cdot\frac{5}{36}+7\cdot\frac{6}{36}+8\cdot\frac{5}{36}+9\cdot\frac{4}{36}+10\cdot\frac{3}{36}+11\cdot\frac{2}{36}+12\cdot\frac{1}{36}=\frac{252}{36}=7.$$

Пусть ξ обозначает *непрерывную* случайную величину с плотностью вероятности $f(x)$.

Определение. Математическим ожиданием $M(\xi)$ абсолютно непрерывной случайной величины ξ называется величина, равная

$$\int_{-\infty}^{+\infty} xf(x)dx, \qquad (2.16)$$

если этот несобственный интеграл сходится абсолютно. В противном случае математическое ожидание случайной величины ξ не существует.

Пример 2.8. Пусть случайная величина ξ имеет равномерный закон распределения. В этом случае плотность вероятности $f(x)$ имеет вид

$$f(x)=\begin{cases}\dfrac{1}{b-a}.\text{если } x\in[a,b],\\ 0,\ \text{если } x\notin[a,b]\end{cases}, \quad . \qquad (2.17)$$

Тогда по формуле (2.16) получаем формулу для вычисления математического ожидания равномерно-распределенной случайной величины

$$M(\xi)=\int_{-\infty}^{+\infty}xf(x)dx=\int_{-\infty}^{a}xf(x)dx+\int_{a}^{b}xf(x)dx+\int_{b}^{+\infty}xf(x)dx=\int_{a}^{b}x\frac{1}{(b-a)}dx=\frac{b^2-a^2}{2(b-a)}=\frac{a+b}{2}.$$

Пусть ξ обозначает дискретную случайную величину с рядом распределения

x_i	$x_1 \quad x_2, \ldots, \quad x_n, \ldots$
p_i	$p_1 \quad p_2, \ldots, \quad p_n, \ldots$

,

и $g(x)$ – некоторая функция переменной x. Новая случайная величина $\eta = g(\xi)$ будет дискретной случайной величиной с рядом распределения

$g(x_i)$	$g(x_1) \quad g(x_2), \cdots, g(x_n)$
p_i	$p_1 \quad p_2, \ \ldots, \quad p_n \ldots$

,

причем вероятности этих значений остаются теми же, что и для случайной величины ξ, а значениями будут числа $g(x_i)$. Тогда математическое ожидание случайной величины $\eta = g(\xi)$ можно вычислить по формуле

$$M(\eta)=\sum_{i=1}^{\infty}g(x_i)p_i \ . \qquad (2.18)$$

Пусть ξ обозначает случайную величину с плотностью вероятности $f(x)$. Тогда математическое ожидание случайной величины $\eta = g(\xi)$ можно вычислить по формуле

$$M(\eta)=\int_{-\infty}^{+\infty}g(x)f(x)dx\,, \qquad (2.19)$$

если несобственный интеграл сходится абсолютно.

Определение. **Математическое ожидание** дискретной случайной величины $\theta = g(\xi,\eta)$ определяется формулой

$$M(\theta) = M(g(\xi,\eta)) = \sum_{i,j} g(x_i, y_j) p_{ij} \ , \qquad (2.20)$$

если числовой ряд (2.20) сходится абсолютно.

Свойства математического ожидания

1. Если случайная величина принимает только одно значение $\xi=C$, то $M(C)=C$.

2. При умножении случайной величины на постоянное число C математическое ожидание случайной величины ξ умножается на это же число, т.е. справедливо равенство $\boldsymbol{M(C\xi)=CM(\xi)}$.

3. Свойство линейности. При сложении случайных величин математическое ожидание алгебраической суммы случайных величин равно алгебраической сумме их математических ожиданий, т.е. справедливо равенство $\boldsymbol{M(\xi+\eta)=M(\xi)+M(\eta)}$.

В частности, $\boldsymbol{M(\xi+C)=M(\xi)+C}$.

4. Мультипликативное свойство. Если случайные величины ξ, η независимы, то $\boldsymbol{M(\xi\eta)=M(\xi)M(\eta)}$.

2.2.2. Дисперсия случайной величины

Из определения математического ожидания следует, что оно определяет среднее значение случайной величины. Дисперсия характеризует среднюю величину отклонения значений случайной величины от математического ожидания.

Пусть ξ обозначает дискретную или абсолютно непрерывную случайную величину.

Определение. Моментом второго порядка случайной величины ξ называется математическое ожидание квадрата этой случайной величины, т. е. число $M(\xi^2)$.

Пусть в формулах (2.18) и (2.19) функция $g(x)=x^2$. Тогда для моментов второго порядка случайной величины ξ имеют место формулы

$$M(\xi^2) = \sum_{i=0}^{\infty} x_i^2 p_i \ , \tag{2.21}$$

$$M(\xi^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx \ . \tag{2.22}$$

Величина ξ-$M(\xi)$ определяет отклонение случайной величины ξ от математического ожидания $M(\xi)$.

Определение. Дисперсией случайной величины ξ называется момент второго порядка случайной величины (ξ - $M(\xi)$). Дисперсию обозначают $D(\xi)$. Таким образом, дисперсия случайной величины ξ определяется формулой

$$D(\xi)=M[(\xi- \mathrm{M}(\xi))^2] \ . \tag{2.23}$$

Стандартным или средним квадратическим отклонением называют величину, равную квадратному корню из дисперсии и обозначают $\sigma(\xi)$: $\sigma(\xi) = \sqrt{D(\xi)}$.

Из равенств (2.18), (2.19) для моментов второго порядка следуют формулы для вычисления дисперсии дискретной и абсолютно непрерывной случайных величин соответственно

$$D(\xi) = \sum_{i=0}^{\infty} (x_i - M(\xi))^2 p_i \tag{2.24}$$

$$D(\xi) = \int_{-\infty}^{+\infty} (x - M(\xi))^2 \varphi(x) dx \tag{2.25}$$

Дисперсия не существует, если ряд (2.24) или несобственный интеграл (2.25) расходятся.

Свойства дисперсии

1. Для любой случайной величины ξ выполняется неравенство $D(\xi) \geq 0$.

2. При умножении случайной величины ξ на постоянное число C дисперсия умножается на квадрат этого числа, т.е. справедливо равенство

$$D(C\xi)=C^2D(\xi).$$

3. Справедлива следующая формула для вычисления дисперсии

$$D(\xi)=M(\xi^2)- M^2(\xi) \ , \tag{2.26}$$

то есть дисперсия случайной величины равна разности второго момента этой величины и квадрата математического ожидания этой же величины.

4. Если случайные величины ξ и η независимы, дисперсия их алгебраической суммы равна сумме дисперсий, т.е.

$$D(\xi+\eta)=D(\xi)+D(\eta). \qquad (2.27)$$

5. Дисперсия постоянной величины C равна нулю.

Пусть ξ обозначает *дискретную* случайную величину с рядом распределения

x_i	$x_1 \quad x_2, \ldots, x_n \ldots$
p_i	$p_1 \quad p_2, \ldots, p_n \ldots$

.

В этом случае согласно свойству 3 дисперсия вычисляется по формуле

$$\mathrm{D}(\xi)=\sum_{i=1}^{\infty} x_i^2 p_i - \left(\sum_{i=1}^{\infty} x_i p_i\right)^2. \qquad (2.28)$$

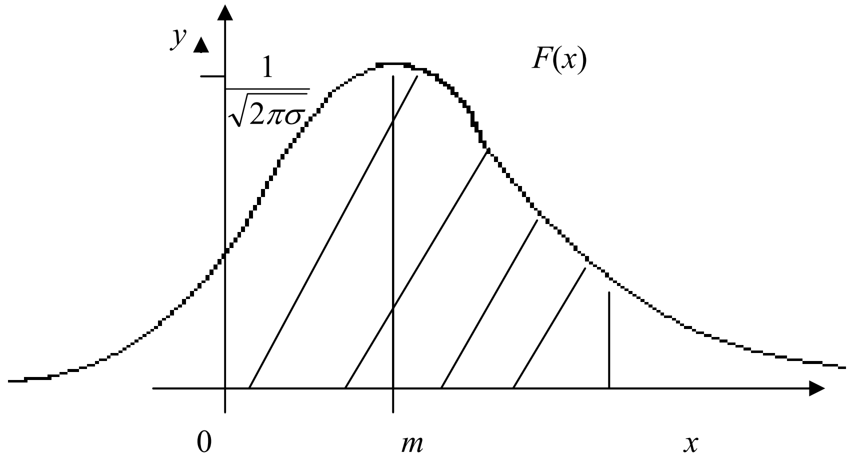
2.2.3. Нормальное распределение случайной величины

Нормальное распределение часто используется для описания случайных явлений, в которых на результат измерения влияет большое число независимых случайных факторов.

Определение. Случайная величина ξ имеет ***нормальное или гауссовское распределение,*** если ее плотность распределения вероятностей при всех x задается равенством

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}. \qquad (2.29)$$

Числа m и σ называются параметрами распределения: параметр m может быть любым действительным числом: $-\infty < m < +\infty$, а параметр σ – положительным: $\sigma>0$. Символическая запись $\xi \sim \mathrm{N}(m,\sigma^2)$ означает, что случайная величина ξ имеет нормальное распределение с параметрами m и σ^2.



Отметим некоторые свойства графика этой функции (кривой нормального распределения).

Во-первых, функция $\varphi(x)$ принимает максимальное значение $\dfrac{1}{\sqrt{2\pi}\sigma}$ при $x=m$.

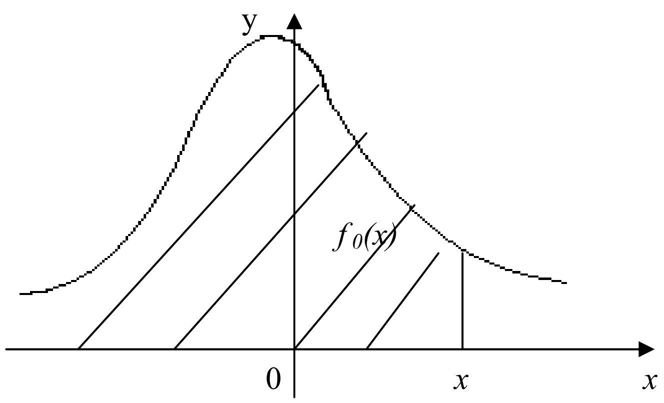
Во-вторых, функция $\varphi(x)$ симметрична относительно вертикальной прямой $x=m$.

В-третьих, асимптотой кривой нормального распределения является ось Ox. Особую роль играет нормальное распределение с параметрами $m=0$, $\sigma=1$, которое часто называют ***стандартным*** **(или нормированным)** нормальным распределением.

Плотность стандартного нормального распределения имеет вид

$$f_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad (2.30)$$

Ниже приведен график $y = f_0(x)$.



Функция распределения случайной величины, имеющей нормальный закон, может быть представлена в виде несобственного интеграла

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{\frac{-(t-m)^2}{2\sigma^2}} dt. \qquad (2.31)$$

Функцию распределения стандартного *нормального* закона

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{\frac{-t^2}{2}} dt \qquad (2.32)$$

часто называют функцией Лапласа, для которой имеются таблицы значений, широко используемые в статистических исследованиях. Рассмотрим свойства нормального распределения.

Свойство 1. Пусть случайная величина ξ имеет нормальное распределение с параметрами m и σ, т.е. $\xi \sim \mathrm{N}(\mathrm{m},\sigma^2)$. Тогда математическое ожидание равно параметру m, а дисперсия равна σ^2, т.е.

$$M(\xi)=m; \qquad D(\xi)=\sigma^2.$$

Свойство 2. Между функциями распределения $F(x)$ и $\Phi(x)$ имеет место следующее соотношение

$$F(x)=\Phi\left(\frac{x-m}{\sigma}\right). \tag{2.33}$$

Таблицы значений функции $\Phi(x)$ не содержат значений при $x<0$. В таких случаях можно использовать следующее свойство.

Свойство 3. При любых значениях x имеет место равенство

$$\Phi(x)=1-\Phi(-x). \tag{2.34}$$

Следующее свойство позволяет вычислить вероятность попадания случайной величины в интервал, используя таблицы значений функции $\Phi(x)$.

Свойство 4. Пусть случайная величина ξ имеет нормальное распределение с параметрами m и σ, т.е. $\xi \sim N(m,\sigma^2)$. Тогда вероятность попадания случайной величины ξ в интервал $[a,b]$ можно найти по формуле

$$P\{a \le \xi \le b\} = F(b)\text{-}F(a) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \tag{2.35}$$

В частности, для симметричного интервала относительно m имеет место формула для любого $\Delta > 0$:

$$P(|\xi-m|<\Delta)=2\Phi\left(\frac{\Delta}{\sigma}\right)-1. \tag{2.36}$$

Формула (2.36) непосредственно следует из (2.35), в которой надо положить $a = m - \Delta,\ \mathrm{b} = \mathrm{m} + \Delta$ и использовать свойство, что $\Phi\left(-\frac{\Delta}{\sigma}\right) = 1 - \Phi\left(\frac{\Delta}{\sigma}\right)$.

Тогда получим

$$P\{|\xi - m < \Delta\} = |P\{m - \Delta \le \xi \le m + \Delta\} = \Phi\left(\frac{\Delta}{\sigma}\right) - \Phi\left(\frac{-\Delta}{\sigma}\right) = 2\Phi\left(\frac{\Delta}{\sigma}\right) - 1.$$

Свойство 5 (закон трех сигм). Пусть случайная величина ξ имеет нормальное распределение с параметрами *m* и σ, т.е. $\xi \sim N(m,\sigma^2)$. Тогда с вероятностью больше 0,99 значения случайной величины содержатся в интервале [*m*-3*σ, m*-3*σ*].

Действительно, по свойству 4, $P(|\xi - m| < 3\sigma) = 2\Phi\left(\frac{3\sigma}{\sigma}\right) - 1 = 2\Phi(3) - 1$. Из таблицы функций $\Phi(x)$ находим значение $\Phi(3)$=0,9987. Отсюда следует, что

$$P(|\xi - m| < 3\sigma) = 2\Phi(3) - 1 \approx 0{,}997.$$

2.2.4. Биномиальное распределение

Биномиальное распределение служит вероятностной моделью для многих явлений. Рассмотрим последовательность *n* независимых испытаний Бернулли, в каждом из которых может произойти один из двух исходов. Обозначим через *p* вероятность «успеха» в отдельном испытании и - *q* вероятность «неудачи». Для каждого отдельного испытания *i* введем случайную величину ξ_{i}, которая может принимать два значения: 1, если испытание закончилось «успехом» и 0, если – «неудачей». Случайная величина η - число «успехов» при *n* независимых испытаниях в схеме Бернулли будет равна сумме независимых случайных величин ξ_{i}, т.е. $\eta = \xi_1 + \xi_2 + \cdots + \xi_{\mathrm{n}}$. (2.37)

Отсюда следует, что случайная величина *η* может принимать возможные значения *m*=0,2,…,*n.* Вероятность того, что случайная величина *η* после завершения всех испытаний примет значение *m* можно найти по формуле Бернулли

$$P(\eta = m) = C_n^m p^m q^{n-m} \quad, m=0, 1, \dots, n. \qquad (2.38)$$

Определение. Случайная величина η имеет ***биномиальное распределение*** с параметрами n и p, если она принимает значения, m=0, 1,…, n с вероятностями по формулам (2.38).

Вычислим математическое ожидание и дисперсию случайной величины, имеющей биномиальное распределение. Из формулы (2.20) и свойств математического ожидания и дисперсии для суммы несовместных случайных величин следует, что

$$M(\eta) = M(\xi_1) + M(\xi_2) + \cdots + M(\xi_n),$$

$$D(\eta) = D(\xi_1) + D(\xi_2) + \cdots + D(\xi_n).$$

Математическое ожидание и дисперсия случайной величины ξ_i будут соответственно равны

$$M(\xi_i)=1p+0q=p,$$

$$D(\xi_i)=1^2 p+0^2 q - p^2=pq.$$

Для случайной величины η, имеющей *биномиальное* распределение с параметрами n и p, математическое ожидание и дисперсия имеют вид:

$$M(\eta)=np, \qquad D(\eta)=npq.$$

Пример 2.9. В коробке находятся 3 однотипных изделия, при этом каждое может быть или бракованным, или стандартным. Рассмотрим случайную величину, которая определяет число бракованных изделий в коробке. «Успехом» отдельного испытания будем считать наличие в коробке некоторого числа бракованных изделий. Тогда случайная величина ξ может принять значения x_1=0, x_2=1, x_3=2, x_4=3 (варианты количества бракованных изделий) и имеет биномиальное распределение. Требуется найти вероятности событий, которые будут соответствовать значениям случайной величины. Последовательно при i=0, 1, 2, 3 из формулы (2.38) получим вероятности для возможных значений этой случайной величины:

$$p_1=P(\xi=0)=C_3^0 p^0 q^3 = q^3 \; ; \qquad p_2=P(\xi=1)= C_3^1 p^1 q^{3-1} = 3pq^2;$$

$$p_3=P(\xi=2)=C_3^2 p^2 q^{3-2} = 3p^2 q \; ; \; p_4=P(\xi=3)= C_3^3 p^3 q^{3-3} = p^3.$$

Теперь можно записать таблицу для ряда распределения:

ξ	0	1	2	3
p_i	q^3	$3pq^2$	$3p^2q$	p^3

Математическое ожидание и дисперсию находим по формулам (2.21) и (2.23) соответственно

$$M(\xi)=3p, \quad D(\xi)=3\,p\,q\,.$$

2.2.5. Распределение Пуассона

Определение. Случайная величина ξ распределена **по закону Пуассона с параметром распределения $\lambda>0$**, если она принимает значения m= 0,1, 2,…
с вероятностями

$$p_m = P\,(\xi = m\,) = \frac{\lambda^m}{m!}e^{-\lambda}\ ,\quad m= 0,1,2,\ldots \qquad (2.39)$$

***Пример* 2.10.** Рассмотрим случайную величину ξ, равную числу покупателей, посетивших супермаркет в промежутке времени от t_0 до T. Появление покупателей - случайные события и происходят в случайные моменты времени.

Сделаем следующие предположения.

1. Вероятность появления одного покупателя за малый промежуток времени Δ пропорциональна Δ, т. е. равна

$$a\Delta + o(\Delta),\ a > 0,$$

где $o(\Delta)$ - бесконечно малая величина при $\Delta \to 0$.

2. Если за малый промежуток времени Δ уже произошло одно событие, то условная вероятность появления в этом же промежутке другого события стремится к 0 при $\Delta \to 0$.

3. События на непересекающихся промежутках времени являются независимыми случайными величинами.

В этих условиях можно доказать, что число покупателей, посетивших супермаркет в промежутке времени от t_0 до T распределено по закону Пуассона с параметром $\lambda = a$T.

Вопросы для самопроверки

1. Как вычислить математическое ожидание и дисперсию:

   а) для дискретной случайной величины?
   б) для непрерывной случайной величины?
2. Как найти плотность вероятности для случайной величины, имеющей равномерное распределение?
3. Что такое нормально распределенная случайная величина?
4. Какие другие законы распределения известны?

Раздел 3. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В различных областях деятельности возникает необходимость в обработке различных объемов экспериментальных и статистических данных. Например, такие задачи как упорядочение результатов измерений, выявление закономерностей случайных явлений, расчет надежности сложных технических систем могут быть решены вероятностно-статистическими методами. Математическая статистика базируется на основных понятиях теории вероятностей.

В этом разделе определяется понятие выборки, генеральной совокупности, доверительного интервала, состоятельной и несмещенной оценок. На конкретном примере показано, как промоделировать дискретную случайную величину методом жребия при помощи n случайных чисел и сделать проверку гипотез о совпадении или различии теоретических данных и полученных экспериментальных данных. Кроме того, рассматриваются точечные и интервальные оценки неизвестных параметров. Для более глубокого изучения этой темы студенты должны выполнить лабораторные работы по темам:

1) моделирование дискретной случайной величины методом жребия;
2) точечные и интервальные оценки.

После проделанной работы студент может приступить к выполнению контрольной задачи № 3 и № 4 из методических указаний к выполнению контрольной работы по вычислительной математике, основам теории вероятностей и элементам математической статистики [8].

Основные определения

Для изучения какого-либо явления необходимо его многократно наблюдать в одинаковых условиях, т.е. надо провести серию испытаний (наблюдений). Предметом математической статистики является изучение случайных величин по результатам наблюдений. При этом выявляемые закономерности, которым подчиняется исследуемая случайная величина, на физическом уровне обусловливается реальным комплексом условий, в которых она наблюдается, а на математическом уровне задается соответствующим законом распределения вероятностей.

Допустим, что исследуется случайная величина ξ.

Конечная или бесконечная совокупность всех мыслимых наблюдений над случайной величиной ξ называется *генеральной совокупностью* значений случайной величины ξ или просто генеральной совокупностью.

Изучение всех элементов генеральной совокупности часто оказывается невозможным. Тогда рассматривается некоторая ее часть.

Часть генеральной совокупности, по которой делается заключение о свойствах случайной величины, называется *выборкой.* Чтобы были сделаны правильные выводы о случайной величине по выборке, она должна быть *представительной* (т.е. иметь большой объем).

Выборку можно рассматривать как последовательность чисел $x_1, x_2, \ldots, x_n$, полученных в результате проведения n независимых испытаний над случайной величиной ξ.

3.2. Систематизация выборки

Пусть $x_1, x_2, \ldots, x_n$ - реализация случайной выборки.

Вариационным рядом называется последовательность элементов реализации случайной выборки, расположенных в неубывающем порядке, при этом одинаковые элементы могут повторяться.

Пусть наблюдаемая в эксперименте случайная величина ξ дискретна и принимает k различных значений $(k < n)$, которые обозначим $z_1, z_2, \ldots, z_k$.

Примечание. В дальнейшем величины, получаемые из опытных данных, будем обозначать теми же символами, что и теоретические аналоги, со знаком * вверху.

Относительной частотой значения z_i (*или статистической вероятностью события* $\{\xi = z_i\}$) называется случайная величина

$$p_i^* = \frac{\mu_i}{n}, \qquad (3.1)$$

где μ_i - частота значения z_i, которое принимает случайная величина (т.е. число элементов выборки $\vec{X}=(X_1, X_2, \ldots, X_n)$,принявших значение z_i). Согласно закону больших чисел, p_i^* сближается с вероятностью $P\{\xi = z_i\}$ при $n \to \infty$, т.е. относительные частоты p_i^* можно рассматривать при больших объемах выборки в качестве приближенных значений (оценок) для неизвестных вероятностей $P\{\xi = z_i\}$.

Статистическим рядом называется последовательность разных значений случайной величины, расположенных в возрастающем порядке, с указанием значений относительных частот. Статистический ряд, как правило, записывается в виде таблицы (табл.3.1).

Таблица 3.1

Z_i	Z_1	Z_2	Z_3	. . .	Z_n
P^*_i	P^*_1	P^*_2	P^*_3	. . .	P^*_n

Важной характеристикой выборки является эмпирическая функция распределения.

Эмпирической (выборочной) функцией распределения $F_n^*(x)$, построенной по случайной выборке, называется относительная частота события $\{\xi =< x\}$:

$$F_n^*(x) = P^*\{\xi \le x\} = \frac{\mu(x)}{n}, \quad -\infty < x < \infty, \tag{3.2}$$

где $\mu(x)$ - случайная величина, равная числу тех наблюдений X_1, X_2, ..., X_n, значения которых не превосходят x.

Чтобы получить значение эмпирической функции распределения при данном значении x для реализации выборки $\vec{x} = (x_1, \ldots, x_n)$, надо подсчитать значение $\mu(x)$ для реализации выборки, т.е. просуммировать значения частот тех элементов z_i, которые меньше x. Получим

$$F_n^*(x) = \frac{1}{n} \sum_{z_i \le x} m_i . \tag{3.3}$$

Свойства эмпирической функции распределения $F_n^*(x)$ аналогичны свойствам обычной функции распределения, а именно:

1) $F_n^*(x)$ - неубывающая функция по x, является ступенчатой со скачками в точках z_k,

2) $0 \le F_n^*(x) \le 1$ для любого x, причем полагают, что

$$F_n^*(-\infty) = 0, \qquad F_n^*(\infty) = 1.$$

На рис. 3.1 приведен пример эмпирической функции распределения.

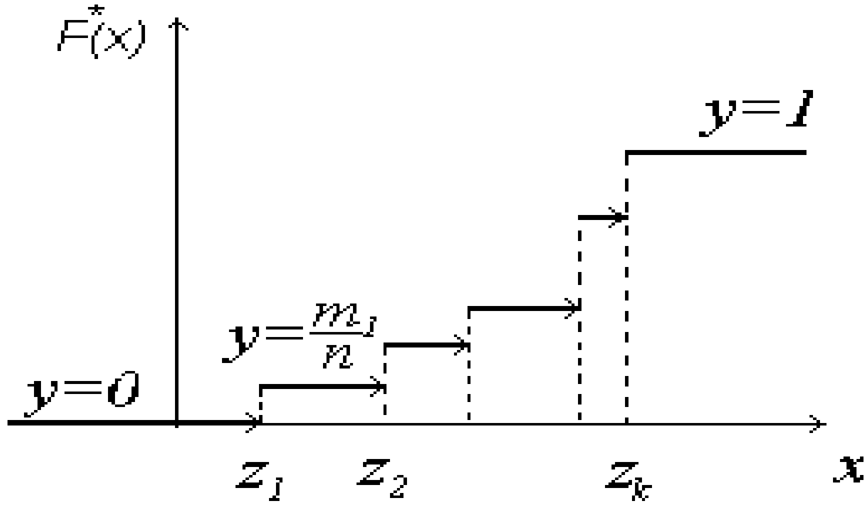


Рис. 3.1

Эмпирическая функция распределения является приближенным значением (т.е. оценкой) теоретической функции распределения наблюдаемой случайной величины ξ:

$$F_n^*(x) \approx F(x).$$

Пример 3.1. Проводятся измерения деталей с точностью до одного миллиметра. Оказалось, что отклонения диаметров изготовленных деталей от заданного размера составили следующую выборку объема $n = 20$: 0, -2, -4, 3, 0, 0, -1, 2, -2, -1, 0, -1, 3, 2, 0, -1, -2, 0, -1, 2.

Построить ***вариационный*** и ***статистический*** ряды, полигон относительных частот, эмпирическую функцию распределения.

Решение. Вариационным рядом заданной выборки будет последовательность: -4, -2, -2, -2, -1, -1, -1, -1, -1, 0, 0, 0, 0, 0, 0, 2, 2, 2, 3, 3.

Статистический ряд представим таблицей (табл.3.2):

Таблица 3.2

z_i	-4	-2	-1	0	2	3
p_i^*	$\frac{1}{20}$	$\frac{3}{20}$	$\frac{5}{20}$	$\frac{6}{20}$	$\frac{3}{20}$	$\frac{2}{20}$

Полигон относительных частот этого распределения изображен на рис.3.2.

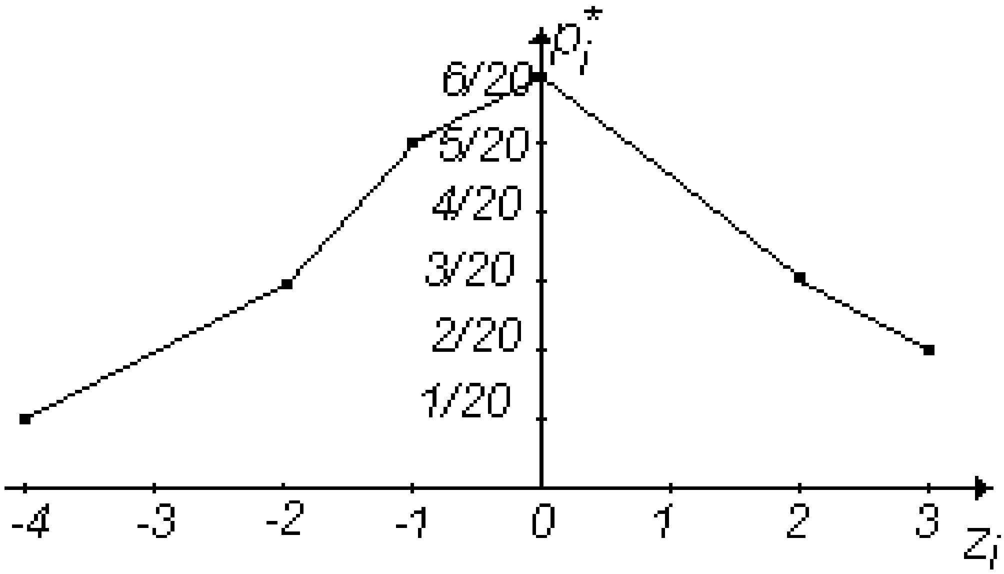


Рис. 3.2

Для полученного статистического ряда вычислим значения эмпирической функции распределения, используя формулу (3.3)

$$F_{20}^*(x) = \begin{cases} 0, \text{ при } x < -4, \\ \dfrac{1}{20}, \text{ при } -4 \le x < -2, \\ \dfrac{4}{20}, \text{ при } -2 \le x < -1, \\ \dfrac{9}{20}, \text{ при } -1 \le x < 0, \\ \dfrac{15}{20}, \text{ при } 0 \le x < 2, \\ \dfrac{18}{20}, \text{ при } 2 \le x < 3, \\ 1, \text{ при } x \ge 3. \end{cases}$$

График этой функции представлен на рис. 3.3.

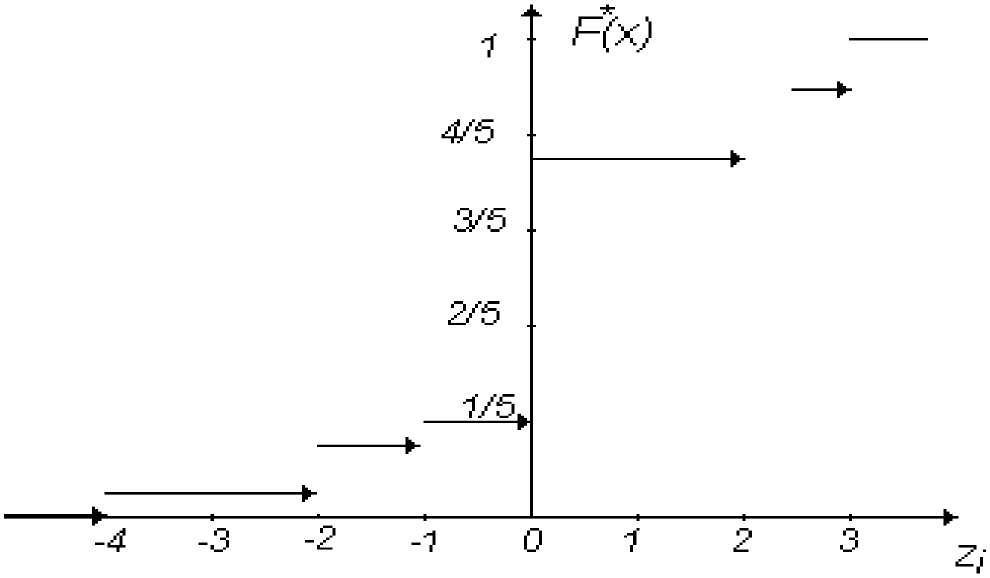


Рис. 3.3

Гистограмма

Если наблюдаемая случайная величина ξ непрерывна или объем выборки n большой, то вариационный и статистический ряды будут трудно обозримыми множествами, практически не будет равных элементов выборки. В этом случае используется процедура группировки выборки, которую рассмотрим для реализации выборки $\vec{x} = (x_1, \ldots, x_n)$. Интервал возможных значений ξ делят точками $z_0 < z_1 < \ldots < z_k$ на k непересекающихся полуинтервалов (разрядов) $\Delta_i = [z_{i-1}, z_i)$, $i = \overline{1,k}$. Для каждого разряда Δ_i подсчитывают частоту m_i - число элементов выборки, попавших в этот разряд. При этом $\sum_{i=1}^{k} m_i = n$. В интервал включают значения, большие или равные нижней границе и меньше верхней границы. Далее находят относительные частоты (статистические вероятности) $p_i^* = \frac{m_i}{n}$. Группированные данные удобно представить в виде интервального статистического ряда – последовательности пар (Δ_i, p_i^*), или в виде таблицы (табл. 3.3). Часто группу элементов выборки, входящих в интервал Δ_i, заменяют средней точкой $\bar{z}_i = \frac{z_{i-1} + z_i}{2}$.

Таблица 3.3

Δ_i	Δ_1	Δ_2	$\ldots$	Δ_k
p^*_i	p_1^*	p_2^*	$\ldots$	p_k^*

Обычно длина разрядов выбирается одинаковой, т.е. равной $h = \dfrac{z_k - z_0}{k}$. Число разрядов k выбирается в зависимости от объема выборки n так, чтобы построенный ряд не был громоздким и в то же время позволял выявить характерные особенности изменения случайной величины. Для определения k можно рекомендовать формулу Стерджеса

$$k = 1 + 1{,}44 \ln n, \qquad (3.4)$$

которая дает нижнюю оценку величины k. В качестве значения k следует брать ближайшее целое число.

Группированный статистический ряд наглядно можно изобразить в виде гистограммы. Для ее построения на оси абсцисс откладывают разряды Δ_i длиною h, и на каждом из них, как на основании, строят прямоугольник. В результате получают ступенчатую фигуру, которую называют ***гистограммой.***

Высота *i*–го частичного прямоугольника при построении *гистограммы частот* равна отношению m_i / h (плотность частоты).

Площадь *i*–го частичного прямоугольника численно равна $\dfrac{m_i}{h} \cdot h = m_i$,

а площадь гистограммы частот численно равна объему выборки, т.е.

$$\sum_{i=1}^{k} m_i = n\ . \qquad (3.5)$$

При построении *гистограммы относительных частот:*

высота *i*–го частичного прямоугольника равна отношению относительной частоты к длине интервала $\dfrac{m_i / n}{h}$ (плотность относительной частоты); площадь *i*–го частичного прямоугольника численно равна $\dfrac{m_i / n}{h} \cdot h = \dfrac{m_i}{n}$; площадь гистограммы относительных частот численно равна

$$\sum_{i=1}^{k} \frac{m_i}{n} = 1. \qquad (3.6)$$

Гистограмма относительных частот является статистическим аналогом плотности распределения наблюдаемой случайной величины ξ.

Гистограмма изображена на рис. 3.4.

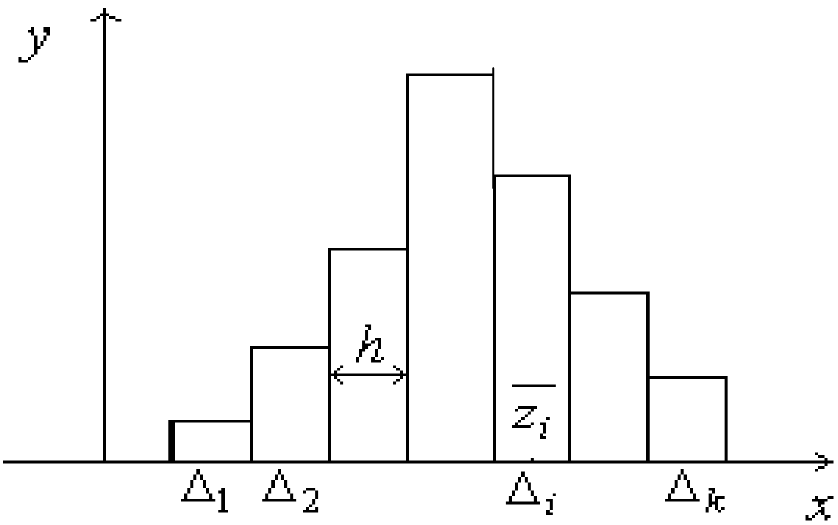


Рис. 3.4

3.3. Точечные оценки параметров распределения

На практике часто удается предсказать или оценить с помощью гистограммы вид распределения наблюдаемой случайной величины ξ с точностью до неизвестного параметра a (или нескольких параметров). Одной из основных задач математической статистики является нахождение оценки (приближенного значения) неизвестного параметра по имеющейся выборке.

Основные понятия

Пусть наблюдается случайная величина ξ с функцией распределения $F(x)$ и плотностью распределения $f(x)$. Случайная выборка представлена вектором $\vec{X} = (X_1, \ldots, X_n)$ с реализацией $\vec{x} = (x_1, \ldots, x_n)$. (3.7)

Параметром распределения a случайной величины ξ называется любая числовая характеристика этой случайной величины (математическое ожидание, дисперсия и т.п.) или любая константа, явно входящая в выражение для функции или плотности распределения.

Если параметр a неизвестен, то его *точечной оценкой* называется произвольная функция элементов выборки

$$a_n{}^* = g(X_1, X_2, \ldots, X_n) = g(\vec{X}). \qquad (3.8)$$

Реализацию оценки, т.е. значение оценки для наблюдавшейся в эксперименте реализации выборки, принимают за приближенное значение неизвестного параметра $a \approx a_n^*(x_1, x_2, \ldots, x_n)$.

Из соотношения (3.8) видно, что a_n * как функция случайных величин сама также является случайной величиной. Закон распределения оценки a_n^* зависит от вида функции g, числа наблюдений и значения оцениваемого параметра.

Ясно, что существует много разных способов построения точечной оценки, и не всякая зависимость $g(\vec{X})$ может давать удовлетворительную оценку неизвестного параметра a. Рассмотрим некоторые свойства, которыми должна обладать оценка, чтобы ее можно было считать хорошим приближением к неизвестному параметру.

Оценка a_n^* параметра a называется ***несмещенной,*** если ее математическое ожидание равно оцениваемому параметру, то есть

$$M\left\lfloor a_n^* \right\rfloor = a.$$

(3.9)

Если свойство (2.2) не выполняется, то есть

$$M\left\lfloor a_n^* \right\rfloor - a = b \neq 0, \qquad (3.10)$$

то оценку a_n^* называют ***смещенной***, при этом величину b называют систематической ошибкой оценки a_n^*.

Требование несмещенности означает, что выборочные значения оценок, полученных в результате повторения выборок, группируются около оцениваемого параметра.

Оценка a_n^* параметра a называется ***состоятельной***, если при $n \to \infty$ она сходится по вероятности к оцениваемому параметру a, т.е. для любого $\varepsilon > 0$ выполняется равенство

$$\lim_{n\to\infty} P\{|a_n * - a| < \varepsilon\} = 1 . \qquad (3.11)$$

Следующая теорема устанавливает достаточные условия состоятельности оценки a_n^* параметра a.

Теорема. Если при $n \to \infty$ $M\left[a_n^*\right] \to a$ и $D\left[a_n^*\right] \to 0$, то оценка a_n^* параметра a является состоятельной.

Состоятельность оценки означает, что, при достаточно большом объеме выборки с вероятностью близкой к единице, отклонение оценки от истинного значения параметра меньше ранее заданной величины.

Обычно в качестве *меры точности* оценки a_n^* используется среднеквадратическая ошибка (среднее значение квадрата ошибки) $M\left[\left(a_n^* - a\right)^2\right]$. Очевидно, чем меньше эта ошибка, тем теснее сгруппированы значения оценки около оцениваемого параметра. Поэтому всегда желательно, чтобы ошибка оценки была по возможности малой. Используя свойства математического ожидания, нетрудно получить

$$M\left[\left(a_n^* - a\right)^2\right] = D\left[a_n^*\right] + b^2. \qquad (3.12)$$

Для несмещенных оценок

$$M\left[\left(a_n^* - a\right)^2\right] = D\left[a_n^*\right] , \qquad (3.13)$$

то есть их мерой точности является дисперсия.

Несмещенная оценка параметра a называется его *эффективной* оценкой, если ее дисперсия $D\left[a_n^*\right]$ является наименьшей среди дисперсий всех возможных оценок параметра a, вычисленных по одному и тому же объему выборки.

Точечные оценки математического ожидания и дисперсии

Пусть случайная выборка $\vec{X} = \left(X_1, \ldots, X_n\right)$ порождена наблюдаемой случайной величиной ξ, математическое ожидание $M[\xi] = m$ и дисперсия

$D[\xi] = D = \sigma^2$ которой неизвестны. В качестве оценок для этих характеристик было предложено использовать выборочное среднее

$$m_n^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

и выборочную дисперсию

$$D_n^* = \sigma_n^{*2} = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2. \qquad (3.14)$$

Рассмотрим некоторые свойства оценок математического ожидания и дисперсии.

1. Вычислим математическое ожидание выборочного среднего:

$$M\left[\overline{X}\right] = M\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} M\left[X_i\right] = \frac{1}{n}\sum_{i=1}^{n} m = m. \qquad (3.15)$$

Следовательно, выборочное среднее является несмещенной оценкой для m.

2. Напомним, что результаты $X_1, X_2, \ldots, X_n$ наблюдений – независимые случайные величины, каждая из которых имеет такой же закон распределения, как и величина ξ, а значит, $M\left[X_i\right] = m$, $D\left[X_i\right] = \sigma^2$, $i = \overline{1,2\ ,...n}$. Будем предполагать, что дисперсия $D[\xi]$ конечна. Тогда, согласно теореме Чебышева о законе больших чисел, для любого $\varepsilon > 0$ имеет место равенство

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M\left[X_i\right]\right| < \varepsilon\right\} = 1,$$

которое можно записать так: $$\lim_{n\to\infty} P\left\{\left|\overline{X} - m\right| < \varepsilon\right\} = 1. \qquad (3.16)$$

Сравнивая (3.16) с определением свойства состоятельности (3.11), видим, что оценка $\overline{X}$ является состоятельной оценкой математического ожидания $M[\xi]$.

3. Найдем дисперсию выборочного среднего:

$$D\left[\overline{X}\right] = D\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} D\left[X_i\right] =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}. \qquad (3.17)$$

Таким образом, дисперсия оценки математического ожидания уменьшается обратно пропорционально объему выборки.

Можно доказать, что если случайная величина ξ распределена нормально, то выборочное среднее $\overline{X}$ является эффективной оценкой математического ожидания $M[\xi]$, то есть дисперсия $D\left[\overline{X}\right]$ принимает наименьшее значение по сравнению с любой другой оценкой математического ожидания. Для других законов распределения ξ это может быть и не так.

Выборочная дисперсия σ_n^{*2} является смещенной оценкой дисперсии σ^2, так как

$$M\left[\sigma_n^{*2}\right] = \frac{n-1}{n}\sigma^2 \neq \sigma^2. \tag{3.18}$$

Действительно, используя свойства математического ожидания и формулу (3.17), найдем

$$M\left[\sigma_n^{*2}\right] = M\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right] = \frac{1}{n}M\left[\sum_{i=1}^{n}\left(\left(X_i - m\right) - \left(\overline{X} - m\right)\right)^2\right] =$$

$$= \frac{1}{n}M\left[\sum_{i=1}^{n}\left(X_i - m\right)^2 - 2\left(\overline{X} - m\right)\sum_{i=1}^{n}\left(X_i - m\right) + \sum_{i=1}^{n}\left(\overline{X} - m\right)^2\right] =$$

$$= \frac{1}{n}M\left[\sum_{i=1}^{n}\left(X_i - m\right)^2 - 2\left(\overline{X} - m\right)\left(n\overline{X} - nm\right) + n\left(\overline{X} - m\right)^2\right] =$$

$$= \frac{1}{n}M\left[\sum_{i=1}^{n}\left(X_i - m\right)^2 - n\left(\overline{X} - m\right)^2\right] =$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n}M\left[\left(X_i - m\right)^2\right] - nM\left[\left(\overline{X} - m\right)^2\right]\right) =$$

$$= \frac{1}{n}\left(n\sigma^2 - n\cdot\frac{\sigma^2}{n}\right) = \frac{n-1}{n}\sigma^2.$$

Чтобы получить несмещенную оценку дисперсии, оценку (3.14) нужно исправить, то есть домножить на $\dfrac{n}{n-1}$. Тогда получим несмещенную выборочную дисперсию

$$S^2 = \frac{n}{n-1}D_n^* = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2. \tag{3.19}$$

Отметим, что формулы (3.14) и (3.19) отличаются лишь знаменателем, и при больших значениях n выборочная и несмещенная дисперсии отличаются мало. Однако при малом объеме выборки $(n < 30)$ следует пользоваться соотношением (3.19).

Для оценки среднего квадратического отклонения случайной величины используют так называемое “исправленное” среднее квадратическое отклонение, которое равно квадратному корню из несмещенной дисперсии: $S = \sqrt{S^2}$.

3.4. Интервальные оценки

В статистике имеются два подхода к оцениванию неизвестных параметров распределений: точечный и интервальный. В соответствии с точечным оцениванием, которое рассмотрено в предыдущем разделе, указывается лишь точка, около которой находится оцениваемый параметр. Желательно, однако, знать, как далеко может отстоять в действительности этот параметр от возможных реализаций оценок в разных сериях наблюдений.

Ответ на этот вопрос – тоже приближенный – дает другой способ оценивания параметров – интервальный. В соответствии с этим способом оценивания находят интервал, который с вероятностью, близкой к единице, накрывает неизвестное числовое значение параметра.

Понятие интервальной оценки

Точечная оценка $a^* = a^*(X_1, \ldots, X_n)$ является случайной величиной и для возможных реализаций выборки принимает значения лишь приближенно равные истинному значению параметра a. Чем меньше разность $\left|a - a^*\right|$, тем точнее оценка. Таким образом, положительное число Δ, для которого $\left|a - a^*\right| \le \Delta$, характеризует точность оценки и называется *ошибкой оценки (или предельной ошибкой).*

Доверительной вероятностью *(или надежностью)* называется вероятность **β**, с которой осуществляется неравенство $\left|a - a^*\right| \le \Delta$, т. е.

$$P\left\{ \left|a - a^*\right| \le \Delta \right\} = \beta . \qquad (3.20)$$

Заменив неравенство $\left|a - a^*\right| \le \Delta$ равносильным ему двойным неравенством $-\Delta \le a - a^* \le \Delta$, или $a^* - \Delta \le a \le a^* + \Delta$, получим

$$P\left\{ a^* - \Delta \le a \le a^* + \Delta \right\} = \beta . \qquad (3.21)$$

Интервал $I_\beta = \left(a^* - \Delta, a^* + \Delta\right)$, накрывающий с вероятностью **β**, $0 < \beta < 1$, неизвестный параметр a, называется ***доверительным интервалом*** *(или интервальной оценкой),* соответствующим доверительной вероятности **β**.

Случайной величиной является не только оценка a^*, но и ошибка Δ: ее значение зависит от вероятности **β** и, как правило, от выборки. Поэтому доверительный интервал случаен и выражение (3.21) следует читать так: “Интервал I_β накроет параметр a с вероятностью **β** ”, а не так: “Параметр a попадет в интервал I_β с вероятностью **β** ”.

Смысл доверительного интервала состоит в том, что при многократном повторении выборки объема n в относительной доле случаев, равной **β**, доверительный интервал, соответствующий доверительной вероятности **β**, накрывает истинное значение оцениваемого параметра. Таким образом, доверительная вероятность **β** характеризует *надежность* доверительного оценивания: чем больше **β**, тем вероятнее, что реализация доверительного интервала содержит неизвестный параметр.

Следует, однако, иметь в виду, что с ростом доверительной вероятности **β** в среднем растет длина доверительного интервала, то есть уменьшается точность доверительного оценивания. Выбор доверительной вероятности определяется конкретными условиями; обычно используются значения **β**, равные 0,90; 0,95; 0,99.

Вероятность
$$\alpha = 1 - \beta \tag{3.22}$$

называется ***уровнем значимости*** и характеризует относительное число ошибочных заключений в общем числе заключений.

В формуле (3.21) границы доверительного интервала симметричны относительно точечной оценки. Однако не всегда удается построить интервал, обладающий таким свойством. Более общим является следующее определение.

Доверительным интервалом (или *интервальной оценкой*) параметра a с доверительной вероятностью **β,** $0< \beta <1$, называется интервал со случайными границами $a_1^* = a_1^*(X_1,\ldots,X_n)$, $a_2^* = a_2^*(X_1,\ldots,X_n)$, накрывающий с вероятностью **β** неизвестный параметр a, т. е.

$$P\{a_1^* \le a \le a_2^*\} = \beta. \tag{3.23}$$

Иногда вместо двусторонних доверительных интервалов рассматривают односторонние доверительные интервалы, полагая $a_1^* = -\infty$ или $a_2^* = \infty$.

Построение интервальных оценок

Доверительный интервал задается своими концами a_1^* и a_2^*. Однако найти функции $a_1^* = a_1^*(X_1,\ldots,X_n)$ и $a_2^* = a_2^*(X_1,\ldots,X_n)$ из условия (3.23) невозможно, поскольку закон распределения этих функций зависит от закона распределения ξ и, следовательно, зависит от неизвестного параметра a. Используют следующий прием, позволяющий в ряде случаев построить доверительный интервал. Подбирается такая функция $G\left(\vec{X},a\right)$, чтобы:

- ее закон распределения был известен и не зависел от неизвестного параметра a;
- функция $G\left(\vec{X},a\right)$ была непрерывной и строго монотонной по a.

Тогда для любого β можно выбрать два числа g_1 и g_2 так, чтобы выполнялось равенство

$$P\{g_1 \le G \le g_2\} = F(g_2) - F(g_1) = \int_{g_1}^{g_2} f_G(t)dt = \beta. \quad (3.24)$$

Отсюда находят g_1 и g_2 как квантили функции распределения $G\left(\vec{X},a\right)$. Границы искомого доверительного интервала выражают через найденные квантили и выборочные данные, используя для этого соотношения, связывающие новую и старую случайные величины.

Если плотность распределения случайной величины $G\left(\vec{X},a\right)$ симметрична, то доверительный интервал симметричен относительно точечной оценки a^*, и для нахождения границ доверительного интервала вместо условия (3.23) можно использовать соотношение (3.21).

Основные статистические распределения

Построение разного рода оценок и статистических критериев часто основывается на использовании ряда специальных распределений случайных величин.

Нормальное распределение. Случайная величина X имеет нормальное распределение с параметрами m и $\sigma^2 > 0$, что обозначается как $X \sim N\left(m;\sigma^2\right)$, если плотность вероятности этой случайной величины имеет вид

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}. \quad (3\ .25)$$

График плотности вероятности случайной величины, имеющей нормальное распределение, представлен на рисунке 3.5, на котором видно, что максимум функции находится в точке $x = m$.

Поскольку нормальное распределение подробно изучается в курсе теории вероятностей, напомним свойства нормальной случайной величины, которые будут использоваться в дальнейшем.

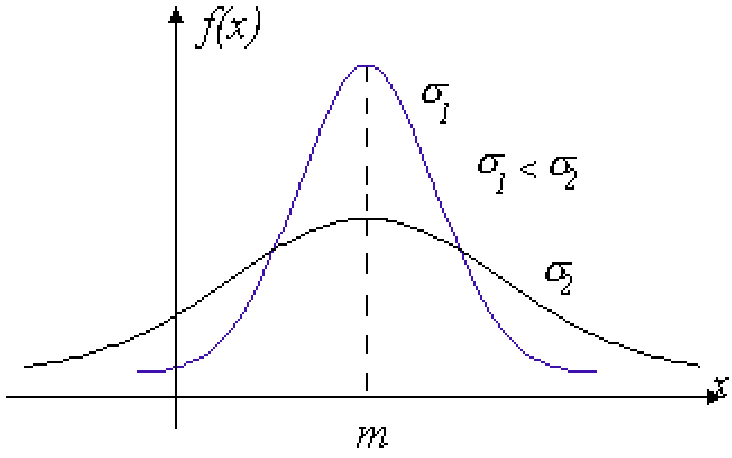


Рис. 3.5

1) $M[X]=m$, $D[X]=\sigma^2$.

2) Случайная величина называется *центрированной*, если ее математическое ожидание равно нулю. Для того чтобы центрировать случайную величину, надо вычесть из нее математическое ожидание:

$$M[X-M[X]]=M[X]-M[X]=0.$$

3) Случайная величина называется *нормированной*, если ее дисперсия равна единице, а математическое ожидание равно нулю.

Для того чтобы нормировать случайную величину, надо ее поделить на среднее квадратическое отклонение:

$$D\left[\frac{X}{\sigma}\right]=\frac{1}{\sigma^2}D[X]=\frac{\sigma^2}{\sigma^2}=1.$$

Центрированная и нормированная нормальная случайная величина называется стандартной. Таким образом, стандартной будет случайная величина

$$Z=\frac{X-m}{\sigma}\sim N(0;\ 1). \qquad (3.26)$$

Вероятность попадания случайной величины X в интервал (α,β) вычисляется по формуле

$$P\{\alpha < X \le \beta\}=\Phi\left(\frac{\beta-m}{\sigma}\right)-\Phi\left(\frac{\alpha-m}{\sigma}\right), \qquad (3.27)$$

где $\Phi(z)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{z}e^{\frac{-t^2}{2}}dt$ - интеграл вероятности, представляющий собой функцию распределения стандартной нормально распределенной случайной величины. Интеграл вероятности табулирован. Его значения приведены в таблице В Приложения.

Для стандартной нормальной случайной величины и симметричного промежутка формула (3.27) принимает следующий вид:

$$P\{-\alpha < X \le \alpha\} = P\{|X| < \alpha\} = 2\Phi(\alpha) - 1. \qquad (3.28)$$

Распределение χ^2 ***(хи-квадрат).*** Если $X_i \sim N(0;\ 1)$, $i=\overline{1,k}$ независимые стандартные нормальные случайные величины, то говорят, что случайная величина

$$Z_k = \sum_{i=1}^{k} X_i^{\,2} \qquad (3.29)$$

имеет распределение хи-квадрат с k степенями свободы, что обозначается как $Z_k \sim \chi^2(k)$. Графики плотности вероятности для двух значений степени свободы приведены на рис.3.6.

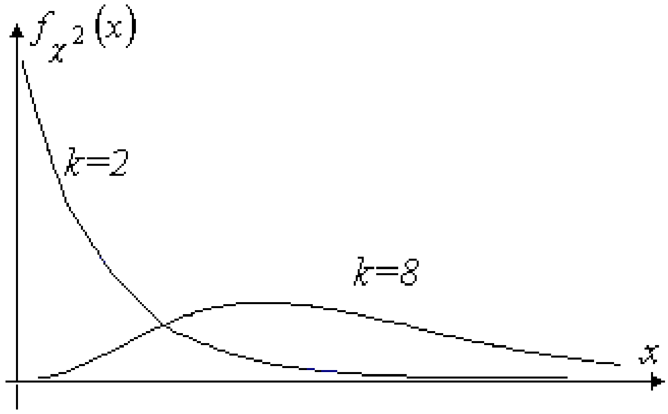


Рис. 3.6

С увеличением числа степеней свободы k плотность вероятности стремится к нормальной. При $k \le 2$ плотность вероятности постоянно убывает, а при $k > 2$ имеет единственный максимум $x = k-2$, $M[Z_k]=k$, $D[Z_k]=2k$.

Распределение Стьюдента. Пусть X_0, X_1, …, X_k - независимые стандартные нормальные случайные величины. Тогда случайная величина

$$T_k = \frac{X_0}{\sqrt{\frac{1}{k}\sum\limits_{i=1}^{k} X_i^{\,2}}} \qquad (3.30)$$

имеет распределение Стьюдента с k степенями свободы, что обозначается как $T_k \sim S(k)$, при этом

$$M[T_k] = 0, \quad D[T_k] = \frac{k}{k-2}.$$

На рис.3.7 приведены кривые стандартного нормального распределения (кривая 1) и плотности распределения Стьюдента (кривая 2).

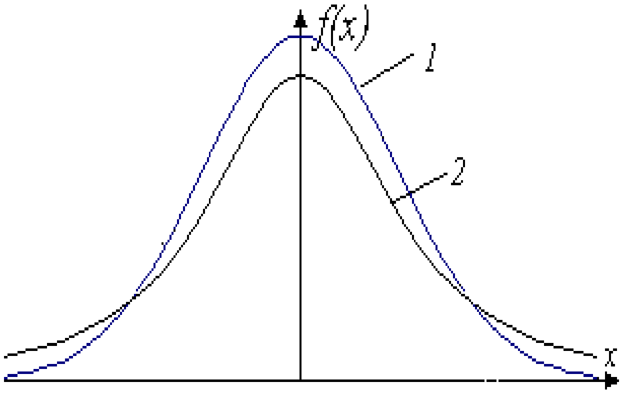


Рис. 3.7

При $n \to \infty$ плотность распределения Стьюдента стремится к плотности стандартной нормальной случайной величины.

На практике, как правило, используется не плотность вероятности, а *квантиль* распределения. Напомним, что *квантилью порядка (или уровня)* p непрерывной случайной величины X называется такое ее значение x_p, которое удовлетворяет равенству $F(x_p) = P\{X \le x_p\} = p$,

где $F(x)$ - функция распределения, а p - заданное значение вероятности. Рис.3.8 поясняет понятие квантили порядка p.

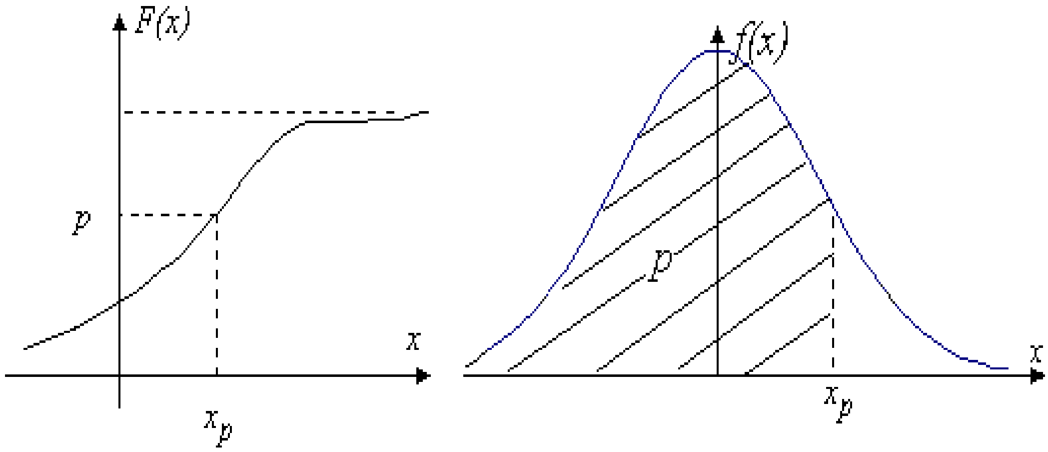


Рис. 3.8

Следующая теорема устанавливает свойства основных выборочных характеристик, вычисленных по выборке, соответствующих нормальному распределению.

Теорема Фишера. Пусть $\vec{X} = (X_1, \dots, X_n)$ - случайная выборка из генеральной совокупности $\xi \sim N\left(m;\ \sigma^2\right)$, тогда выборочное среднее $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ и несмещенная выборочная дисперсия $S^2 = \frac{1}{n-1}\sum\limits_{i=1}^{n}\left(X_i - \overline{X}\right)^2$ независимы, и при этом

1) случайная величина $\left(\overline{X} - m\right)\sqrt{n}/\sigma$ имеет распределение $N\left(0;\ 1\right)$;

2) случайная величина $(n-1)S^2/\sigma^2$ имеет распределение $\chi^2(n-1)$;

3) случайная величина $\left(\overline{X} - m\right)\sqrt{n}/S$ имеет распределение $S(n-1)$.

Доказательство теоремы приведено в [2].

Интервальные оценки математического ожидания нормального распределения

Интервальная оценка математического ожидания при известной дисперсии. Построим доверительный интервал для математического ожидания наблюдаемой случайной величины $\xi \sim N\left(m;\ \sigma^2\right)$ при известной дисперсии $D = \sigma^2$ по выборке $\vec{X} = (X_1, \dots, X_n)$.

Образуем вспомогательную случайную величину $Z = \dfrac{m^* - m}{\sigma / \sqrt{n}}$, где $m^* = \overline{X} = \dfrac{1}{n}\sum\limits_{i=1}^{n} X_i$ - точечная оценка математического ожидания m. Согласно утверждению 1 теоремы Фишера, случайная величина Z имеет нормальное распределение $N(0;\ 1)$ и ее функция распределения $\Phi(z) = P\{Z \le z\}$ не зависит от неизвестного параметра.

Доверительный интервал, соответствующий надежности β, определяется из условия (3.20), которое в нашем случае имеет вид

$$P\left\{ \left|m^* - m\right| \le \Delta \right\} = \beta. \tag{3.31}$$

Неравенства $\left|m^* - m\right| \le \Delta$ и $\left|\dfrac{m^* - m}{\sigma / \sqrt{n}}\right| \le \dfrac{\Delta}{\sigma / \sqrt{n}} \equiv z$ являются равносильными, то есть для любой выборки $\vec{x}$ они выполняются или не выполняются одновременно, поэтому соотношение (3.31) можно записать в виде

$$P\left\{ \left|\frac{m^* - m}{\sigma / \sqrt{n}}\right| \le z \right\} = \beta. \tag{3.32}$$

Поскольку случайная величина Z имеет стандартное нормальное распределение, вероятность в левой части формулы (3.32) можно выразить через нормальную стандартную функцию распределения по формуле (3.7):

$$P\left\{ \left|\frac{m^* - m}{\sigma / \sqrt{n}}\right| \le z \right\} = 2\Phi(z) - 1 \ . \tag{3.33}$$

Приравняв правую часть формулы (3.33) заданной доверительной вероятности β, получим уравнение $\Phi(z) = \dfrac{1+\beta}{2}$. Решение этого уравнения z_γ является квантилью порядка $\gamma = \dfrac{1+\beta}{2} = 1 - \dfrac{\alpha}{2}$ стандартного нормального распределения и определяется по таблице значений стандартной нормальной функции распределения (см. табл. В Приложения). Предельная ошибка Δ вычисляется по формуле $\Delta = \dfrac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}$. Таким образом, доверительным интервалом математического ожидания, соответствующим надежности β, является интервал

$$I_\beta = \left(m^* - \Delta,\ m^* + \Delta\right) = \left(m^* - \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}},\ m^* + \frac{\sigma}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right). \quad (3.34)$$

<u>Интервальная оценка математического ожидания при неизвестной дисперсии.</u>

По выборке $\vec{X} = (X_1, \ldots, X_n)$ из нормального распределения $N(m;\ \sigma^2)$ требуется построить доверительный интервал для неизвестного математического ожидания m при неизвестной дисперсии D=σ^2. Введем новую случайную величину $T = \dfrac{m^* - m}{S/\sqrt{n}}$, где S - несмещенная выборочная дисперсия.

Статистика T согласно утверждению 3 теоремы Фишера имеет распределение Стьюдента с $n-1$ степенями свободы. Рассуждая аналогично случаю, когда дисперсия известна, получим следующий доверительный интервал для математического ожидания:

$$I_\beta = \left(m^* - \Delta, m^* + \Delta\right) = \left(m^* - \frac{S}{\sqrt{n}} t_\gamma, m^* + \frac{S}{\sqrt{n}} t_\gamma\right), \quad (3.35)$$

где t_γ - квантиль порядка $\gamma = \dfrac{1+\beta}{2}$ распределения Стьюдента. В отличие от доверительного интервала (3.34) длина интервала (3.35) случайна и зависит от случайной величины S. Поскольку с увеличением числа степеней свободы распределение Стьюдента быстро приближается к нормальному, то для больших выборок $(n > 30)$ интервалы (3.34) и (3.35) практически совпадают.

Пример 3.2. По результатам 9 измерений напряжения батареи получено среднее арифметическое значение $\bar{x} = 30{,}6$В. Точность вольтметра характеризуется средним квадратическим отклонением 0,2В. Требуется найти доверительный интервал для истинного значения напряжения батареи, соответствующий доверительной вероятности β=0,95, предполагая, что контролируемый признак имеет нормальный закон распределения.

Решение. Для нахождения доверительного интервала воспользуемся формулой (3.34). Квантиль порядка $\dfrac{1+\beta}{2} = 1 - \dfrac{\alpha}{2} = 0{,}975$ найдем по таблице А приложения: $z_{1-\frac{\alpha}{2}} = z_{0,975} = 1{,}96$. Поскольку предельная ошибка $\Delta = 1{,}96\dfrac{0{,}2}{\sqrt{9}} \approx 0{,}1$, то доверительный интервал имеет вид

$$I_{0,95} = (30,6-0,1,\ 30,6+0,1) = (30,5,\ 30,7)\ .$$

Интервальная оценка дисперсии нормального распределения

Построим доверительный интервал для дисперсии D=σ^2 наблюдаемой случайной величины $\xi \sim N(m;\ \sigma^2)$ по случайной выборке $\vec{X} = (X_1, \ldots, X_n)$ при неизвестном математическом ожидании.

Введем случайную величину (статистику)

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}, \qquad (3.36)$$

которая согласно утверждению 2 теоремы Фишера имеет распределение χ^2 с $n-1$ степенями свободы. Поскольку плотность распределения этого закона асимметрична, доверительный интервал, соответствующий надежности β, найдем из формулы (3.31) в виде:

$$P\{c_1 \le \chi^2 \le c_2\} = \beta. \qquad (3.37)$$

Обычно доверительный интервал I_β для случайной величины χ^2 выбирают так, чтобы вероятность ее попадания за пределы этого интервала влево и вправо была одинаковой (рис. 3.9):

$$P\{\chi^2 < c_1\} = P\{\chi^2 > c_2\} = \frac{1-\beta}{2}.$$

Тогда условия для определения значений c_1 и c_2 будут иметь вид:

$$P\{\chi^2 < c_1\} = \frac{1-\beta}{2}, \quad P\{\chi^2 < c_2\} = 1 - \frac{1-\beta}{2} = \frac{1+\beta}{2}. \qquad (3.38)$$

По таблице квантилей χ^2 - распределения (табл. С приложения) найдем

$$c_1 = \chi^2_{\frac{1-\beta}{2}}, \qquad c_2 = \chi^2_{\frac{1+\beta}{2}}. \qquad (3.39)$$

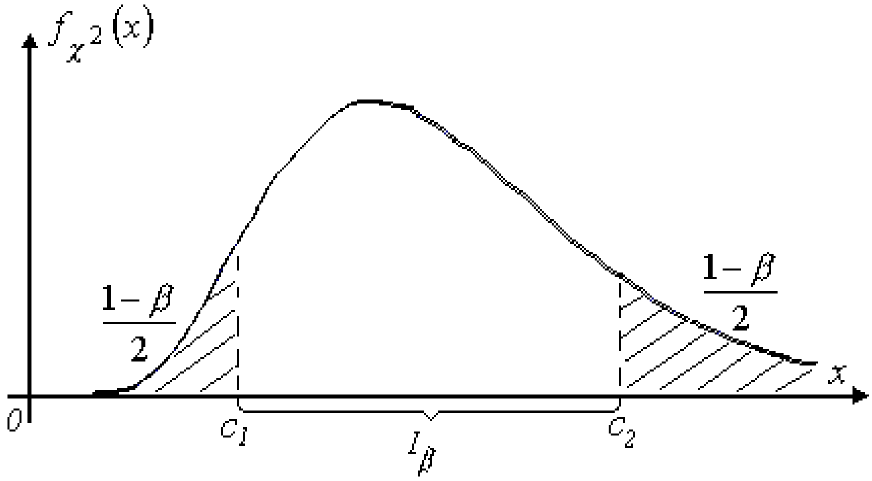


Рис. 3.9

Неравенства $c_1 \le \dfrac{(n-1)S^2}{\sigma^2} \le c_2$ эквивалентны неравенствам $\dfrac{(n-1)S^2}{c_2} \le \sigma^2 \le \dfrac{(n-1)S^2}{c_1}$, поэтому

$$P\left\{\frac{(n-1)S^2}{\chi^2_{\frac{1+\beta}{2}}} < \chi^2 < \frac{(n-1)S^2}{\chi^2_{\frac{1-\beta}{2}}}\right\} = \beta.$$

Следовательно, интервал

$$I_\beta = \left(\frac{(n-1)S^2}{\chi^2_{\frac{1+\beta}{2}}},\ \frac{(n-1)S^2}{\chi^2_{\frac{1-\beta}{2}}}\right) \qquad (3.40)$$

является доверительным интервалом дисперсии, соответствующим доверительной вероятности β.

Пример 3.3. По данным выборочного контроля найти выборочное математическое ожидание и несмещенную оценку дисперсии нормальной случайной величины ξ. Найти доверительные интервалы для них, соответствующие доверительной вероятности $\beta = 0{,}98$.

Таблица 3.4

x_i	42	43	45	46	48	51	52	54
m_i	1	2	3	6	4	3	1	1

Решение. Выборочное математическое ожидание найдем по формуле (3.14), используя табл.3.4

При $n = 21$ $\qquad m^* = \dfrac{1}{21}\sum_{i=1}^{8} x_i m_i \approx 47{,}1.$

Несмещенную выборочную дисперсию вычислим по формуле (3.19):

$$S^2 = \frac{1}{20}\sum_{i=1}^{8}(x_i - 47{,}1)\cdot m_i \approx 14{,}3, \quad S = 3{,}78.$$

Доверительный интервал для математического ожидания определим по формуле (3.35). При $k = n - 1 = 20$ из таблицы А приложения находим квантиль распределения Стьюдента $t_{\frac{1+\beta}{2}} = t_{0,99} = 2{,}53$. Вычислив предельную ошибку

$$\Delta = 2{,}53 \cdot \frac{3{,}78}{\sqrt{21}} \approx 2{,}09,$$

получим искомый доверительный интервал для математического ожидания:

$$I_\beta = (47{,}1 - 2{,}09,\ 47{,}1 + 2{,}09) = (45{,}0,\ 49{,}2).$$

Границы доверительного интервала для дисперсии определим по формуле (3.20). По таблице квантилей распределения χ^2 (см. табл. С приложения) при $k = n - 1 = 20$ определим квантили:

$$\chi^2_{\frac{1+\beta}{2}} = \chi^2_{0,99} = 37{,}6, \qquad \chi^2_{\frac{1-\beta}{2}} = \chi^2_{0,01} = 8{,}3.$$

Подставив эти значения, а также S и n в формулу (3.20), получим искомый доверительный интервал для дисперсии

$$I_{0,98} = \left(\frac{20\cdot 14{,}3}{37{,}6},\ \frac{20\cdot 14{,}3}{8{,}3}\right) = (7{,}60,\ 34{,}5).$$

Вопросы для самопроверки

1.Что называется выборкой?

2.Как произвести оценку выборочного математического ожидания и выборочной дисперсии?

3.Как найти функцию распределения для дискретной случайной величины?

4.Что такое несмещенная оценка параметра?

5.Дайте определение состоятельной оценки.

6.Что такое интервальная оценка?

Заключение

В результате изучения выше приведенного материала студент может приступить к выполнению контрольной работы и проверить свои ответы на вопросы самоконтроля. Затем после выполнения лабораторных работ может приступить к ответам на вопросы экзаменационного теста и получить оценку за проделанную работу.

Глоссарий

Биномиальное распределение с параметрами n и p – вычисление вероятности того, что случайная величина принимает значения m=0, 1,…, n.

Вариационный ряд – последовательность элементов выборки, расположенных в неубывающем порядке (одинаковые элементы записываются последовательно друг за другом).

Вероятность произведения двух независимых событий – произведение вероятностей этих событий.

Вероятность события -- отношение числа исходов m события A к общему числу элементарных событий N.

Возможные значения случайной величины – числа $f(\omega)$.

Выборка – последовательность значений из генеральной совокупности;

- **объема *k* -** часть, состоящая из k элементов генеральной совокупности;
- **репрезентативная –** позволяет адекватно описать случайную величину
- **случайная объема *n* –** последовательность n независимых случайных величин из генеральной совокупности.

Выборочная дисперсия – величина, равная сумме квадратов разностей между значением случайной величины и ее математическим ожиданием, деленная на объем выборки.

Выборочное среднее – число, равное сумме значений случайной величины, деленной на объем выборки.

Генеральная совокупность – конечная или бесконечная совокупность наблюдений над случайной величиной.

Геометрическое определение вероятности – отношение площади $S(A)$, соответствующей событию A, к площади всей области Ω.

Гипергеометрическое распределение – вычисление вероятности того, что случайная величина примет заданное значение через число сочетаний.

Гипотеза альтернативная – гипотеза, конкурирующая с основной;

-основная – гипотеза, которая проверяется;

-статистическая – предположение относительно параметров или закона распределения случайной величины.

Гистограмма – представление статистического ряда на плоскости.

Дискретная случайная величина - множество возможных значений образует конечную или бесконечную последовательность чисел, т.е. конечно или счетно.

Дисперсия случайной величины ξ – момент второго порядка случайной величины $(\xi - M(\xi))$.

Доверительная вероятность – вероятность, с которой производится оценка параметров.

Доверительный интервал – область значений, при которых основная гипотеза принимается.

Дополнение множества *A* – разность между всем множеством S и множеством A, которое является частью S.

Достоверное событие Ω – всегда наступает в условиях данного эксперимента.

Закон трех сигм – значения случайной величины ξ, имеющей нормальное распределение с параметрами m и σ, содержатся в интервале $[\, m\text{-}3\sigma,\ m\text{-}3\sigma]$.

Кривая распределения – график плотности вероятности.

Критерий значимости – вероятность ошибки 1-го рода.

Критерий

- согласия – правило, в соответствии с которым принимается решение;

- Колмогорова – проверка гипотезы о совпадении функций распределения.

Математическое ожидание дискретной случайной величины ξ – сумма ряда из произведений возможных значений x_i на их вероятности p_i.

Множество – некоторая совокупность объектов, называемых элементами множества.

Множество конечное – состоящие из конечного числа элементов, в противном случае – **бесконечное множество.**

Момент второго порядка случайной величины ξ – математическое ожидание квадрата этой случайной величины.

Моргана формулы или соотношения двойственности – правило для записи выражения, соответствующего «отрицанию» функции.

Невозможное событие – это такое, которое не может наступить в условиях данного эксперимента, т.е. это событие имеет пустое множество благоприятствующих исходов.
Независимые *с*обытия A и B – событие A происходит независимо от того, происходит событие В или нет.
Несовместные события A и B – не могут происходить одновременно.
Нормальное или гауссовское распределение – случайная величина ξ имеет плотность распределения вероятностей при всех x

$$\varphi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad .$$

- хи-квадрат (Пирсона) – проверка гипотезы о совпадении дисперсий.
Относительная частота события ***А*** – показывает долю опытов, в которых наступило событие A при N экспериментах.
Оценка интервальная – доверительный интервал:
- несмещенная – математическое ожидание случайной величины в этом случае равно оцениваемому параметру;
- точечная – произвольная функция элементов выборки, когда параметр неизвестен.
Ошибка второго рода – событие, состоящее в том, что гипотеза принимается, когда на самом деле она неверна.
Ошибка первого рода – событие, состоящее в том, что гипотеза отвергается, когда на самом деле она верна.

Показательное распределение ***с параметром*** $\lambda > 0$ – это такое распределение, плотность вероятности которого задается равенством $\varphi(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{если } x \geq 0, \\ 0, & \text{если } x < 0. \end{cases}$.

Произведение или пересечение множеств A **и** B – множество, состоящее из всех общих элементов этих множеств.
Пространство элементарных событий – множество всех исходов данного эксперимента.
Противоположное событие – это событие, которое происходит в том случае, если не происходит событие A.
Пустое множество – множество, не содержащее элементов**.**
Равномерное распределение - случайная величина ξ на промежутке **[*a*,*b*]** имеет постоянную плотность распределения вероятностей.
Размещение из *n* элементов по *k* элементов – упорядоченные выборки объема k без возвращения элементов.
Разность множеств А и *В* – множество, состоящее из всех элементов множества A, которые не содержатся в множестве B.
Ряд распределения – статистический ряд, записанный в виде таблицы.

Случайная величина – функция f, которая каждому элементарному событию ω ставит в соответствие число $f(\omega)$**.**

Событие – некоторое высказывание о результатах рассматриваемого эксперимента.

Сочетание из ***n*** элементов по ***k*** элементов - неупорядоченные выборки объема k без возвращения элементов.

Стандартное или средне-квадратическое отклонение - квадратный корень из дисперсии.

Статистика – результат наблюдения над случайной величиной.

Статистический ряд – последовательность различных значений, расположенных в возрастающем порядке, с указанием относительных частот.

Сумма или объединение множеств *А* и *В* – множество, состоящее из всех элементов, принадлежащих хотя бы одному из этих множеств.

Уровень значимости статистического критерия – величина, определяющая степень достоверности вычислений.

Условие нормировки – площадь криволинейной трапеции под всей кривой распределения равна 1.

Условная вероятность – вероятность события A при условии, что событие B произошло.

Функция Лапласа - функция распределения стандартного нормального закона.

Функция распределения *F(x)* случайной величины ξ - вероятность того, что случайная величина примет значение, меньшее заданного x**.**

Частный случай – если при каждом осуществлении события A происходит и событие B**,** то говорят, что событие А влечет событие B.

Частота события *А* – число экспериментов $m_n(A)$, в которых наступило событие A.

Элементарные события – исходы (результаты) эксперимента.

Эмпирическая функция распределения – относительная частота события, заключающегося в том, что случайная величина примет значение, меньшее чем заданное число.

4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ВЫПОЛНЕНИЮ ЛАБОРАТОРНЫХ РАБОТ

К выполнению лабораторных работ допускаются студенты, изучившие основные теоретические положения. Каждый студент должен выполнить две лабораторные работы, задания для которых сформулированы в описании.

Организация безопасной работы студентов при выполнении лабораторных работ на кафедре информатики и прикладной математики производится в соответствии с требованиями ГОСТ 12.1.030-81 «Электробезопасность. Защитное заземление, зануление», а также правил использования персональных компьютеров.

Перед выполнением лабораторных работ на ПК все студенты проходят нструктаж по технике безопасности, о чем производится запись в соответствующем журнале, которая подтверждается собственноручными подписями студентов и лицом, проводящим инструктаж.

При обнаружении неисправностей в ПК во время выполнения лабораторной работы следует немедленно прекратить работу, отключить ПК и сообщить преподавателю.

В случае завершения работы надо отключить напряжение электропитания и привести в порядок рабочее место.

Запрещается:

- находиться в помещении в верхней одежде;
- оставлять без надзора лабораторную технику;
- выполнять работу в отсутствии преподавателя или дежурного лаборанта;
- складывать сумки, одежду и другие вещи на рабочие столы и на лабораторную технику.

Студенты, нарушающие правила техники безопасности, отстраняются от выполнения лабораторных работ.

Библиографический список

1. Вентцель, Е.С. Теория вероятностей и ее инженерные приложения / Е.С. Вентцель, Л.А. Овчаров. - М.: Наука, 1988.
2. Гмурман, В.Е. Теория вероятностей и математическая статистика / В.Е. Гмурман. - М.: Высш. шк., 2004.
3. Шабаева, М.Б. Вычислительная математика. Элементы математической статистики / М.Б. Шабаева. – СПб.: Изд-во СЗТУ, 2004.

Лабораторная работа 1

ОПИСАНИЕ СЛУЧАЙНЫХ ВЕЛИЧИН

ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ

Основные теоретические положения изложены в разделе 3 опорного конспекта данного пособия, а также в [2] , с.24-32, [3], гл.16, [6], с.126-138.

Целью работы является изучение методики статистического оценивания параметров закона распределения, освоение инструментов статистического анализа в в MS Excel для студентов всех форм обучения и по ДОТ в том числе.

Задача ставится следующим образом: случайная величина имеет закон распределения определенного вида, зависящий от параметра, значение которого неизвестно. Требуется на основании опытных данных оценить значение этого параметра.

Лабораторная работа состоит из двух частей. Первая часть включает в себя выполнение заданий 1 и 2 и представляет собой контрольный пример, решение которого приведено ниже. Во второй части самостоятельно выполняется индивидуальная работа, состоящая из заданий 3 и 4.

Порядок выполнения лабораторной работы

Задание 1

1.1. Получить с использованием *Пакета анализа* в MS Excel выборку объема $n = 40$ из генеральной совокупности, в которой случайная величина ξ распределена по нормальному закону с параметрами $m = 4$ и $\sigma = 0{,}5$.

1.2. Найти точечные оценки математического ожидания, дисперсии, стандартного отклонения, а также стандартную ошибку оценки математического ожидания:

а) по данным малой выборки $(n = 20)$;

б) данным большой выборки $(n = 40)$.

Задание 2

2.1. Найти доверительные интервалы, соответствующие доверительной вероятности β=0,95, для оценок математического ожидания, полученных в задании 1.2, без использования MS Excel.

2.2. Выполнить задание 2.1 с использованием MS Excel.

2.3.Построить с использованием инструмента **Описательная Статистика** *Пакета анализа* статистический отчет для выборки малого объема.

Задание 3

3.1. При помощи *Пакета анализа* смоделировать выборку случайной величины ξ, распределенной по нормальному закону с параметрами m и σ. Значения параметров и объема выборки n следует выбрать в соответствии с последней цифрой шифра из табл. 4.1 (случайное рассеивание взять равным предпоследней цифре шифра).

Таблица 4.1

Последняя цифра шифра	0	1	2	3	4	5	6	7	8	9
m, Σ	10,4	10,5	20,5	20,4	30,5	30,6	50,6	60,7	70,7	80 7
n	42	44	46	48	50	49	47	45	43	41

3.2. Выполнить с использованием MS Excel точечное и интервальное оценивание математического ожидания генеральной совокупности по данным выборки, полученной в задании 3.1. Значение доверительной вероятности взять равным 0,90.

Задание 4 . По результатам выполнения заданий 1 и 2 сформулировать и обосновать выводы о том, как изменяется точность найденных параметров в зависимости от объема выборки.

Выполнение задания 1.1. Инструмент **Генерация случайных чисел** *Пакета анализа* предназначен для проведения статистического моделирования. Для моделирования выборки заданного закона распределения выполните следующие действия:

1. Выберите команду **Сервис – Анализ данных**. Откроется окно диалога “Анализ данных” (рис. 4.1).
2. Из списка инструментов анализа выберите **Генерация случайных чисел** и нажмите кнопку **OK**. На экране появится окно диалога “Генерация случайных чисел” (рис. 4.2).
3. Введите в поле **Число переменных** - 1 (поскольку требуется смоделировать один закон распределения), а в поле **Число случайных чисел** - 40 (объем выборки).
4. Закон распределения моделируемой случайной величины задается параметром **Распределение**: выберите из списка **Нормальное**. Откроется окно диалога для ввода параметров нормального распределения.
5. Введите в поле **Среднее** число 4, в поле ввода **Стандартное отклонение -** 0,5, а в поле **Случайное рассеивание -** число 4.
6. В разделе “Параметры вывода” установите переключатель **Выходной интервал**. В одноименном поле укажите адрес ячейки $A$6, начиная с которой будут выводиться на экран выборочные значения.
7. Нажмите кнопку **OK**. В столбце А, начиная с ячейки А6, появятся генерируемые значения выборки.

8. Значения ячеек A26:A45 перенесите в ячейки B6:B25 для удобства дальнейшей обработки.

Создаваемая электронная таблица представлена в табл. 4.2 в режиме формул и в табл. 4.3 в режиме вычислений.

Завершая выполнение первого задания, оформите таблицу так, как указано:

Ячейка	**Значение**
B1	ТОЧЕЧНОЕ И ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ
A2	выборка из нормального распределения с параметрами

Ячейка	**Значение**	**Ячейка**	**Значение**
F2	мат. ожид.	G2	4
F3	станд .откл .	G3	0,5
F4	дов. вер.	G4	0,95
F5	объем выборки	G5	40

Назовите ярлык рабочего листа **Оценки**.

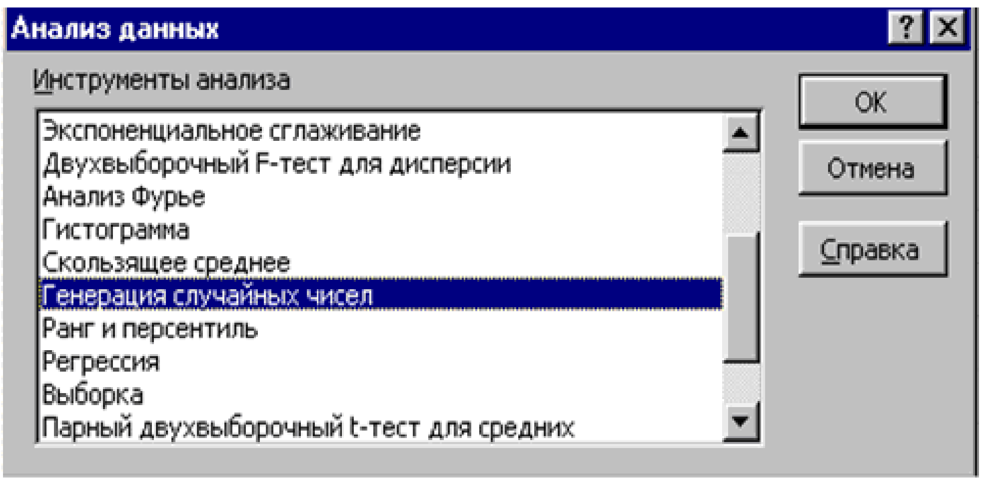


Рис. 4.1. Окно диалога «Анализ данных»

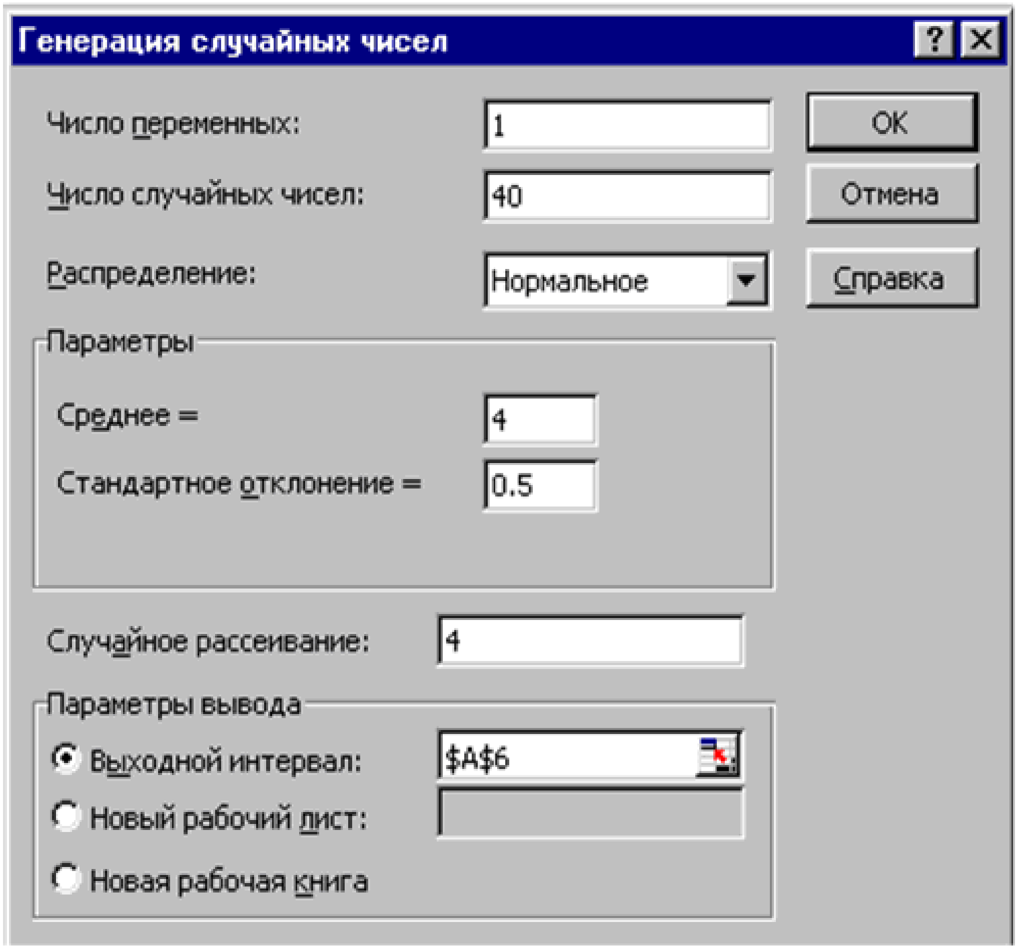


Рис. 4.2. Окно диалога «Генерация случайных чисел»

В табл. 4.2 приводится текст программы в Excel (режим показа формул) выполнения задания 1.1; в табл. 4.3 – программа в режиме вычислений.

Таблица 4.2.

	ТОЧЕЧНОЕ И ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ							
выборка из нормального распределения с параметрами					мат ожид	4		
					станд откл	0,5		
					дов вер	0,95		
					объем выборки	40		
2,52	3,0		**ТОЧЕЧНЫЕ ОЦЕНКИ**					
4,05	4,5	объем	выб среднее	несм дисп	несм станд откл	выб дисп	выб ст	станд ошибка
4,47	3,9	20	=СРЗНАЧ(A6:A25)	=ДИСП(A6:A25)	=СТАНДОТКЛОН	=ДИСПР(A6:A25)	=СТА	=F8/КОРЕНЬ(C8)
2,86	4,1	40	=СРЗНАЧ(A6:B26)	=ДИСП(A6:B26)	=СТАНДОТКЛОН	=ДИСПР(A6:B26)	=СТА	=F9/КОРЕНЬ(C9)
4,03	3,8							
4,49	3,8		**ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ**					
4,17	3,9	объем	квантиль	пред ошибка	ниж гран	верх гран		
4,16	3,2	20	=СТЬЮДРАСПОБР(1-$G$4;C13-	=D13*F8/КОРЕНЬ(C13)	=D8-E13	=E8+F13		
3,78	4,0	40	=НОРМСТОБР((1+$G$4)/2)	=D14*F9/КОРЕНЬ(C14)	=D9-E14	=D9+E14		
3,54	3,9			=ДОВЕРИТ(1-$G$4;F9;C14)				
3,50	3,9							
4,54	4,1		*Столбец1*					
4,31	4,3							
3,04	4,2		Среднее	3,81148556571497				
3,78	4,2		Стандартная ошибка	0,124975720503504				
4,12	3,9		Медиана	3,91239235441753				
4,05	4,9		Мода	#Н/Д				
3,18	3,2		Стандартное отклонение	0,558908413165698				
3,75	4,7		Дисперсия выборки	0,312378614307399				
3,76	3,4		Эксцесс	0,0828435006460544				
			Асимметричность	-0,81089891377659				
			Интервал	2,02506839741545				
			Минимум	2,52181645110249				
			Максимум	4,54688484851795				
			Сумма	76,2297113142995				
			Счет	20				
			Уровень надежности(95,0%)	0,261577270496961				

Программа в режиме вычислений

Таблица 4.3

	ТОЧЕЧНОЕ И ИНТЕРВАЛЬНОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ							
Выборка из нормального распределения с параметрами					Мат. ожид.	4		
					Станд. откл.	0,5		
					Дов. вер.	0,95		
					Объем выборки	40		
2,521816	3,085		**ТОЧЕЧНЫЕ ОЦЕНКИ**					
4,059869	4,547	Объем выб	Выб. среднее	Несм. дисп	Несм. станд. откл.	Выб. дисп.	Выб. станд. откл.	Станд. Ошибка
4,47482	3,992	20	3,811485566	0,312378614	0,5589084	0,29676	0,54476	0,124975721
2,867888	4,188	40	3,904166415	0,274648532	0,5240692	0,26778	0,51748	0,082862617
4,03681	3,802							
4,495959	3,861		**ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ**					
4,172886	3,978	Объем выб	Квантиль	Пред. ошибка	Ниж. гран.	Верх. гран.		
4,163213	3,296	20	2,09302405	0,261577189	3,5499084	3,86229		
3,786216	4,021	40	1,959963985	0,162407745	3,7417587	4,06657		
3,549364	3,967			0,162407745				
3,503916	3,926							
4,546885	4,124		*Столбец1*					
4,317657	4,301							
3,047441	4,272		Среднее	3,811485566				
3,787975	4,234		Стандартная ошибка	0,124975721				
4,129709	3,983		Медиана	3,912392354				
4,058636	4,952		Мода	#Н/Д				
3,183965	3,201		Стандартное отклонение	0,558908413				

3,759025	4,774		Дисперсия выборки	0,312378614				
3,765662	3,433		Эксцесс	0,082843501				
			Асимметричность	-0,810898914				
			Интервал	2,025068397				
			Минимум	2,521816451				
			Максимум	4,546884849				

Выполнение задания 1.2. Для нахождения значений точечных оценок воспользуемся функциями, которые содержатся в категории **Статистические**. Функция СРЗНАЧ возвращает значение оценки математического ожидания, вычисленное по формуле

$$m_\xi^* = \frac{1}{n}\sum_{i=1}^{n} X_i . \qquad (4.1)$$

Выборочная дисперсия

$$D_\xi^* = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - m_\xi^*\right)^2$$

(4.2)

является смещенной оценкой дисперсии генеральной совокупности. Несмещенной и состоятельной оценкой дисперсии является величина

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - m_\xi^*\right)^2 . \qquad (4.3)$$

Функции ДИСПР и СТАНДОТКЛОНП – значения выборочной дисперсии и стандартного отклонения. Для вычисления значений несмещенной оценки дисперсии и оценки стандартного отклонения воспользуемся функциями ДИСП и СТАНДОТКЛОН соответственно.

Введите формулы для вычисления значений точечных оценок параметров закона распределения для малой выборки в строку 8 (ячейки D8:H8), для большой выборки – в строку 9 (ячейки D9:H9).

Ячейка	**Значение**
D8	=СРЗНАЧ(A6:A25)
D9	= СРЗНАЧ(A6:B25)
E8	=ДИСП(A6:A25)
E9	=ДИСП(A6:B25)
F8	=СТАНДОТКЛОН(A6:A25)
F9	= СТАНДОТКЛОН(A6:B25)
G8	=ДИСПР(A6:A25)
G9	=ДИСПР(A6:B25)
H8	=СТАНДОТКЛОН(A6:A25)
H9	= СТАНДОТКЛОН(A6:B25)

В ячейках C8 и C9 укажите объемы большой и малой выборок – числа 20 и 40.

В ячейках I8 и I9 наберите формулу для расчета стандартной ошибки

$$\mu_{m^*} = \sqrt{\frac{\sigma^2}{n}}\,.$$

Ячейка	Значение
I8	=F8/КОРЕНЬ(C8)
I9	=F9/КОРЕНЬ(C9)

Чтобы электронная таблица была удобной для анализа результатов ычислений, введите следующие поясняющие заголовки:

Ячейка	Значение	Ячейка	Значение
D6	ТОЧЕЧНЫЕ ОЦЕНКИ	C7	Объем выб.
D7	Выб. среднее	E7	Несм. дисп.
F7	Несм. станд. откл.	G7	Выб. дисп.
H7	Выб. станд. откл.	I7	Станд. ошибка

Выполнение задания 2.1. Задача о нахождении доверительного интервала для математического ожидания нормально распределенной случайной величины решена в *примере3.3*. В расчетах следует использовать вычисленные при выполнении задания 1.2 значения точечных оценок математического ожидания и дисперсии.

Выполнение задания 2.2

А. Доверительный интервал для математического ожидания в случае малой выборки определяется по формуле

$$I_\beta = \left(m^* - \varepsilon, m^* + \varepsilon\right) = \left(m^* - \frac{s}{\sqrt{n}} t_\beta, m^* + \frac{s}{\sqrt{n}} t_\beta\right).$$

Расчет выполняется в ячейках C13:G13.

А1. Введите значения объемов малой и большой выборок (числа 20 и 40) в ячейки C13 и C14 соответственно.

А2. Функция СТЬЮДРАСПОБР возвращает квантиль порядка $\frac{1+\beta}{2}$ распределения Стьюдента с $n-1$ степенью свободы t_β. В качестве аргументов функции следует указать уровень значимости $\alpha=1-\beta$ и число степеней свободы. Введите в ячейку D13 формулу

=СТЬЮДРАСПОБР(1-$G$4;C13-1).

А3. Для вычисления предельной ошибки $\varepsilon = \frac{s}{\sqrt{n}} t_\beta$ введите в ячейку E13 формулу =D13*F8/КОРЕНЬ(C13).

А4. Нижнюю и верхнюю границы доверительного интервала вычислите в ячейках F13 и G13.

Ячейка	**Значение**
F13	=D8-E13
G13	=D8+E13

В. Расчет границ доверительного интервала для математического ожидания большой выборки по формуле (3.3) выполните в ячейках D14:G14. Для нахождения квантили нормального стандартного распределения порядка **(1+β)/2** воспользуйтесь функцией НОРМСТОБР. Введите следующие формулы:

Ячейка	**Значение**
D14	= НОРМСТОБР((1+$G$4)/2)
E14	= D14*F9/КОРЕНЬ(C14)
F14	= D9-E14
G14	= D9+E14

Введите комментарий так, как указано ниже.

Ячейка	**Значение**

D11 ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ

Ячейка	**Значение**	**Ячейка**	**Значение**
C12	объем. выб.	D12	квантиль
E12	пред. ошибка	F12	ниж. гран.
G12	верх. гран.		

С. Функция ДОВЕРИТ возвращает предельную ошибку для заданной доверительной вероятности. Аргументами этой функции являются уровень значимости, несмещенное стандартное отклонение и объем выборки. Введите в ячейку E15 формулу

=ДОВЕРИТ(1-$G$4;F9;C14).

Убедитесь в том, что значения в ячейках E14 и E15 совпадают.

Выполнение задания 2.3. Инструмент **Описательная статистика** позволяет построить статистический отчет для входных данных (заданной выборки). Выходная таблица содержит два столбца: левый столбец содержит названия статистических данных, правый – статистические данные.

Для получения результатов обработки выборки в выходной таблице выполните следующие действия:

1. Выберите команду **Сервис–Анализ данных – Описательная статистика**.
2. В открывшемся окне диалога “Описательная статистика” задайте параметры. Введите входной интервал $A$6:$A$25. Для параметра **Группирование** установите переключатель **по столбцам.**
3. Установите флажки **Итоговая статистика, Уровень надежности,** в одноименное поле введите **95 %**.
4. В разделе **Параметры вывода** установите переключатель **Выходной интервал** и в поле **Выходной интервал** введите адрес ячейки $D$17.
5. Параметры окна диалога “Описательная статистика” установлены. Нажмите кнопку **OK**. На экране появится таблица, созданная инструментом **Описательная статистика** для заданной выборки.

Рассмотрите внимательно эту таблицу. Сравните приведенные в ней значения для среднего, дисперсии и стандартного отклонения малой выборки, а также стандартной ошибки с результатами расчетов тех же величин с использованием формул (со значениями в ячейках D8, E8, F8, I8). Уровнем надежности в таблице названа предельная ошибка.

В таблице приведены и другие характеристики выборки. *Медианой* называют варианту, которая делит вариационный ряд на две части, равные по числу вариант. *Интервалом* названа разность между наибольшим и наименьшим значениями выборки. *Максимум* – это наибольший элемент выборки, а *сумма* и *счет* – это сумма значений элементов выборки и ее объем.

Подготовить отчет

1. Название работы и задание.
2. Результаты ручного счета по заданию 2.1.
3. Две распечатки таблицы, созданной при выполнении задания 3 (одна содержит результаты вычислений, другая - сами формулы).
4. Результаты анализа данных по заданию 4.

Лабораторная работа 2

Основные определения. Систематизация выборки. Точечные оценки параметров распределения. Интервальные оценки

Цель работы – освоение методики проверки гипотезы о нормальном законе распределения с использованием инструментов статистического анализа MS Excel.

Задача проверки гипотез ставится так: на основании некоторых данных известно, что закон распределения исследуемой случайной величины есть функция $f(x)$. Требуется проверить, совместимы ли опытные данные с гипотезой о том, что случайная величина действительно имеет распределение $f(x)$.

Для выполнения этой работы необходимо изучить теоретические материалы, содержащиеся в литературе: [3], гл. 19; [2], с. 55-62; [6], с. 144-148.

Понятие статистической гипотезы о виде распределения

Часто необходимо знать закон распределения генеральной совокупности. Если закон распределения неизвестен, но имеются основания предположить, что он имеет определенный вид (например, $f(x)$), выдвигают гипотезу: генеральная совокупность распределена по закону $f(x)$. Выдвинутую гипотезу называют нулевой (основной) и обозначают H_0. Наряду с выдвинутой гипотезой рассматривают и альтернативную ей гипотезу H_1, исключающую основную гипотезу.

Выдвинутая гипотеза H_0 может быть правильной или неправильной, поэтому возникает необходимость ее проверки. Гипотезу проверяют на основании выборки, полученной из генеральной совокупности. Из-за случайности выборки в результате проверки гипотезы в двух случаях может быть принято неверное решение, т. е. могут быть допущены ошибки двух родов. *Ошибка первого рода* состоит в том, что H_0 отвергают, хотя в действительности она верна. *Ошибка второго рода* состоит в том, что принимают H_0, хотя в действительности верна H_1.

Допустимая вероятность ошибки первого рода обозначается через α и называется *уровнем значимости*. Значение ά обычно мало и устанавливается самим исследователем в зависимости от характера и важности решаемых задач. Уровень значимости, например, ά = 0,05, означает, что в среднем в 5 случаях из 100 имеется риск отвергнуть верную гипотезу H_0.

Решение относительно гипотезы H_0 принимается по значению некоторой случайной величины K, которая называется *критерием согласия*. Это такая специально подобранная величина, которая подчиняется при выполнении гипотезы H_0 некоторому известному закону распределения. Значения K зависят от выборочных данных и позволяют судить о “расхождении выборки с гипотезой H_0”.

Множество значений критерия согласия K можно разделить на два непересекающихся подмножества: подмножество значений критерия, при которых гипотеза H_0 отвергается (отклоняется), называют *критической областью*; подмножество значений критерия, при которых гипотеза H_0 не отклоняется, называется *областью принятия гипотезы*. *Критическим значением* $K_{кр}$ называют значение, отделяющее критическую область от области принятия гипотезы. Критическое значение критерия $K_{кр}$ определяется исходя из требования, чтобы при условии справедливости нулевой гипотезы вероятность того, что критерий K примет значение, большее $K_{кр}$, была равна принятому уровню значимости:

$$P(K > K_{кр}) = \alpha .$$

Для каждого критерия имеются соответствующие таблицы, по которым и находят критическую точку, удовлетворяющую этому требованию.

Принцип проверки гипотезы и принятия заключения о совместимости выборочных данных с выдвинутой гипотезой состоит в следующем: если наблюдаемое значение критерия, вычисленное по выборке, принадлежит критической области $(K_{набл} > K_{кр})$ *– нулевую гипотезу отвергают, как не согласующуюся с результатами наблюдений, если наблюдаемое значение критерия принадлежит области принятия гипотезы* $(K_{набл} < K_{кр})$ *– расхождение выборочных данных с предполагаемым законом распределения не существенно, нет оснований отвергнуть нулевую гипотезу.*

Вывод "гипотеза H_0 не отвергается" не означает, что H_0 является единственно подходящей гипотезой: просто расхождение между выборочными данными и гипотезой H_0 невелико, или иначе H_0 не противоречит результатам наблюдений; однако таким же свойством наряду с H_0 могут обладать и другие гипотезы.

Критерий согласия Пирсона

Предположим, что выполнено n измерений некоторой случайной величины ξ:

$$x_1,\ x_2, ..., x_n, \tag{4.4}$$

и есть основания полагать, что результаты распределены нормально с плотностью вероятности

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}. \tag{4.5}$$

Параметры закона распределения m и σ обычно неизвестны. Вместо неизвестных параметров подставляют значения их оценок, которые вычисляют по следующим формулам:

$$m^* = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{4.6}$$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - m^*\right)^2}. \tag{4.7}$$

В качестве критерия проверки выдвинутой гипотезы примем критерий согласия Пирсона (критерий согласия "хи- квадрат")

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(m_i - m_i^T\right)^2}{m_i^T}, \tag{4.8}$$

где k – число интервалов, на которые разбито выборочное распределение, m_i - частоты эмпирического распределения; m_i^T – частоты теоретического распределения. Из формулы вытекает, что критерий характеризует близость эмпирического и теоретического распределений: чем меньше различаются m_i и m_i^T, тем меньше значение χ^2.

Доказано, что при $n \to \infty$ закон распределения случайной величины (4.8) независимо от того, какому закону распределения подчинена генеральная совокупность, стремится к закону распределения χ^2 с r степенями свободы. Число степеней свободы определяется равенством $r = k - 1 - s$, где k - число частичных интервалов; s – число параметров предполагаемого распределения, которые были оценены. Для нормального распределения оцениваются два параметра (математическое ожидание и среднее квадратическое отклонение), поэтому $r = k - 1 - 2 = k - 3$.

В соответствии с процедурой проверки гипотезы следует вычислить наблюдаемое значение критерия. Чтобы вычислить частоты эмпирического распределения, весь интервал наблюдаемых значений делят на k частичных интервалов (бинов) точками z_k:

$$-\infty = z_0 < z_1 < z_2 < \ldots < z_{k-1} < z_k = \infty. \tag{4.9}$$

m_i определяют, подсчитав число измерений (4.4), которые попадают в i- й интервал $\left(z_{i-1}, z_i\right)$.

Используя теоретический закон распределения (4.5) можно рассчитать ожидаемое число m_i^T результатов измерений для каждого интервала i.

Вероятность того, что результат одного измерения попадает в интервал $\left(z_{i-1}, z_i\right)$, равна

$$p_i = P\left(z_{i-1} \le \xi < z_i\right) = F_N\left(z_i\right) - F_N\left(z_{i-1}\right), \tag{4.10}$$

где $F_N(z)$ – интегральный закон нормального распределения:

$F_N(z) = \int\limits_{-\infty}^{z} f(t)dt$. Учитывая, что функция распределения $F_N(x)$ с параметрами m и σ связана со стандартной нормальной функцией формулой $F_N(x) = \Phi\left(\dfrac{x-m}{\sigma}\right)$, соотношение (4.10) можно записать в следующем виде:

$$p_i = \Phi\left(\frac{z_i - m^*}{s}\right) - \Phi\left(\frac{z_{i-1} - m^*}{s}\right). \qquad (4.11)$$

Поскольку проводится не одно, а n измерений и эти измерения независимы, то их можно рассматривать как n испытаний Бернулли, в которых "успехом" считается попадание результата измерения в интервал (z_{i-1}, z_i). Тогда числа m_i^T вычисляются по формуле

$$m_i^T = n \cdot p_i \qquad (4.12)$$

(математическое ожидание числа "успехов" при n испытаниях).

Для заданного уровня значимости по таблицам определяют критическое значение критерия. Сравнивая наблюдаемое и критическое значения критерия делают, вывод о соответствии экспериментальных данных предполагаемому закону распределения.

Пример 4.1. Проверить с помощью критерия χ^2 при уровне значимости 0,05 гипотезу о том, что выборка объема $n = 50$, представленная интервальным вариационным рядом в таблице 4.4, извлечена из нормальной генеральной совокупности.

Таблица 4.4

Номер интервала i	Границы интервала	Частота m_i
1	0 – 2	5
2	2 – 4	11
3	4 –6	17
4	6 – 8	10
5	8 – 10	7

Решение. 1. Сформулируем нулевую и альтернативную гипотезы: H_0 – эмпирическое распределение соответствует нормальному; H_1 - эмпирическое распределение не соответствует нормальному.

Для проверки нулевой гипотезы необходимо рассчитать наблюдаемое значение критерия $\chi^2_{\text{набл}}$ по формуле (4.8) и сравнить его с критическим значением $\chi^2_{\text{кр}}$.

2. Определим параметры предполагаемого (теоретического) нормального закона распределения.

Найдем середины интервалов $\overline{x_i} = \frac{z_{i-1} + z_i}{2}$ и относительные частоты $p_i{}^* = \frac{m_i}{n}$. Получим следующие значения:

$\overline{x_i}$	1	3	5	6	7
p_i^*	$\frac{5}{50}$	$\frac{11}{50}$	$\frac{17}{50}$	$\frac{10}{50}$	$\frac{7}{50}$

Оценку математического ожидания найдем по формуле (4.1):

$$m^* = \left(1\cdot\frac{5}{50} + 3\cdot\frac{11}{50} + 5\cdot\frac{17}{50} + 7\cdot\frac{10}{50} + 9\cdot\frac{7}{50}\right) =$$

$$= \frac{1}{50}(5 + 33 + 85 + 70 + 63) = \frac{256}{50} = 5{,}12\ .$$

Оценки дисперсии и стандартного отклонения вычислим по формулам (4.2) и (4.3):

$$s^2 = \left((1 - 5{,}12)^2 \cdot 5 + (3 - 5{,}12)^2 \cdot 11 + (5 - 5{,}12)^2 \cdot 17 +\right.$$

$$\left. + (7 - 5{,}12)^2 \cdot 10 + (9 - 5{,}12)^2 \cdot 7\right)\cdot\frac{1}{49} = \frac{1}{49}\cdot 275{,}27 = 5{,}62\ ;$$

$$s = \sqrt{5{,}62} = 2{,}37\ .$$

3. Выполним расчет теоретических частот m_i^T по формуле (4.12). Для вычисления вероятностей p_i по формуле (4.11) воспользуемся таблицей В Приложения со значениями нормальной стандартной функции распределения. При этом наименьшее значение, т. е. z_0, полагаем равным $-\infty$, а наибольшее, т.е. z_5, полагаем равным ∞. Последовательно находим для интервала (-∞, 2)

$$p_1 = \Phi\left(\frac{2 - 5{,}12}{2{,}37}\right) - \Phi(-\infty) = \Phi(-1{,}3) - 0 = 1 - \Phi(1{,}3) = 1 - 0{,}9 = 0{,}1$$

и $m_1^T = 50 \cdot 0{,}1 = 5$;

для интервала $(2{,}4)$ находим

$$p_2 = \Phi\left(\frac{4-5{,}12}{2{,}37}\right) - \Phi\left(\frac{2-5{,}12}{2{,}37}\right) = \Phi(-0{,}5) - \Phi(-1{,}3) = 0{,}90 - 0{,}69 = 0{,}21$$

и $m_2^T = 50 \cdot 0{,}21 = 10{,}5$;

для интервала (4,6) соответственно :

$$p_3 = \Phi\left(\frac{6-5{,}12}{2{,}37}\right) - \Phi\left(\frac{4-5{,}12}{2{,}37}\right) = \Phi(0{,}4) - \Phi(-0{,}5) = 0{,}66 - 0{,}31 = 0{,}35;$$

$$m_3^T = 50 \cdot 0{,}35 = 17{,}5;$$

для интервала (6,8):

$$p_4 = \Phi\left(\frac{8-5{,}12}{2{,}37}\right) - \Phi\left(\frac{6-5{,}12}{2{,}37}\right) = \Phi(1{,}2) - \Phi(0{,}4) = 0{,}88 - 0{,}66 = 0{,}22$$

и $m_4^T = 50 \cdot 0{,}22 = 11;$

для интервала $(8, \infty)$ вычислим

$$p_5 = \Phi(\infty) - \Phi\left(\frac{8-5{,}12}{2{,}37}\right) = \Phi(\infty) - \Phi(1{,}2) = 1 - 0{,}88 = 0{,}12;$$

$$m_5^T = 50 \cdot 0{,}12 = 6.$$

4. По формуле (4.8) найдем значение $\chi^2_{набл}$:

$$\chi^2_{набл} = \sum_{i=1}^{5} \frac{\left(m_i - m_i^T\right)^2}{m_i^T} = \frac{(5-5)^2}{5} + \frac{(11-10{,}5)^2}{10{,}5} +$$

$$+ \frac{(17-17{,}5)^2}{17{,}5} + \frac{(10-11)^2}{11} + \frac{(7-6)^2}{6} = 0{,}29.$$

5. По таблице квантилей распределения χ^2 (см. таблицу С Приложения) с числом степеней свободы $r = k - 3 = 5 - 3 = 2$ находим, что $\chi^2_{кр} = 6{,}0$ для $\alpha = 0{,}05$.

Поскольку $\chi^2_{набл} < \chi^2_{кр}$ $(0{,}29 < 6{,}0)$, то можно считать, что гипотеза о нормальном распределении генеральной совокупности не противоречит опытным данным.

Порядок выполнения лабораторной работы

В данной лабораторной работе задания 1 и 2 представляют собой контрольный пример, решение которого приводится ниже. Задания 3 и 4 составляют индивидуальное задание.

Задание 1. Для выборки из 40 значений случайной величины ξ, полученной в задании 1 работы 1, оценить близость эмпирического распределения к нормальному распределению:
а) построить интервальный вариационный ряд и гистограмму частот;
б) построить на одном графике гистограмму относительных частот и график плотности нормального распределения.

Задание 2. При уровне значимостит 0,05 проверить гипотезу о нормальном распределении генеральной совокупности с использованием χ^2 - критерия как критерия согласия.

Задание 3. Для выборки нормальной случайной величины, смоделированной в задании 3 работы 1, построить, на выбор, либо гистограмму частот, либо гистограмму относительных частот.

Задание 4. Пользуясь критерием Пирсона при уровне значимости 0,01 проверить, согласуется ли с нормальным распределением статистическое распределение из задания 3 работы 1.

Выполнение задания 1

1. Подготовьте рабочий лист в EXCEL. Для этого выполните следующее:

- перейдите на новый лист и введите в ячейку B1 название таблицы *ПОСТРОЕНИЕ ГИСТОГРАММЫ;*
- назовите ярлык листа **Гистограмма**;
- разместите в ячейках A5:B24 выборку, которая была получена при выполнении задания 1.1 лабораторной работы 1 (40 значений нормальной случайной величины с параметрами $m = 4$ и σ=0,5), либо выполнив копирование значений с листа **Оценки**, либо повторив процедуру моделирования выборки заданного закона распределения.

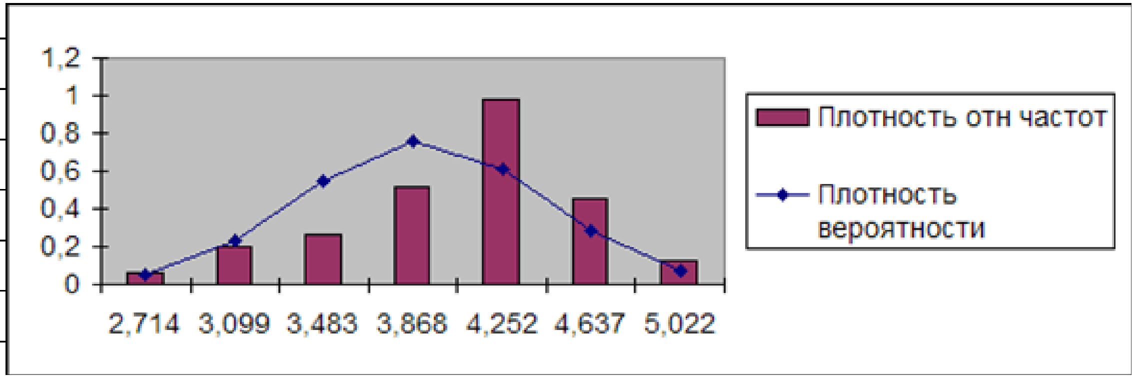
Создаваемая электронная таблица представлена в табл. 4.5 в режиме вычислений и в табл. 4.6 в режиме формул.

Таблица 4.5

	ПОСТРОЕНИЕ ГИСТОГРАММЫ							
Вариац .ряд		Объем выб.	мин	макс	k	Вел. инт-ла	Оценка мат. ож.	Несм.станд откл
2,522	3,085	40	2,5	5	6,32	0,38457	3,90417	0,52407
4,06	4,547							
4,475	3,9922							
2,868	4,1885	Левый кон.	Правый кон.	*Карман*	*Частота*	Отн. част.	Плот. отн. част.	
4,037	3,8017	2,3	2,7	2,7	1	0,025	0,06501	0,05778
4,496	3,8608	2,7	3,1	3,1	3	0,075	0,19502	0,23364
4,173	3,9781	3,1	3,5	3,5	4	0,1	0,26003	0,55136
4,163	3,296	3,5	3,9	3,9	8	0,2	0,52006	0,75941
3,786	4,0209	3,9	4,3	4,3	15	0,375	0,97511	0,61045
3,549	3,9667	4,3	4,6	4,6	7	0,175	0,45505	0,2864
3,504	3,9258	4,6	5	5	2	0,05	0,13002	0,07842
4,547	4,1243			Еще	0	контроль	1	
4,318	4,3007							
3,047	4,272							
3,788	4,2343							
4,13	3,983							
4,059	4,9519							
3,184	3,2014							
3,759	4,7737							
3,766	3,433							

Таблица 4.6

	ПОСТРОЕНИЕ ГИСТОГРАММЫ				
Вариац ряд		Объем выб	Мин	Макс	k
2,52181645110249	3,08504378254293	40	=МИН(A5:B24)	=МАКС(A5:B24)	=1+3,32*LOG10(C5)
4,05986862561258	4,54702354645997				
4,47481989895459	3,99217777713056				
2,86788828007411	4,18848368199542	левый кон	правый кон	*Карман*	*Частота*
4,0368095243175	3,80174720803916	=$D$5-$G$5/2	=C9+$G$5	2,71410160827802	1
4,49595882956055	3,86077227731585	=D9	=C10+$G$5	3,09867192262907	3
4,17288584786002	3,97809538854199	=D10	=C11+$G$5	3,48324223698012	4
4,16321280327247	3,29598743619863	=D11	=C12+$G$5	3,86781255133117	8
3,78621588070382	4,02087062966893	=D12	=C13+$G$5	4,25238286568222	15
3,54936356516554	3,96671760982281	=D13	=C14+$G$5	4,63695318003327	7
3,50391611491796	3,92577329521737	=D14	=C15+$G$5	5,02152349438432	2
4,54688484851795	4,12433702067938			Еще	0
4,31765694075148	4,30070737011556				
3,04744072409812	4,27199803298572				
3,78797518451756	4,23425229755958				
4,12970929219591	3,9829947228136				
4,05863626029168	4,951854417508				
3,18396497378125	3,20142727205529				
3,75902483129175	4,77369577411446				



2. Для построения интервального вариационного ряда выполните следующие действия:

2.1. Произведите расчет длины частичных интервалов в ячейках C5:G5 по указанным в ячейках формулам и комментариям так, как указано ниже.

Ячейка	**Значение**	**Ячейка**	**Значение**
C5	40	C4	объем выб.
D5	=МИН(A5:B24)	D4	минимум
E5	=МАКС(A5:B24)	E4	максимум
F5	=1+3,32*LOG10(C5)	F4	k
G5	=(E5-D5)/F5	G4	вел. инт-ла
H5	=СРЗНАЧ(A5:B24)	H4	оценка мат. ож.
I5	=СТАНДОТКЛОН(A5:B24)	I4	несм. станд.откл.

2.2. Разместите массив значений границ интервалов в ячейках C9:D15 (в столбце C – значения левых границ, в столбце D – значения правых границ).

Выполните это так:

- для определения левой границы первого частичного промежутка введите в ячейку C9 формулу - =$D$5-$G$5/2;
- для определения правой границы введите в ячейку D9 формулу C9+$G$5;
- поскольку левая граница последующего частичного промежутка совпадает с правой границей предыдущего введите в ячейку C10 формулу - =D9;
- перенесите автозаполнением формулу из ячейки C10 на диапазон C11:C15, а формулу из D9 – в ячейки D10:D15;
- в ячейку C8 введите текст *левый кон*, в ячейку D8 – *правый кон.*

3. Для построения гистограммы частот воспользуемся инструментом анализа **Гистограмма**. Выполните команду **Сервис – Анализ данных – Гистограмма**. В окне “Гистограмма” задайте параметры;

- введите в поле **Входной интервал** $A$5:$B$24, в поле **Интервал карманов** – $D$9:$D$15, в **Выходной интервал** – $E$8;
- установите флажок **Вывод графика;**
- нажмите **OK**.

На экране появятся выходная таблица и гистограмма. В левом столбце таблицы размещен *карман* – так в MS Excel называется набор граничных значений частичных интервалов. Правый столбец содержит вычисленные значения частот.

Поместите полученную диаграмму (выделите и перетащите) так, чтобы левый верхний конец находился в ячейке J8.

4. Подготовим исходные данные для построения гистограммы относительных частот и графика плотности вероятности.

4.1. Расчет относительных частот произведите в ячейках G9:G15, для этого введите в ячейку G9 формулу =F9/$C$5 и перенесите ее на диапазон G10:G15.

4.2. При построении гистограммы используются значения плотности относительных частот. Выполните расчет этих значений в ячейках H9:H15. Введите в ячейку H9 формулу =G9/$G$5 и скопируйте ее в ячейки H10:H15. Озаглавьте столбцы: введите в G8 текст *отн. част.*, в H8 – *плот. отн. част.*

4.3. Сформируйте в ячейках I9:I15 массив значений плотности вероятности, по которым будет построен график. Указанные значения вычислите с использованием функции НОРМРАСП в граничных точках частичных интервалов, размещенных в ячейках D9:D15. Введите в I9 формулу

=НОРМРАСП(D9;$H$5;$I$5;0)

и перенесите ее на диапазон I10:I15.

5. Как отмечалось выше, площадь гистограммы относительных частот численно равна единице. Введите для контроля правильности вычислений в ячейку G16 текст *контроль*, а в ячейку H16 – формулу =СУММ(H9:H15)*$G$5.

6. Для построения гистограммы и графика выполните следующие действия:

- выделите ячейки H9:I15, в которых размещены данные;
- нажмите кнопку **Мастер диаграмм**, откроется окно диалога;
- выберите вкладку “Нестандартные” и вид графика **График! гистограмма**, нажмите кнопку **Далее**;
- на втором шаге построения диаграммы выберите вкладку “Ряд”. Измените текст легенды (условного обозначения для рядов данных): в

разделе **Ряд** выделите **Ряд 1**, перейдите в поле **Имя** и введите текст *Плотность вероятности*, затем выделите **Ряд 2** и в поле **Имя** наберите *Плотность отн. частот*;

- введите в поле “Подписи оси X” диапазон D9:D15 и нажмите кнопку **Далее**;
- оформление гистограммы на третьем шаге можно опустить (либо выполните по своему желанию);
- на четвертом шаге задайте место размещения гистограммы –**имеющийся лист** и нажмите **OK**.

Выполнение задания 2

1. Подготовьте рабочий лист. Для этого выполните следующие действия:

- перейдите на новый лист и введите в ячейку C1 название таблицы *ПРОВЕРКА ГИПОТЕЗЫ ПО КРИТЕРИЮ ПИРСОНА*;
- назовите ярлык листа **Крит Пирсона**;
- занесите в ячейку E2 значение заданного уровня значимости 0,05, а в C2 - *уровень значимости*;
- перенесите содержимое столбцов A, B, C, D, а также четвертой и пятой строк с листа **Гистограмма** на лист **Крит Пирсона**.

Создаваемая электронная таблица представлена в табл. 4.7 в режиме формул и в табл. 4.8 в режиме вычислений.

Чтобы вычислить наблюдаемое значение критерия по формуле (4.7), для каждого частичного интервала необходимо найти значения эмпирической и теоретической частот.

2. Частоту появления значений выборки в построенных частичных интервалах (эмпирическую частоту) вычислите с помощью функции ЧАСТОТА, которая возвращает распределение частот в виде вертикального массива. Эта функция подсчитывает для данного множества значений и данного множества карманов (интервалов, в математическом смысле), сколько исходных значений попадает в каждый интервал. Выполните следующие действия:

- выделите ячейки E9:E15, в которые будет введена функция ЧАСТОТА (данная функция возвращает массив, поэтому она должна задаваться в качестве формулы массива);

- нажмите кнопку **Вставка функции**;
- в открывшемся окне диалога “Мастер функций” выберите функцию ЧАСТОТА из категории **Статистические** и нажмите кнопку **OK**;
- укажите в поле **Массив данных** диапазон $A$5:$B$24, в поле **Двоичный массив** – $D$9:$D$15 (массив верхних границ интервалов);
- не выходя из строки формул, одновременно нажмите клавиши *Ctrl+Shift+Enter*;
- введите в ячейку E7 текст *эмп. частота*, в D16 – *число бинов*, а в E16 – формулу для подсчета числа бинов

=СЧЕТ(E9:E15).

3. Расчет теоретической частоты по формулам (4.10) и (4.12) произведите в ячейках F9:H15. Выполните следующее:

- определите значения интегральной функции распределения на правом конце для каждого частичного промежутка, для чего введите в ячейку F9 формулу =НОРМРАСП(D9;$H$5;$I$5;1)
- и перенесите ее автозаполнением на диапазон F10:F14 (в ячейку F15 введите 1, поскольку $F(\infty) = 1$);
- вычислите вероятность того, что результат одного измерения попадет в частичный интервал, для чего введите в ячейку G9 формулу: =F9-F8 и скопируйте ее на диапазон G10:G15;
- сосчитайте теоретические частоты, введя в ячейку H9 формулу:

=$C$5*G9

и автозаполнением перенесите ее на диапазон H10:H15;

Таблица 4.7

	ПРОВЕРКА ГИПОТЕЗЫ ПО КРИТЕРИЮ ПИРСОНА			
		Уровень значимости		0,05
Вариац. ряд		Объем выб.	Мин.	Макс.
2,52181645110249	3,08504378254293	40	=МИН(A5:B24)	=МАКС(A5:B24)
4,05986862561258	4,54702354645997			
4,47481989895459	3,99217777713056	Левый кон.	Правый кон.	Эмп. частота
2,86788828007411	4,18848368199542			
4,0368095243175	3,80174720803916	=\$D\$5-\$G\$5/2	=C9+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
4,49595882956055	3,86077227731585	=D9	=C10+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
4,17288584786002	3,97809538854199	=D10	=C11+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
4,16321280327247	3,29598743619863	=D11	=C12+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
3,78621588070382	4,02087062966893	=D12	=C13+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
3,54936356516554	3,96671760982281	=D13	=C14+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
3,50391611491796	3,92577329521737	=D14	=C15+\$G\$5	=ЧАСТОТА(A5:B24;D9:D15)
4,54688484851795	4,12433702067938		число бинов	=СЧЁТ(E9:E15)
4,31765694075148	4,30070737011556			
3,04744072409812	4,27199803298572			
3,78797518451756	4,23425229755958			
4,12970929219591	3,9829947228136			
4,05863626029168	4,951854417508			
3,18396497378125	3,20142727205529			
3,75902483129175	4,77369577411446			
3,76566243731213	3,43298575999506			

Окончание таблицы 4.7

K	вел инт-ла	оценка мат ож	несм станд откл
=1+3,32*LOG10(C5)	=(E5-D5)/F5	=СРЗНАЧ(A5:B24)	=СТАНДОТКЛОН(A5:B24)
ф р на пр конце	вер	теор частота	
=НОРМРАСП(D9;$H$5;$I$5;1)	=F9-F8	=$C$5*G9	=(E9-H9)^2/H9
=НОРМРАСП(D10;$H$5;$I$5;1)	=F10-F9	=$C$5*G10	=(E10-H10)^2/H10
=НОРМРАСП(D11;$H$5;$I$5;1)	=F11-F10	=$C$5*G11	=(E11-H11)^2/H11
=НОРМРАСП(D12;$H$5;$I$5;1)	=F12-F11	=$C$5*G12	=(E12-H12)^2/H12
=НОРМРАСП(D13;$H$5;$I$5;1)	=F13-F12	=$C$5*G13	=(E13-H13)^2/H13
=НОРМРАСП(D14;$H$5;$I$5;1)	=F14-F13	=$C$5*G14	=(E14-H14)^2/H14
1	=F15-F14	=$C$5*G15	=(E15-H15)^2/H15
		набл зн критерия	=СУММ(I9:I15)
		крит зн критерия	=ХИ2ОБР($E$2;$E$16-3)

Таблица 4.8

		ПРОВЕРКА ГИПОТЕЗЫ ПО КРИТЕРИЮ ПИРСОНА						
		Уровень значимости		0,05				
Вариац. ряд		Объем выб.	Мин.	Макс	k	Вел. инт-ла	Оценка мат ожидания	Несм. станд. откл.
2,522	3,08504	40	2,52182	4,951854	6,318839171	0,3845703	3,904166	0,5240692
4,06	4,54702							
4,475	3,99218	левый кон.	правый кон.	Эмп. частота	Ф. р. на пр..конце	Вер.	Теор. Частота	
2,868	4,18848							
4,037	3,80175	2,3295	2,7141	1	0,011579055	0,0115791	0,463162	0,622233
4,496	3,86077	2,7141	3,09867	3	0,062146626	0,0505676	2,022703	0,4721948

4,173	3,9781	3,0987	3,48324	4	0,210934098	0,1487875	5,951499	0,6398973
4,163	3,29599	3,4832	3,86781	8	0,472348174	0,2614141	10,45656	0,577121
3,786	4,02087	3,8678	4,25238	15	0,746798016	0,2744498	10,97799	1,473542
3,549	3,96672	4,2524	4,63695	7	0,918983001	0,172185	6,887399	0,0018409
3,504	3,92577	4,637	5,02152	2	1	0,081017	3,24068	0,4749888
4,547	4,12434		число бинов	7			Набл. зн. .критерия	4,2618178
4,318	4,30071						Крит. зн. критерия	9,487729
3,047	4,272							
3,788	4,23425							
4,13	3,98299							
4,059	4,95185							
3,184	3,20143							
3,759	4,7737							
3,766	3,43299							

- поясните полученные результаты, для этого в ячейку F7 введите текст *ф. р. на пр. конце*, в ячейку G7 – *вер.*, а в H7 – *теор. частота*.

4. Вычислите слагаемые критерия Пирсона, для чего введите в ячейку I9 формулу

=(E9-H9)^2/H9

и автозаполнением перенесите эту формулу в ячейки I10:I15.

5. Наблюдаемое значение критерия вычислите по формуле (4.6) в ячейке I16, для чего введите формулу =СУММ(I9:I15).

6. Критическое значение критерия "хи-квадрат" для уровня значимости 0,95 и числа степеней свободы $r = 3$ выведите в ячейке I17, набрав формулу

=ХИ2ОБР($E$2;$E$16-3).

Функция ХИ2ОБР возвращает обратную функцию для χ^2-распределения.

В ячейку H16 введите текст *Набл. зн. критерия*, а в H17 – *Крит. зн. критерия*.

Так как наблюдаемое значение критерия, равное 4,26, меньше критического значения, равного 9,49, то гипотезу о нормальном распределении генеральной совокупности не отвергаем. Другими словами, расхождение эмпирических и теоретических частот незначительное. Следовательно, смоделированные значения случайной величины согласуются с гипотезой о распределении случайной величины с заданным законом распределения.

Подготовить отчет:

1. Название работы и задание.
2. По две распечатки таблиц, созданных при выполнении заданий 3 и 4 (одна распечатка содержит результаты вычислений, другая – сами формулы).
3. Выводы по результатам выполнения задания 4.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

Задание на контрольную работу

В контрольной работе студенту надо выполнить четыре задачи (по одной из каждого задания), при этом номера задач нужно выбрать в соответствии с последней и предпоследней цифрами шифра, а также первой буквой фамилии как показано в приведенной ниже таблице.

Посл. цифра шифра	1	2	3	4	5	6	7	8	9	0
№.задач	1	2	3	4	5	6	7	8	9	10
Предпосл. цифра шифра	1	2	3	4	5	6	7	8	9	0
№ задач	11	12	13	14	15	16	17	18	19	20
Первая буква фамилии	А,И,Т	Б,О,Ц	В,М	Г,Ф. Ч	Д,З Л,Х	Е, Н	Ж,С, Р	К,Э	П,Щ	У,Ш, Ю,Я
№ задач	30	29	28	27	26	25	24	23	22	21

Задание 1

1. Бросаются три игральных кубика. Найти вероятности событий:
A – на всех кубиках одинаковое число очков;
B – на всех кубиках выпало в сумме три очка;
C – на всех кубиках выпало в сумме более трех очков.

2. Бросаются три игральных кубика. Найти вероятности событий:

A – на всех кубиках в сумме выпало ровно четыре очка;
B – на всех кубиках в сумме выпало не менее четырех очков;
C – на всех кубиках в сумме выпало более четырех очков.

3. В коробке лежат 5 красных шаров, 6 синих и 3 желтых шара. Из коробки наугад вынимают по одному шару, не возвращая их обратно. Найти вероятность того, что при трехразовом изъятии шаров окажутся вынутыми в 1-й раз – желтый шар, во 2-й раз – красный шар и в 3-й раз – синий шар.

4. В каждой из трех коробок находится по три белых и пять красных шаров. Из каждой коробки наудачу вынимается по одному шару, не возвращая назад. Найти вероятности событий:
A – все шары белые;
B – только один шар белый;
C – хотя бы один шар белый.

5. В каждой из трех коробок находится по три белых и пять красных шаров. Из каждой коробки наудачу вынимается по одному шару. Найти вероятности событий:
A – все шары красные;
B – только один шар красный;
C – хотя бы один шар красный.

6. На сборку поступило десять деталей, среди которых четыре бракованные. Сборщик на удачу берет три детали. Найти вероятности событий:
A – все взятые детали стандартные;
B – только одна деталь среди взятых стандартная;
C – хотя бы одна из взятых деталей стандартная.

7. На сборку поступило десять деталей, среди которых четыре бракованные. Сборщик наудачу берет три детали. Найти вероятности событий: *A* – все взятые детали бракованные;

B – только одна деталь среди взятых бракованная;

C – хотя бы одна из взятых деталей бракованная.

8. В группе спортсменов два мастера спорта, шесть кандидатов в мастера и восемь перворазрядников. По жребию выбирается четыре спортсмена. Найти вероятности событий:

A – все четыре выбранные спортсмена оказались перворазрядниками;

B – среди выбранных спортсменов хотя бы один оказался перворазрядником;

C – среди выбранных спортсменов ровно половина оказалась перворазрядниками.

9. В группе спортсменов два мастера спорта, шесть кандидатов в мастера и восемь перворазрядников. По жребию выбирается четыре спортсмена. Найти вероятности событий:

A – все четыре выбранные спортсмена оказались кандидатами в мастера спорта;

B – среди выбранных спортсменов хотя бы один оказался кандидатом в мастера спорта;

C – среди выбранных спортсменов оказалось два мастера спорта и два кандидата в мастера спорта.

10. В ящике находятся **30** деталей, выполненных первым рабочим, **40** деталей, изготовленных вторым рабочим и **50** деталей, сделанных третьим рабочим. Известно, что брак, который рабочие могут допустить при работе, составляет 10%, 20% и 30% соответственно для каждого рабочего.

Найти вероятность того, что одна деталь, вынутая из ящика, окажется бракованной.

Задание 2

11. Известна плотность вероятности случайной величины

$$f(x)=\frac{1}{\sqrt{4.5\pi}}\,e^{\frac{-x^2-4x-4}{4.5}}.$$

Найти её математическое ожидание, дисперсию; построить кривую вероятности; найти вероятности событий: *A* – случайная величина примет только положительные значения; *B* – случайная величина попадает в интервал, длиной в два средних квадратических отклонения, симметричный относительно математического ожидания.

12. Случайная величина распределена по нормальному закону; среднее квадратическое отклонение её равно 5 , $P\{X<3\}=0.2$. Найти математическое ожидание, дисперсию; построить кривую вероятности; найти вероятность события: А – случайная величина попадает в интервал $(m+\sigma;\ m+2\sigma)$.

13. Случайная величина распределена по нормальному закону с математическим ожиданием m=-3, и известно, что P{X>3}=0.15. Найти её дисперсию; построить кривую вероятности; вычислить вероятность того, что случайная величина будет принимать отрицательные значения.

14. Плотность вероятности случайной величины

$$f(x)=\frac{1}{\sqrt{8\pi}}\,e^{\frac{-x^2-3x-2.25}{8}}.$$

Найти математическое ожидание случайной величины, её дисперсию; построить кривую вероятности; найти вероятности событий: *A* – случайная величина примет значение меньше 1, *B* – случайная величина примет значения, большие, чем (–2).

15. Случайная величина распределена по нормальному закону с математическим ожиданием 5 и вероятностью попадания в интервал (7;∞) равной 0,4. Найти её дисперсию; построить кривую вероятности; вычислить вероятность попадания в интервал (m-σ; m+σ).

16. Случайная величина распределена по нормальному закону с σ = 8, и известно, что вероятность попадания в интервал (-∞;4) равна 0,3. Найти её математическое ожидание, дисперсию; построить кривую вероятности; вычислить вероятности событий: *A* – случайная величина принимает положительные значения, *B* – случайная величина попадает в интервал длиной четыре средних квадратических отклонения., симметричный относительно математического ожидания,

17. Плотность вероятности случайной величины

$$f(x)=\frac{1}{\sqrt{50\pi}}\,e^{\frac{-x^2+6x-9}{50}}.$$

Найти её математическое ожидание, дисперсию; построить кривую вероятности; найти интервал, симметричный относительно математического ожидания, вероятность попадания в который равна 0,8.

18. Случайная величина распределена по нормальному закону со средним квадратическим отклонением σ = 5 и вероятностью принять значение, больше чем 10, равной 0,4. Найти её математическое ожидание, дисперсию; построить кривую вероятности; найти вероятность попадания случайной величины в интервал (-2;8).

19. Математическое ожидание случайной величины, распределенной по нормальному закону, равно –2, а вероятность попадания значений случайной величины в интервал $|\eta + 2| < 4$ равна 0,4. Найти её дисперсию; построить кривую вероятности; вычислить вероятности событий: *A* – случайная величина примет значение, большее чем $m+\sigma$; *B* – случайная величина примет отрицательные значения.

20. Плотность вероятности случайной величины имеет вид:

$$f(x)=\frac{1}{\sqrt{18\pi}}\,e^{\frac{-x^2+4x-4}{18}}.$$

Найти её математическое ожидание, дисперсию, построить кривую вероятности; найти вероятности событий: А – случайная величина примет только отрицательные значения, В – случайная величина попадает в интервал длиной в три средних квадратических отклонения, симметричный относительно математического ожидания.

Задание 3

В заданиях 21 – 30 рассматривается прибор, состоящий из двух независимо работающих блоков *A* и *B*, каждый из которых состоит из нескольких элементов. Известны вероятности отказов каждого из элементов: $p_1=0.3$, $p_2=0.2$, $p_3=0.1$, $p_4=0.1$, $p_5=0.2$, $p_6=0.2$, $p_7=0.3$. При отказе блока он подлежит полной замене, причем стоимость замены блока *A* составляет C_1, блока *B* – C_2 единиц стоимости. Предполагается, что за период времени *T* замененный блок не выйдет ещё раз из строя.

1. Найти случайную величину η – стоимость восстановления прибора за период времени *T*:

1.1. построить её ряд и функцию распределения;

1.2. вычислить математическое ожидание, дисперсию и среднее квадратическое отклонение.

2. Построить модель найденной случайной величины для двадцати приборов (методом жребия получить её 20 значений):

2.1. найти экспериментальные ряд и функцию распределения;

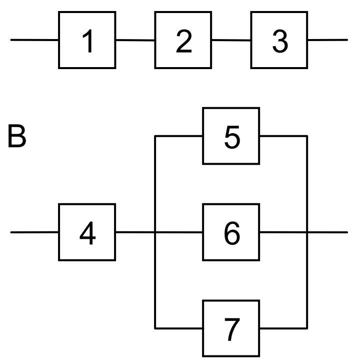
2.2. найти оценки математического ожидания, дисперсии и среднего квадратического отклонения;

2.3. построить графики теоретического и экспериментального ряда и функции распределения.

3. С помощью критерия Пирсона оценить соответствие экспериментального и теоретического распределений при уровне значимости α = 0,05.

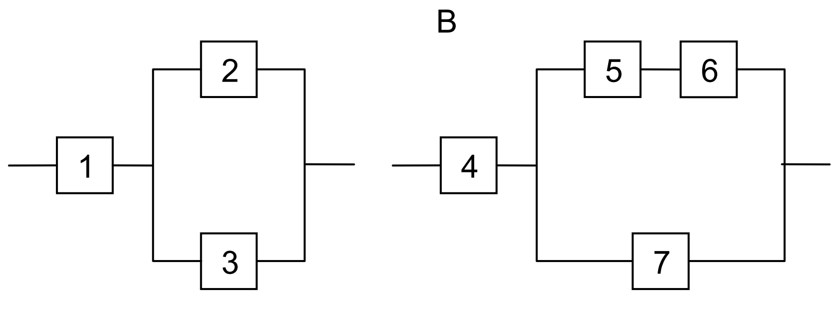
Замечание. Расчеты произвести с точностью до четырех знаков после запятой.

21. A



$C_1=3$, $C_2=5$

22. A

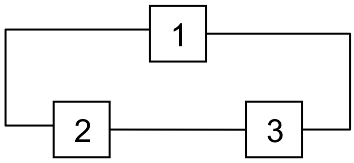


$C_1=4$, $C_2=7$

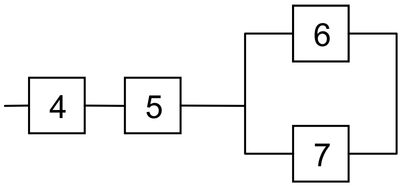
23.

24.

A

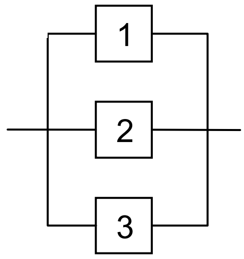


B

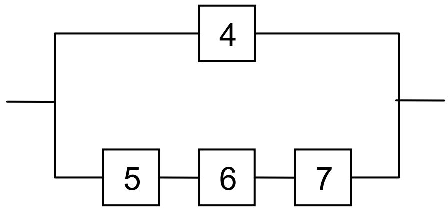


$C_1=5$, $C_2=8$

A



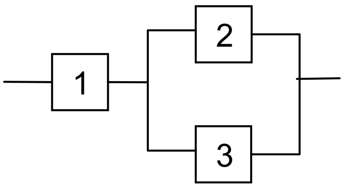
B



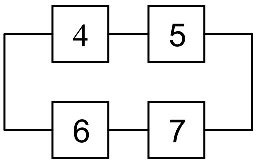
$C_1=4$, $C_2=8$

25.

A



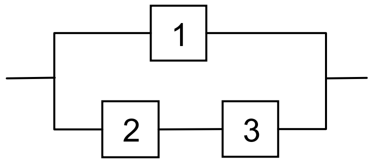
B



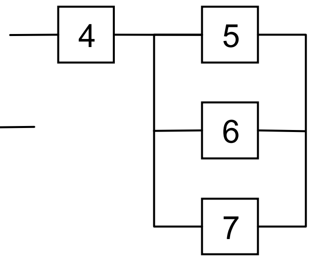
$C_1=7$, $C_2=10$

26.

A

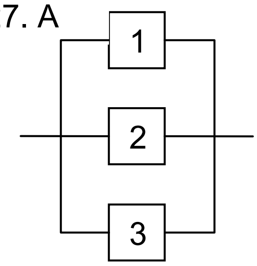


B

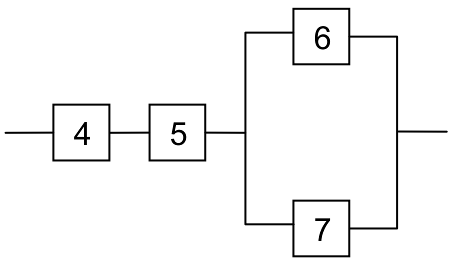


$C_1=3$, $C_2=8$

27. A

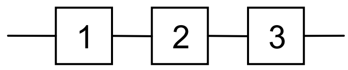


B

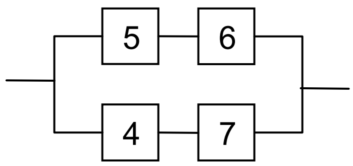


$C_1=7$, $C_2=12$

28.A

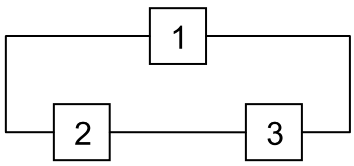


B

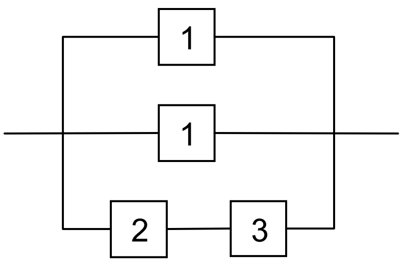


$C_1=6$, $C_2=10$

29.A

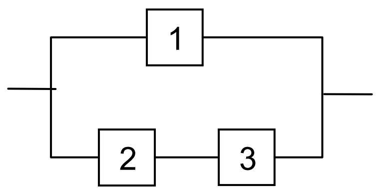


B

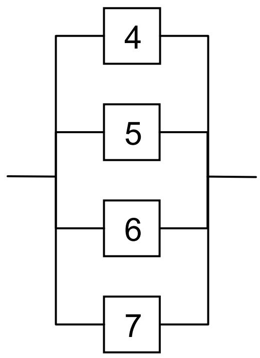


$C_1=5$, $C_2=10$

30.A



B



$C_1=8$, $C_2=10$

Задание 4

В четвертом задании предполагается, что случайная величина распределена по нормальному закону. По выборке объёмом n=20 вычислены оценки математического ожидания $m*$ и дисперсии s^2. При заданной доверительной вероятности β найти предельную ошибку оценки математического ожидания и дисперсии. Определить, какими будут эти величины, если при выборке объёмом n=40 получены такие же величины оценок. Исходные величины следует взять из таблицы, приведенной ниже.

Последняя цифра шифра	0	1	2	3	4	5	6	7	8	9
m^*	-2	-3	-4	-1	-5	-4	-3	-2	-1	-6
s^2	0,8	0,9	0,7	0,6	0,3	0,5	0,4	1,1	1,2	1,3
Предпоследняя цифра шифра	0; 5	1; 6	2; 7	3; 8	4; 9	-	-	-	-	-
Доверительная вероятность β	0,9	0,95	0,98	0,99	0,999	-	-	-	-	-

Пояснения к выполнению контрольной работы

События и их вероятности

Прежде чем приступить к решению первой задачи рекомендуется внимательно изучить понятия, теоремы и формулы, связанные с событиями. Для этого можно воспользоваться любым учебником по теории вероятностей. Кроме того, полезно разобрать примеры 5.1- 5.4, приведенные в данном комплексе.

Пример 5.1. Бросаются две игральные кости. Найти вероятность того, что сумма выпавших очков будет больше десяти.

Решение. Обозначим A событие – появление более десяти очков**,** E_{ij} – событие, состоящее в том, что на первой кости выпадает i очков, а на второй j (i,j= 1,2,…,6). События E_{ij} попарно несовместны, равновозможны, появление одного из них – событие достоверное. А это означает, что они образуют пространство элементарных событий **Ω**, позволяющее использовать классическое определение вероятности:

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

Число элементов **Ω** равно $n = 6{\cdot}6 = 36$, из них благоприятствуют событию A – появлению более десяти очков (являются его частью, влекут) всего 3 (E_{56}, E_{65}, E_{66}) и, следовательно

$$P(A)=\frac{m}{n}=\frac{3}{36}=\frac{1}{12}\approx 0{,}08.$$

Пример 5.2. На полке расставлены 15 учебников, причем 5 из них в жестком переплете. Библиотекарь берет наугад три учебника. Найти вероятность того, что хотя бы один из них окажется в жестком переплете.

Решение. Обозначим события:
A - хотя бы один учебник в жестком переплете;
A_1- один учебник в жестком переплете, два в мягком переплете;
A_2- два учебника в жестком переплете, один в мягком;
A_3 – все три учебника в жестком переплете.

По определению суммы событий искомое событие

$$A = A_1 + A_2 + A_3.$$

События A_1, A_2, A_3 несовместны, следовательно

$$P(A) = P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3).$$

Для вычисления каждой из вероятностей событий A_j (j = 1,2,3) можно применить формулу классической вероятности, так как возможность взять любую из пятнадцати книг одинакова.

При определении общего числа элементов пространства элементарных событий и числа элементов, благоприятствующих исследуемому событию, используются элементы комбинаторики. Пространство элементарных событий будут образовывать неупорядоченные последовательности книг, отличающиеся друг от друга хотя бы одной книгой. Число таких последовательностей вычисляется как число сочетаний C_r^k (читается «число сочетаний из r элементов по k элементов»). Оно может быть вычислено, в частности, по формуле

$$C_r^k = \frac{r!}{k!(r-k)!}.$$

Напомним, что 0!=1.

В решаемой задаче r = 15 (общее число книг). При вычислении общего числа элементов пространства элементарных событий k=3 (выбираются наудачу три книги), т.е.

$$n = C_{15}^3.$$

Определим число элементов, благоприятствующих событию A_1: один учебник в переплете можно выбрать только из учебников в переплете, т.е. C_5^1 способами; два учебника без переплета выбираются из 10 учебников без переплета, т.е. C_{10}^2 способами; и поскольку к любой из C_5^1 книг может присоединиться любая из пар книг, общее число возможных последовательностей m_1= $C_5^1\ C_{10}^2$. Таким образом

$$P(A_1)= \frac{C_5^1 \cdot C_{10}^2}{C_{15}^3} = \frac{5!\cdot 10!\cdot 3!\cdot 12!}{1!\cdot 4!\cdot 2!\cdot 8!\cdot 15!} = \frac{45}{91}.$$

Рассуждая аналогичным образом, получим

$$P(A_2)= \frac{C_5^2 \cdot C_{10}^1}{C_{15}^3} = \frac{5!\cdot 10!\cdot 3!\cdot 12!}{2!\cdot 3!\cdot 1!\cdot 9!\cdot 15!} = \frac{20}{91};$$

$$P(A_3)= \frac{C_5^3}{C_{15}^3} = \frac{5!\cdot 3!\cdot 12!}{3!\cdot 2!\cdot 15!} = \frac{2}{91}.$$

Окончательно

$$P(A)= P(A_1) + \mathrm{P}(A_2) + \mathrm{P}(A_3) = \frac{45}{91} + \frac{20}{91} + \frac{2}{91} = \frac{67}{91} \approx 0{,}736.$$

Однако существует более простое решение этой задачи.
Обозначим B – ни одна взятая книга не оказалась в жестком переплете. Это событие противоположно событию А – хотя бы одна книга в жестком переплете ($B=\overline{A}$ или $A=\overline{B}$). Поэтому

$$P(A) = P(\overline{B}) = 1 - P(B).$$

$$P(B) = \frac{C_{10}^3}{C_{15}^3} = \frac{10!\cdot 3!\cdot 12!}{3!\cdot 7!\cdot 15!} = \frac{24}{91}.$$

Тогда $P(A) = 1- \frac{24}{91} = \frac{67}{91} \approx 0{,}736.$
Естественно, результат оказался таким же. Заметим, что существуют и другие способы решения этой задачи, зависящие от выбора пространства элементарных событий.

Пример 5.3. Стрелок производит шесть выстрелов по мишени с вероятностью попадания при каждом выстреле 0,7. Для выхода в следующий этап соревнований нужно иметь не меньше четырех попаданий. Найти вероятность выхода стрелка в следующий этап.

Решение. Обозначим A – стрелок вышел в следующий этап, A_i – стрелок попал i раз (i= 0, 1, …, 6). Стрелок должен попасть 4, 5 или 6 раз, т.е. $A = A_4 + A_5 + A_6$. В силу несовместности событий A_i получим

$$P(A) = \mathrm{P}(A_4 + A_5 + A_6) = P(A_4) + P(A_5) + P(A_6).$$

Найдем вероятности событий A_i. Вероятность того, что стрелок попадет в цель ровно i раз, следует вычислять исходя из того, что серия опытов образует схему Бернулли, т.е. по формуле

$$P(A_i) = C_6^i\, p^i\, q^{6-i},$$

где $q = 1-p$, $i = 0,1,2,\ldots 6$.

Таким образом
$P(A) = C_6^4\, p^4\, q^2 + C_6^5\, p^5\, q + C_6^6\, p^6 = 15\cdot 0.7^4\cdot 0.3^2 + 6\cdot 0.7^5\cdot 0.3 + 0.7^6 \approx 0.324 + 0.303 + 0.118 \approx 0{,}745$.

Пример 5.4. Из четырех купленных билетов лотереи, каждый может быть как выигрышным, так и проигрышным. Найти вероятность того, что наудачу взятый билет окажется выигрышным, если количество таких билетов среди купленных равновероятны. Найти вероятность того, что если взят выигрышный билет, то и все остальные оказались такими же.

Решение. Пусть событие A – взят выигрышный билет. Обозначим через B_k событие: «k билетов выигрышные», k = 0, 1, 2, 3, 4. Эти события «предшествуют» событию A и могут рассматриваться как гипотезы. Поскольку все они равновероятны, их вероятности

$$P(B_k) = \frac{1}{5} = 0{,}2.$$

Вероятность события A вычисляется по формуле полной вероятности

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A/B_i),$$

где $P(A/B_i)$– условная вероятность события A при условии, что событие B_i произошло. Для решаемой задачи эта формула примет вид

$P(\mathrm{A}) = P(B_0)P(A/B_0) + P(B_1)P(A/B_1) + P(B_2)P(\mathrm{A}/B_2) + P(B_3)P(A/B_3) + + + $
$+P(B_4)P(A/B_4)$.

Вычислим условные вероятности появления выигрышного билета, если произошло событие B_i:

$P(A/B_0) = 0$, так как среди билетов нет ни одного выигрышного,

$P(A/B_1) = \frac{1}{4}$, так как среди четырех билетов один выигрышный.

Соответственно

$$P(A/B_2) = \frac{2}{4} = \frac{1}{2}; \quad P(A/B_3) = \frac{3}{4}; \quad \text{P}(A/B_4) = 1.$$

Подставив найденные вероятности в формулу полной вероятности, получим

$$P(A) = 0.2\cdot 0 + 0.2\cdot\frac{1}{4} + 0.2\cdot\frac{1}{2} + 0.2\cdot\frac{3}{4} + 0.2\cdot 1 = 0{,}5.$$

Переоценим доопытные, априорные, вероятности количества выигрышных среди купленных билетов после взятия выигрышного билета. Для этого воспользуемся формулой Байеса.

$$P(H_j/A) = \frac{P\left(H_j\right)P\left(A/H_j\right)}{P\left(A\right)},$$

где H_j – гипотезы, «предшествующие» событию A. Учитывая, что в задаче $H_j = B_4$, получим

$$P(H_4/A) = \frac{P\left(H_4\right)P\left(A/H_4\right)}{P\left(A\right)} = \frac{0.2\cdot 1}{0.5} = 0{,}4.$$

Таким образом, вероятность того, что все купленные билеты оказались выигрышными, увеличилась в два раза.

Случайные величины

Перед решением задач по этой теме следует усвоить основные понятия, связанные со случайными величинами: дискретные и непрерывные случайные величины, законы их распределения; изучить примеры распределений; оценить роль числовых характеристик случайных величин.

Кроме того, следует разобрать приведенные ниже примеры 5.5 – 5.8.

Пример 5.5. Прибор состоит из двух независимо работающих блоков *A* и *B*, каждый из которых собран из нескольких независимых элементов (рис.5.1), вероятности отказов которых

$$p_1 = p_2 = 0.2, \quad p_3 = p_4 = p_7 = 0.3; \quad p_5 = p_6 = 0.25, \quad p_8 = 0.278.$$

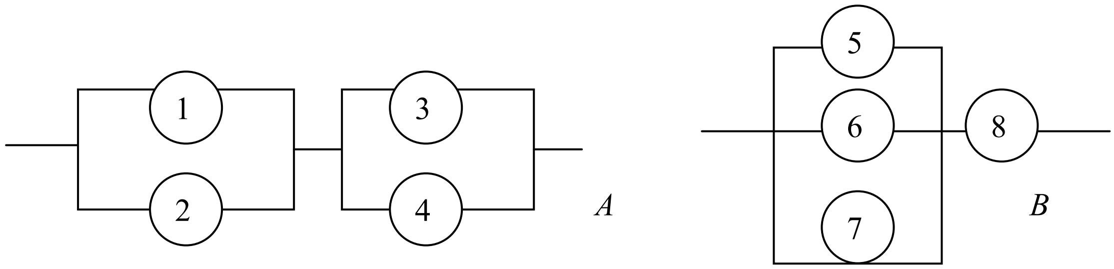


Рис.5.1

При отказе блока он подлежит полной замене, причем стоимость замены блока *A* равна $C_1 = 5$ единицам стоимости, блока *B* – $C_2 = 10$ единицам. Предполагается, что за определенный период времени *T* ни один блок не потребует повторной замены.

Найти случайную величину η – стоимость восстановления прибора за период времени *T*:

1) построить ряд и функцию распределения,
2) вычислить математическое ожидание, дисперсию и среднее квадратическое отклонение,
3) построить многоугольник распределения и график функции распределения.

Решение.

1. Определим значения случайной величины η, которая является дискретной. Случайная величина η «стоимость ремонта» может принимать только четыре значения:

$x_1 = 0$ – ни один блок не потребует замены;
$x_2 = C_1 = 5$ – только блок *A* потребует замену;
$x_3 = C_2 = 10$ – только блок В потребует замену;
$x_4 = C_1 + C_2 = 15$ – оба блока потребуют замену.

Чтобы вычислить вероятность каждого из значений x_i, следует сначала найти вероятности выхода из строя блоков *A* и *B*.

Обозначим A – выход из строя блока A, A_i – отказ i-го элемента (i = 1,2,3,4). Блок A откажет, если откажет хотя бы одна из его частей (первая состоит из элементов 1 и 2, вторая – 3 и 4). Первая часть откажет, если откажут оба элемента, т.е. произойдет событие A_1A_2, вторая - если произойдет $A_3\,A_4$. По определению суммы событий

$$A = A_1\,A_2 + A_3\,A_4.$$

В силу теоремы сложения вероятностей совместных событий

$$P(A) = P(A_1\,A_2 + A_3\,A_4) = P(A_1\,A_2) + P(\,A_3\,A_4) - P(A_1\,A_2\,A_3\,A_4).$$

В силу независимости событий A_i, получим

$$P(A) = P(A_1)\cdot P(A_2) + P(A_3)\cdot P(A_4) - P(A_1)\cdot P(A_2)\cdot P(A_3)\cdot P(A_4) =$$
$$= 0.2\cdot 0.2 + 0.3\cdot 0.3 - 0.2\cdot 0.2\cdot 0.3\cdot 0.3 = 0{,}1264.$$

Определим вероятность того, что блок A не откажет за время T (событие $\overline{A}$)

$$P(\overline{A}) = 1 - P(A) = 1 - 0{,}1264 = 0{,}8736.$$

Обозначим B – выход из строя блока B, а B_i – отказ i-го элемента (i = 5,6,7,8). Блок B потребует ремонта, если откажут все элементы ветви, состоящей из элементов 5,6 и 7, или элемент 8, а также если откажут все четыре элемента, т.е. событие B может быть записано следующим образом

$$B = B_5\,B_6\,B_7 + B_8.$$

В силу совместности и независимости событий B_i (i = 5,6,7,8) вероятность события B определяется формулой

$$P(B) = P(B_5)\,P(B_6)\,P(B_7) + P(B_8) - P(B_5)\,P(B_6)\,P(B_7)\,P(B_8).$$

Таким образом,

$P(B) = p_5p_6p_7 + p_8 - p_5p_6p_7p_8 = 0.25\cdot 0.25\cdot 0.3 + 0.278 - 0.25\cdot 0.25\cdot 0.3\cdot 0.278 = 0{,}2915.$

Найдем вероятность безотказной работы блока B:

$$P(\overline{B}) = 1 - P(B) = 1 - 0{,}2915 = 0{,}7085.$$

Найдем вероятности значений случайной величины η.

Случайная величина имеет значение $x_1 = 0$, если произойдет событие $\overline{A}\cdot\overline{B}$ (оба блока исправны за время T). События $\overline{A}, \overline{B}$ независимы, поэтому

$$P(\eta=0) = P(\overline{A})P(\overline{B}) = 0{,}8736 \cdot 0{,}7085 = 0{,}6189$$

(ограничение при вычислениях в четвертом знаке после запятой).

Значение $x_2 = 5$ принимается, если отказывает блок A и не отказывает блок B, т.е.

$$P(\eta=5) = P(A)P(\overline{B}) = 0.1264 \cdot 0.7085 = 0.0896.$$

$P(\eta=10) = P(\overline{A}\;B)$, так как должен отказать только блок B, т.е.

$P(\eta=10) = 0.8736 \cdot 0.2915 = 0.2547$. И последнее значение

$P(\eta=15) = P(A\,B) = P(A)P(\mathrm{B}) = 0.1264 \cdot 0.2915 = 0.0368.$

Запишем полученные результаты в табл.5.1, которая и будет являться рядом распределения рассматриваемой случайной величины η.

Таблица 5. 1

x_i	0	5	10	15	Σ
p_i	0.6189	0.0896	0.2547	0.0368	1.0000

Замечание. Просуммировав все вероятности и получив 1, убедимся, что избежали грубых ошибок при вычислениях.

Построим многоугольник распределения (рис.5.2):

по оси абсцисс откладываем значения случайной величины x_i;

по оси ординат значения их вероятностей p_i.

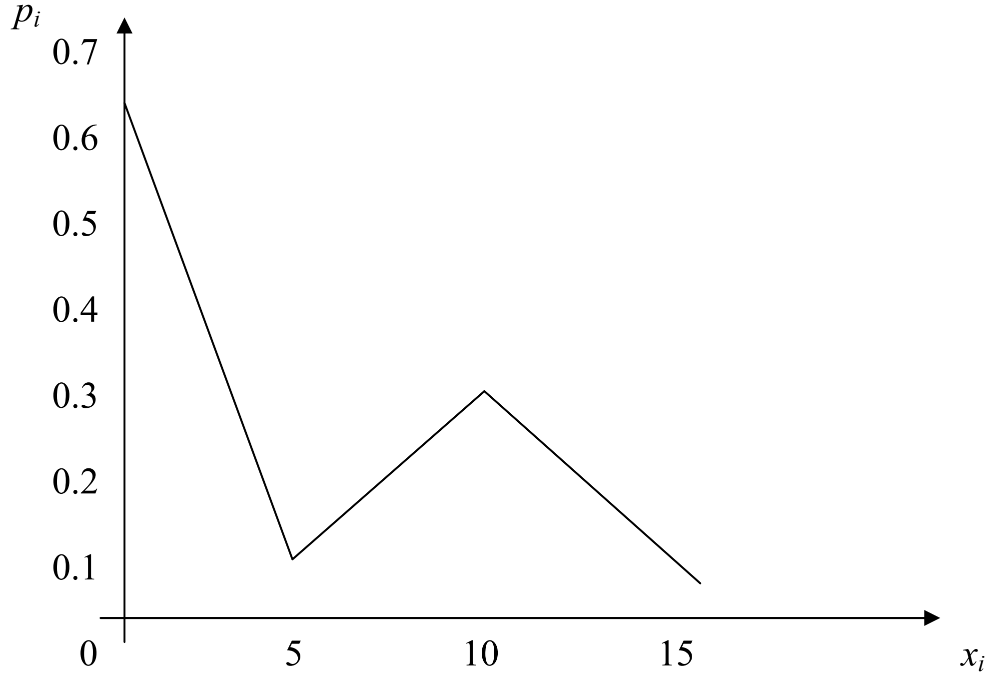


Рис. 5.2

Найдем функцию распределения случайной величины, используя соотношение:

$$F(x) = \sum_{x_i<x} p_i\,.$$

При $x<0$ $F(x)= 0$;

При $0\le x<5$ $F(x)= P(\eta=x_1)=p_1=0.6189$;

При $5\le x<10$ $F(x)= P(\eta=x_1) + P(\eta=x_2)=p_1+p_2=0.7085$;

При $10\le x<15$ $F(x)=P(\eta=x_1)+P(\eta=x_2)+ P(\eta=x_3)=p_1+p_2+p_3=0.9632$;

При $x \geq 15$ $$F(x)=\sum_{i=1}^{4} P(\eta = x_i) = 1.0000.$$

$$\text{Таким образом } F(x) = \begin{cases} 0 \text{ при } x < 0, \\ 0.6189 \text{ при } 0 \leq x < 5, \\ 0.7085 \text{ при } 5 \leq x < 10, \\ 0.9632 \text{ при } 10 \leq x < 15, \\ 1.000 \text{ при } x \geq 15. \end{cases}$$

График функции распределения изображен на рис.5. 3.

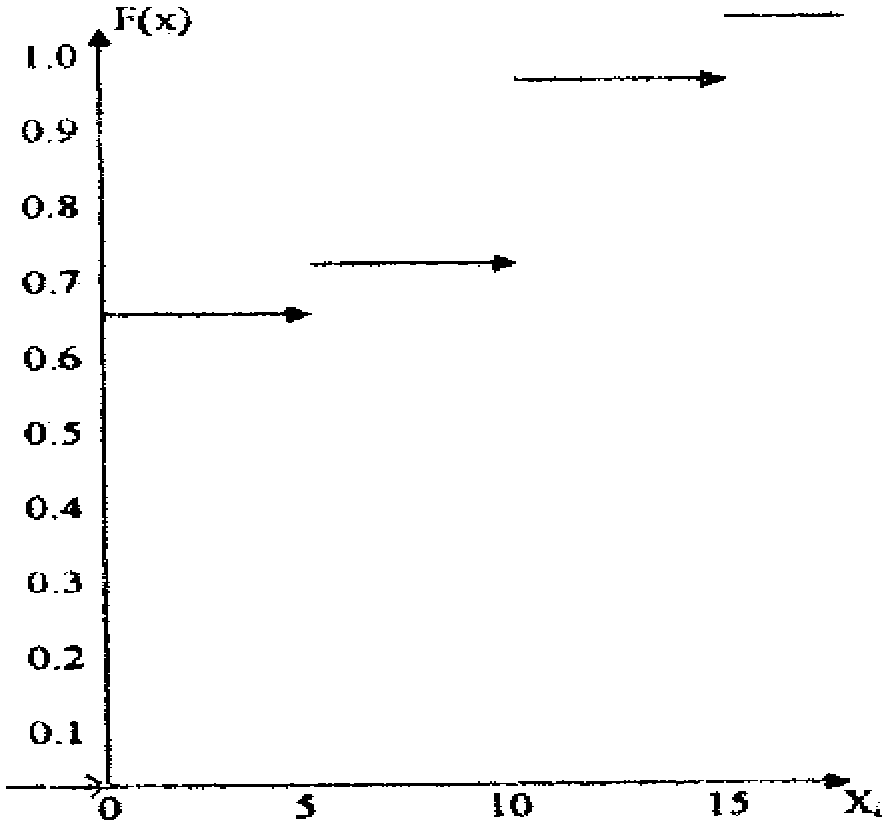


Рис. 5.3

Найдем математическое ожидание $M[\eta]$, дисперсию $D[\eta]$ и среднее квадратическое отклонение σ_η исследуемой случайной величины, воспользовавшись формулами:

$$M[\eta]= \sum_{i=1}^{n} x_i \, p_i, \tag{5.1}$$

$$D[\eta]=\sum_{i=1}^{n} \left(x_i - M[\eta]^2\right)^2 p_i, \tag{5.2}$$

$$\sigma_\eta=\sqrt{D[\eta]}. \tag{5.3}$$

Заметим, что дисперсию удобнее вычислять не по формуле (5.2), а по формуле

$$D[\eta]=M\left[\eta^2\right]-(M[\eta])^2, \qquad (5.4)$$

которая является одним из свойств дисперсии (“Дисперсия есть разность математического ожидания квадрата случайной величины и квадрата ее математического ожидания”). Все вычисления удобно записать в табл. 5. 2.

Таблица 5. 2

x_i	0	5	10	15	$\sum$
p_i	0.6189	0.0896	0.2547	0.0368	1.0000
$x_i\,p_i$	0	0.4480	2.547	0.5520	3.5470 = $M[\eta]$
$x_i^2\,p_i$	0	2.2400	25.47	8.2800	35.1900 = $M[\eta^2]$
					12.5812 = $(M[\eta])^2$
					23.4080 = $D[\eta]$

σ_η= 4.8382

Процесс вычисления достаточно ясен из самой таблицы. Первые две строки – ряд распределения случайной величины. Третья строка – произведение значений случайной величины на их вероятности; сумма в этой строке и даст математическое ожидание случайной величины согласно формуле (5.1). Четвертая строка – произведение квадрата значений случайной величины на их вероятности (достаточно умножить элементы третьей строки на элементы первой); сумма их равна математическому ожиданию квадрата случайной величины. Вычитая из этой величины квадрат математического ожидания, получим дисперсию, извлекая корень квадратный из которой найдем величину, равную среднему квадратическому отклонению.

Итак, случайная величина “стоимость ремонта” имеет среднее значение 3,547 денежных единиц со среднеквадратическим отклонением 4,8382.

Пример 5.6. Плотность вероятности случайной величины

$$f(x)=\begin{cases} 0 \text{ при } x<0, \\ 0.08\,x \text{ при } 0\le x\le 5, \\ 0 \text{ при } x>5. \end{cases}$$

Показать, что предложенная функция может быть плотностью вероятности некоторой случайной величины η. Найти её функцию распределения $F(x)$, математическое ожидание $M[\eta]$, дисперсию $D[\eta]$, среднее квадратическое отклонение σ_η и вероятность попадания случайной величины в интервал [1,3]. Построить графики плотности вероятности и функции распределения.

Решение. Покажем, что данная *f(x)* может быть плотностью вероятности.

Действительно, прежде всего, *f(x)* ≥ 0. Проверим выполнение свойства $\int_{-\infty}^{\infty} f(x)dx=1.$

$$\int_{-\infty}^{\infty} f(x)dx=\int_0^5 0.08\cdot x dx=0.08\cdot\left.\frac{x^2}{2}\right|_0^5=0.04\cdot 25=1.$$

Следовательно, предложенную функцию можно рассматривать как плотность вероятности.

Найдем функцию распределения $F(x)$ из соотношения

$$\boldsymbol{F(x)}= \int\limits_{-\infty}^{x} f(x)dx\,.$$

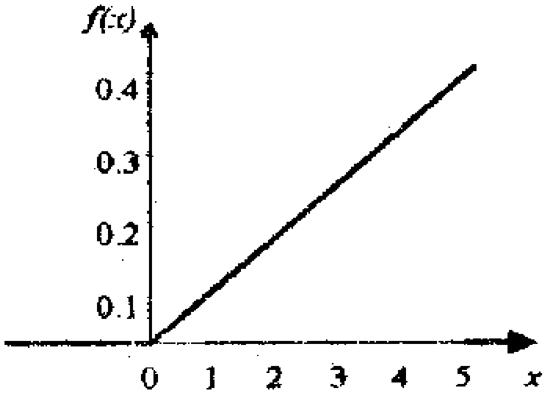
Для $x\in(-\infty;0)$ $\quad F(x)= \int\limits_{-\infty}^{x} 0dx = 0$,

для $x\in[0;5]$ $\quad F(x)= \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0\cdot dx + \int\limits_{0}^{x} 0.08xdx = 0{,}04\,x^2$,

для $x\in(5;+\infty)$ $\quad F(x)= \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0dx + \int\limits_{0}^{5} 0{,}08xdx + \int\limits_{5}^{\infty} 0dx = 1$.

Окончательно: $F(x)= \begin{cases} 0 \text{ при } x<0, \\ 0{,}04\,x^2 \text{ при } 0\le x\le 5, \\ 1 \text{ при } x>5. \end{cases}$

Построим графики плотности вероятности и функции распределения (рис. 5.4).



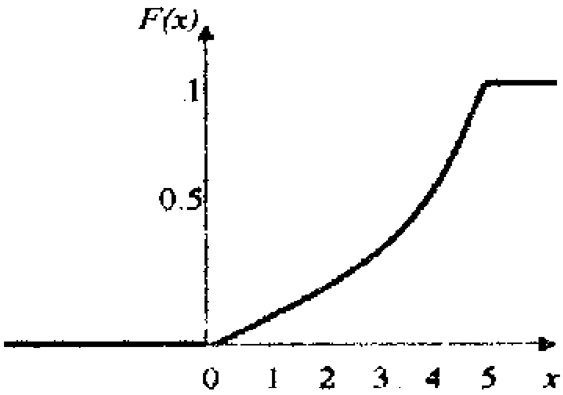


Рис. 5.4.

$$M[\eta]= \int\limits_{-\infty}^{\infty} x\cdot f(x)dx = \int\limits_{0}^{5} x\cdot 0.08xdx = 0.08\cdot \left.\frac{x^3}{3}\right|_0^5 = \frac{10}{3} \approx 3.33.$$

Для вычисления $D[\eta]$, воспользуемся её свойством

$$D[\eta]=M[\eta^2] - (M[\eta])^2 = \int\limits_{-\infty}^{\infty} x^2 f(x)dx - \left(\frac{10}{3}\right)^2 =$$

$$= \int\limits_{0}^{5} 0.08\,x^3\,dx - \frac{100}{9} = 0.02\,x^4\Big|_0^5 - \frac{100}{9} = 12.5 - \frac{100}{9} = \frac{25}{18} \approx 1.39.$$

Среднее квадратическое отклонение

$$\sigma_\eta = \sqrt{D} = \frac{5\sqrt{2}}{6} \approx 1.18\,.$$

Вероятность попадания в интервал [1,3] можно вычислить любым из двух способов, используя либо плотность вероятности, либо функцию распределения. В первом случае

$$P(1 \le \eta \le 3) = \int_1^3 f(x)dx = 0.08\int_1^3 xdx = 0.04\,x^2\Big|_1^3 = 0.32\,.$$

Во втором случае

$$P(1 \le \eta \le 3) = F(3) - F(1) = 0.04 \cdot 3^2 - 0.04 \cdot 1^2 = 0.32.$$

Пример 5.7. Функция распределения случайной величины η

$$F(x) = \begin{cases} 0 \quad \text{при } x < -\dfrac{\pi}{2}, \\ \dfrac{1}{2}(1 + \sin x) \quad \text{при } -\dfrac{\pi}{2} \le x \le \dfrac{\pi}{2}, \\ 0 \quad \text{при } x > \dfrac{\pi}{2}. \end{cases}$$

Показать, что эта функция может быть функцией распределения некоторой непрерывной случайной величины; найти её плотность вероятности $f(x)$, математическое ожидание $M[\eta]$, дисперсию $D[\eta]$, среднее квадратическое отклонение σ_η , построить графики функции распределения и плотности вероятности; найти вероятность попадания в интервал $\left[0; \dfrac{\pi}{6}\right]$.

Решение. Чтобы *F(x)* могла быть функцией распределения непрерывной случайной величины она, во-первых, должна быть определена на всей числовой оси. Действительно, *F(x)* определена на интервалах $\left(-\infty; -\dfrac{\pi}{2}\right), \left(-\dfrac{\pi}{2}; \dfrac{\pi}{2}\right), \left(\dfrac{\pi}{2}; +\infty\right)$, и в точках границ интервалов односторонние пределы функции равны её значениям. Во-вторых, её область значений должна быть [0;1]. Действительно, все значения $F(x)$ принадлежат интервалу [0,1], ($0 \le \dfrac{1}{2}(1 + \sin x) \le 1$, так как $\sin x \in [-1;1]$).

И, наконец, $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to +\infty} F(x) = 1$.

Из всего сказанного следует, что предложенная функция может быть функцией распределения непрерывной случайной величины.

Плотность вероятности найдем из соотношения $f(x) = F'(x)$. Заметим, что $F(x)$ задана разными выражениями в области её определения, следовательно, и $f(x)$ также будет описана разными выражениями на этих интервалах

$$f(\mathrm{x}) = \begin{cases} 0 \quad \text{при } x < -\dfrac{\pi}{2}, \\ \dfrac{1}{2}\cos x \quad \text{при } -\dfrac{\pi}{2} \le x \le \dfrac{\pi}{2}, \\ 0 \quad \text{при } x > \dfrac{\pi}{2}. \end{cases}$$

Построим графики $f(x)$ и $F(x)$ (рис.5.5)

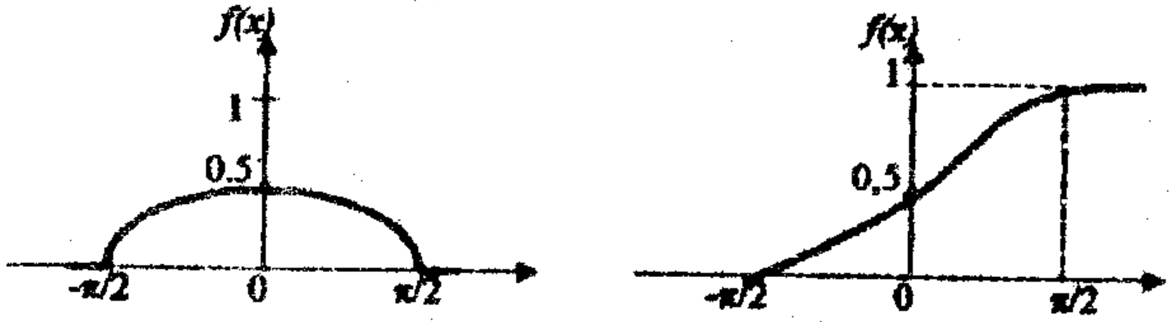


Рис.5.5

Найдем математическое ожидание:

$$M[\eta]= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x \cdot \frac{1}{2}\cos xdx = \frac{1}{2}\left[x\sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin xdx \right] = 0 .$$

Найдем дисперсию:

$$D[\eta]= \int_{-\infty}^{\infty} \left(x - M[\eta]\right)^2 f(x)dx = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x^2 \cdot \frac{1}{2}\cos xdx = \frac{1}{2}\left[x^2 \sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} - 2\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x\sin xdx \right] =$$

$$= \frac{1}{2}\left[\frac{\pi^2}{4} - \left(-\frac{\pi^2}{4}\right) - 2\left(-x\cos x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos xdx \right)\right] = \frac{1}{2}\left[\frac{\pi^2}{2} - 2\sin x \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \right] =$$

$$= \frac{1}{2}\left[\frac{\pi^2}{2} - 4\right] = \frac{\pi^2 - 8}{4} \approx 0{,}467.$$

Среднее квадратическое отклонение

$$\sigma_{\eta} = \sqrt{D[\eta]} = \frac{\sqrt{\pi^2 - 8}}{2} \approx 0{,}684.$$

Вероятность попадания в интервал $\left[0, \dfrac{\pi}{6}\right]$

$$P\left(0 \le \eta \le \frac{\pi}{6}\right) = F\left(\frac{\pi}{6}\right) - F(0) = \frac{1}{2}\left(1 + \sin\frac{\pi}{6}\right) - \frac{1}{2}\left(1 + \sin 0\right) = \frac{1}{4} = 0{,}25.$$

Пример 5.8. Случайная величина распределена по нормальному закону с плотностью вероятности

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\left(2x^2+6x+4.5\right)}.$$

Найти её математическое ожидание, дисперсию, среднее квадратическое отклонение, вероятность попадания в интервал [-2;-1]. Построить кривую плотности вероятности этой случайной величины.

Решение

1. По виду формулы плотности вероятности определяем, что случайная величина распределена по нормальному закону, для которого плотность вероятности $f(x) = \dfrac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma^2}}$. Приведем заданную функцию к стандартному виду:

$$f(x) = \frac{2}{\sqrt{2\pi}} e^{-\left(2x^2+6x+4.5\right)} = \frac{1}{\sqrt{2\pi} \cdot 0.5} e^{-\frac{(x+1.5)^2}{2 \cdot 0.5 \cdot 0.5}}.$$

Отсюда следует, что m = -1.5; σ = 0.5. Известно, что параметр m – математическое ожидание $M[\eta]$, а σ - среднее квадратическое отклонение σ_η. Следовательно, $M[\eta]$ = -1.5, σ_η=0.5, $D[\eta] = \sigma_\eta^2$ =0.25.

2. Найдем вероятность попадания заданной случайной величины в интервал [-2,-1]. По свойствам функции распределения вероятность попадания случайной величины в интервал $[\alpha;\beta)$

$$(\alpha \le \eta < \beta) = F(\beta) - F(\alpha),$$

где $F(x)$ – функция распределения случайной величины. Для нормально распределенной случайной величины функция распределения $F(x)$ может быть выражена через её нормированную функцию $\Phi(x)$ формулой:

$$F(x) = \Phi\left(\frac{x-m}{\sigma}\right). \qquad (5.5)$$

Функция $\Phi(x)$ табулирована (см. табл. В приложения).
Таким образом

$$P\left(\alpha \le \eta < \beta\right) = \Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right). \qquad (5.6)$$

Для решаемой задачи: $\alpha = -2, \beta = -1, m = -1.5$, σ=0,5 т.е.

$$P(-2 \le \eta < -1) = \Phi\left(\frac{-1-(-1.5)}{0.5}\right) - \Phi\left(\frac{-2-(-1.5)}{0.5}\right) = \Phi(1) - \Phi(-1).$$

Учитывая, что Φ(-x)=1- Φ(x), и найдя в табл.В приложения Φ(1)=0.8413, получим

$$P(-2 \le \eta < -1) = 2\Phi(1) - 1 = 0,6826.$$

3. Построим кривую плотности вероятности. Для этого на графике построим сначала кривую нормированной плотности вероятности (на рис. 5.6 штриховая линия 1), т.е. $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. Затем сожмем её по оси ординат и растянем по оси абсцисс в σ раз (т.е. максимум увеличится в два раза). Получим пунктирную линию 2. И, наконец, сдвинем по оси абсцисс на величину m влево, т.е. в данном случае максимум графика будет в точке x=-1,5. Окончательный результат на рисунке изображен сплошной линией.

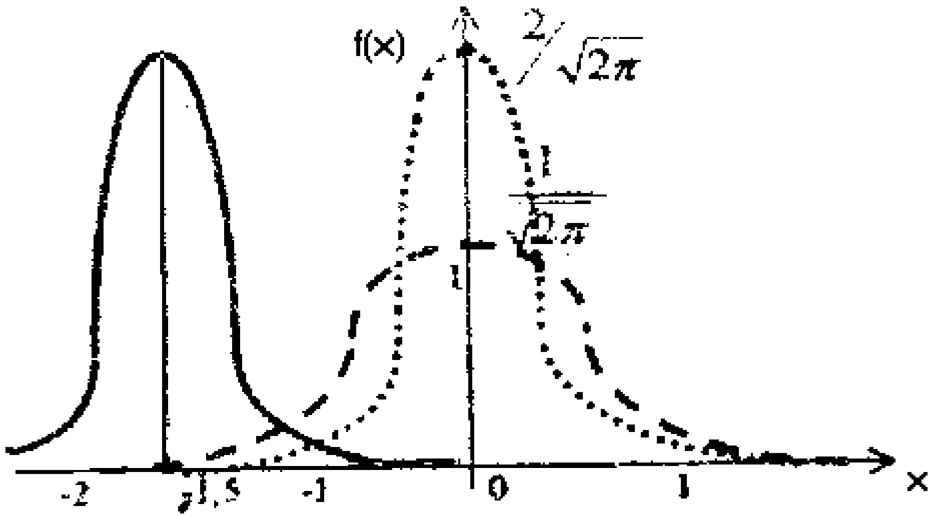


Рис. 5.6

Пример 5.9. Найти математическое ожидание и среднее квадратическое отклонение случайной величины, распределенной по нормальному закону, если $P\{X>60\}$=0,98 и $P\{X<90\}$=0,84.

Решение. Для определения искомых числовых характеристик следует найти параметры распределения предлагаемой случайной величины, так как для нормально распределенной случайной величины математическое ожидание совпадает с параметром m, а среднее квадратическое отклонение с параметром σ. Для этого воспользуемся формулой, выражающей вероятность попадания случайной величины в данные в условиях интервалы через функцию распределения. Преобразуем задания в условии задачи равенства:

из $P\{x>60\}$= 0,98 получим $p\{x \le 60\}$ = 1- $p(x>60)$ = 1-0,98. Отсюда

$$P\{x \le 60\} = 0,02.$$

По формуле (5.5) преобразуем левую часть , получим

$$F(60) = \Phi\left(\frac{60-m}{\sigma}\right) = 0,02.$$

Теперь по таблицам $\Phi(x)$ (табл. В приложения) необходимо найти значение x, при котором $\Phi(x)$ равняется 0,02. Такого значения в таблице нет, это означает, что искомое значение – отрицательное. Используя формулу

$$\Phi(-x)= 1-\Phi(x), \tag{5.7}$$

можно записать

$$\Phi(\frac{60-m}{\sigma})= 1-\Phi(\frac{m-60}{\sigma})= 0{,}02,$$

т.е. $\Phi(\frac{m-60}{\sigma})= 0{,}98$.

По табл.В приложения находим, что $\Phi(x)= 0{,}98$ соответствует значению x=2,056, т.е. $\frac{m-60}{\sigma}= 2{,}056$.

Таким образом $m\text{-}2.056\,\sigma =60$.

Из второго условия следует $P\{X<90\}=F(90)=\Phi(\frac{90-m}{\sigma}) = 0{,}84$; по табл. В Приложения находим аргумент для значения функции 0,84 и получаем $\frac{90-m}{\sigma}$=0,995 , отсюда $m+0{,}995\,\sigma =90$. Таким образом получаем систему уравнений относительно параметров m и σ:

$$\begin{cases} m-2.056\sigma = 60, \\ m+0{,}995\sigma = 90 \end{cases}$$

Находим из системы искомые параметры: $3{,}051\,\sigma =30$, $\sigma \cong 9{,}83$, $m=60+2{,}05\cdot 9{,}83 \cong 80{,}15$.

Итак, $M[\xi]$=80,15, а σ =9,83.

Элементы математической статистики

Перед тем как приступить к решению второй половины третьей задачи, следует изучить такие понятия, как: выборка, случайные числа; случайные числа распределенные по определенному закону; эмпирические (экспериментальные) ряд и функция распределения; оценки параметров распределения; метод жребия моделирования дискретной величины; критерии Пирсона оценки достоверности гипотезы; доверительные вероятности и интервалы. Следует разобрать примеры 5.10, 5.11.

Коротко рассмотрим, в чем заключается метод жребия моделирования дискретной случайной величины. Пусть событие A может произойти с вероятностью p, и пусть очередное значение случайного числа – r_i (случайное число – значение непрерывной случайной величины, равномерно распределенной в интервале [0,1]). Если $r_i \le p$, то оно принадлежит интервалу $[0, p]$, поэтому считаем, что событие A наступило. Если $r_i>p$, то считается, что событие A не наступило.

Поскольку значения случайной величины ни что иное как случайные события, процедура моделирования дискретной случайной величины с заданным законом распределения аналогична моделированию случайного события.

Пусть дискретная случайная величина задана теоретическим рядом распределения.

η	x_1	x_2	x_3	…	x_n
p_i	$p(x_1)$	$p(x_2)$	$p(x_3)$	…	$p(x_n)$

Присваиваем случайной величине η значение x_1, если значение случайного числа $r_i \le p(x_1)$, значение x_2, если $p(x_1) < r_i \le p(x_1) + p(x_2)$, т.е. в общем случае, если $\sum_{i=1}^{m-1} p(x_i) < r_j \le \sum_{i=1}^{m} p(x_i)$, то случайной величине η присваивается значение x_m.

Пример 5.10. Дискретная случайная величина задана рядом распределения, приведенным в табл. 5.3.

Таблица 5.3

x_i	0	5	10	15
p_i	0.6189	0.0896	0.2547	0.0368

1. Построить модель этой случайной величины для партии из 25 приборов (методом жребия получить её 25 значений); найти экспериментальные ряд и функцию распределения, построить их графики.

2. Найти оценки математического ожидания, дисперсии и среднего квадратического отклонения.

3. С помощью критерия Пирсона оценить соответствие экспериментального распределения теоретическому при уровнях значимости α_1=0.01, α_2= 0.05.

Решение. 1. Приступим к построению модели данной случайной величины. Этот процесс будем осуществлять методом жребия с помощью случайных чисел r_j, т.е. значений случайной величины равномерно распределенной в интервале $[0,1)$. Эти значения приведены в табл.Д Приложения. Моделируемые значения случайной величины обозначим Z_j (j = 1, 2, …, 25). Заметим, что каждое из них следует рассматривать как случайную величину.

Для рассматриваемой случайной величины правило моделирования заключается в том, чтобы определить какое значение будет принимать случайная величина в зависимости от попадания случайного числа в интервал.
η примет значение:

0, если $r_j \le 0.6189$,
5, если $0.6189 \le r_j < 0.7085$,
10, если $0.7085 \le r_j < 0.9631$,

15, если r_i>= 0.9631.

Для удобства использования правило можно свести в табл. 5.4 или изобразить на рис. 5.7.

Таблица 5. 4

	Интервал	z_j
1	0;0.619	0
2	0.619; 0.708	5
3	0.708; 0.963	10
4	0.963; 1.000	15

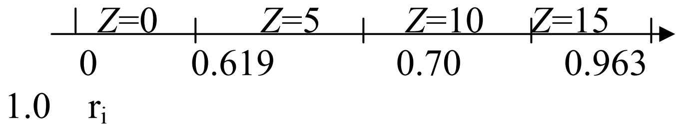


Рис. 5.7

Замечание. Поскольку в табл. 5.4 даны только 2 знака мантиссы, значения границ интервалов округлили до трех знаков после запятой.

Приступая к моделированию η, возьмем первое число из табл. Д Приложения. Для того, чтобы начало было случайным, воспользуемся днем рождения решающего задачу. Допустим, он родился 9 марта. Поэтому начнем с 9-й строки 3-го столбца. Это число 67, следовательно, r_1 = 0.67, оно принадлежит второму интервалу $[0.619;0.708]$, поэтому x_1 = 5. Таким образом, найдена стоимость ремонта первого прибора. Аналогично моделируются стоимости остальных приборов. Далее случайные числа будем выбирать двигаясь, например, по строкам влево или вправо. Второе число 43, т.е. r_2 = 0.43, оно из интервала $[0;0.619]$, поэтому x_2=0. Сведем процесс нахождения реализаций η в табл. 5.5.

Таблица 5.5

j	r_j	Интервал	z_j	j	r_j
1	0.67	0.619;0.708	5	13	0.35
2	0.43	0;0.619	0	14	0.98
3	0.97	0.963;1.000	15	15	0.95
4	0.04	0;0.619	0	16	0.11
5	0.43	0;0.619	0	17	0.68
6	0.62	0.619;0.708	5	18	0.77
7	0.76	0.705;0.963	10	19	0.12
8	0.59	0;0.619	0	20	0.17
9	0.63	0.619;0.708	5	21	0.17
10	0.57	0;0.619	0	22	0.68
11	0.33	0;0.619	0	23	0.33
12	0.21	0;0.619	0	24	0.73
				25	0.79

Найдем экспериментальный ряд распределения, для чего подсчитаем частоты m_i, равные числу приборов с данной стоимостью ремонта, т.е. числу появлений значений x_j, вычислим их относительные частоты, т.е. оценки вероятностей $p_i^* = \frac{m_i}{25}$, и занесем результаты в табл. 5.6.

Таблица 5.6

x_i	0	5	10	15	Σ
m_i	13	5	5	2	25
p_i^*	0.52	0.20	0.20	0.08	1.00

Найдем экспериментальную функцию распределения $F^*(x) = \sum\limits_{x<x_i} p_i^*$:

$$F^*(x) = \begin{cases} 0 & \text{при } \; x < 0, \\ 0.52 & \text{при } \; 0 \le x < 5, \\ 0.72 & \text{при } \; 5 \le x < 10, \\ 0.92 & \text{при } \; 10 \le x < 15, \\ 1.00 & \text{при } \; \text{x} \; \ge 15 \end{cases} .$$

Построим экспериментальные многоугольник распределения и функцию распределения (рис. 5.8). Для наглядности сравнения теоретической и экспериментальной кривых построим штриховыми линиями теоретические кривые.

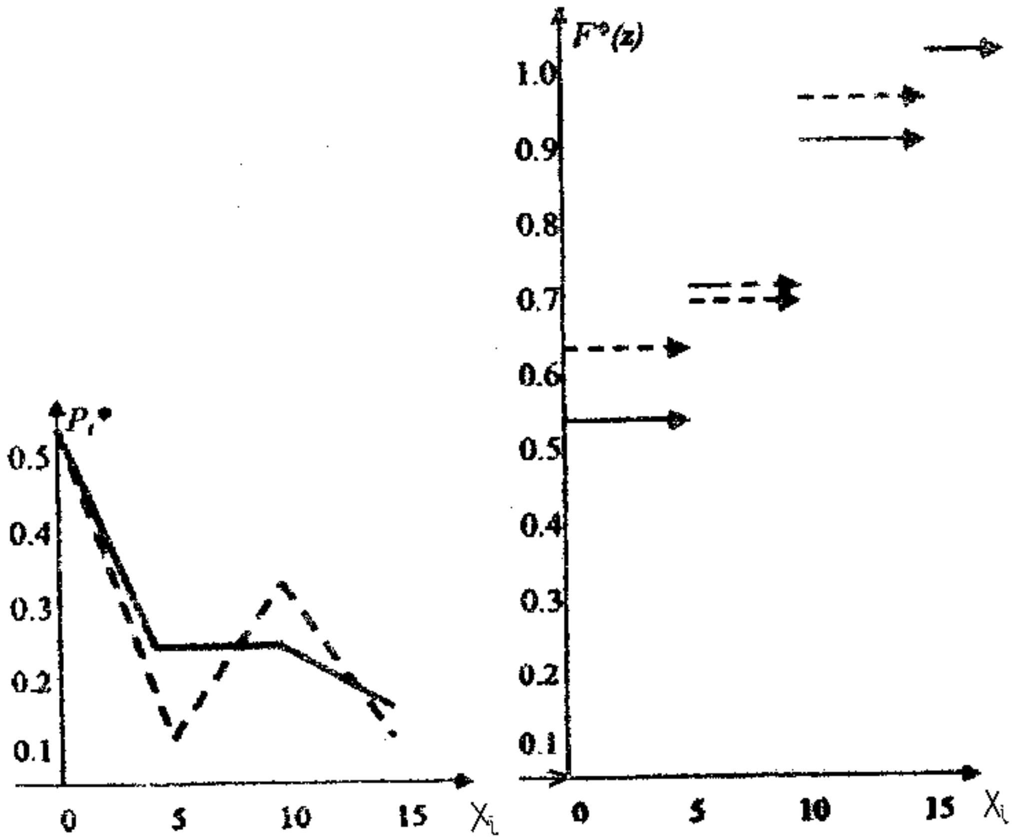


Рис. 5.8

2. Найдем оценки числовых характеристик. Для вычисления оценок математического ожидания, дисперсии и среднего квадратического отклонения воспользуемся формулами:

$$m^* = \sum_{i=1}^{k} x_i p_i \ , \tag{5.8}$$

где k – число различных значений случайной величины;

$$D[\eta] = \sum_{i=1}^{k} \left(x_i - m^*\right)^2 \ p_i \tag{5.9}$$

или

$$D^*[\eta] = M^*\left[\eta^2\right] - \left(m^*\right)^2 = \sum_{i=1}^{k} x_i^2 \ p_i - \left(m^*\right)^2 \ . \tag{5.10}$$

Поскольку формулы (5.9) и (5.10) дают смещенную оценку дисперсии, несмещенную оценку найдем по формуле

$$D^*[\eta]_{несм} = \frac{n}{n-1} D^*[\eta]_{смещ} \ . \tag{5.11}$$

Замечание. При больших значениях n коэффициент $\frac{n}{n-1}$ очень близок к единице, и можно считать оценку, вычисленную по формулам (5.9) или (5.10), оценкой несмещенной дисперсии.

Вычисления запишем в табл. 5.7.

Таблица 5.7

x_i	0	5	10	15	Σ	
p_i^*	0.52	0.20	0.20	0.08	1.00	
$x_i\, p_i^*$	0	1.00	2.00	1.20	4.20	$= m^*$
$x_i^2\, p_i^*$	0	5.00	20.00	18.00	43.00	$= M^*\left[\eta^2\right]$
					17.64	$= \left(m^*\right)^2$
					25.36	$= D^*[\eta]_{см}$

$$D^*_{несм}[\eta] = \frac{25}{25-1} D^*_{см} = 1.0417 \cdot 25.36 = 26.418\,,$$

$$\sigma^*_\eta = 5.140\,.$$

Сравнив полученные результаты с теоретическими (см. пример 5.5), видим, что экспериментальные характеристики отличаются от полученных из исходного ряда распределения. Для того чтобы получить более близкие результаты, следует существенно увеличить число реализаций случайной величины (например в два раза).

3. Проверим соответствие закона распределения полученной случайной величины $F^*(x)$ заданному закону распределения $F(x)$, используя критерий Пирсона.

Для этого определяется случайная величина

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(m_i - np_i\right)^2}{np_i}\,,$$

где k – число значений случайной величины;

m_i – число появлений значений случайной величины η;

p_i – теоретическая вероятность значения;

n – объем моделируемой выборки (np_i – ожидаемое число появлений значения x_i при n реализациях случайной величины). Величина χ^2, называемая “хи-квадрат”, служит показателем того, насколько хорошо согласуются моделируемое и ожидаемое распределения.

В статистических расчетах число степеней свободы для дискретной случайной величины определяется как $r = k - \ell - 1$, где k – число значений случайной величины, ℓ – число параметров, которые были вычислены по результатам наблюдений.

Введем понятие «критическое значение» $C = \chi^2_{\alpha,r}$ следующим образом: если при проверяемой гипотезе вероятность события $\{\chi^2 > C\}$ мала, $P(\chi^2 > C) = \alpha$, то C называется «критическим значением», а α – «уровнем значимости» критерия χ^2. Уровень значимости является вероятностью отвергнуть правильную гипотезу. Выбор его определяется решаемой задачей. Как правило, полагают α = 0.01 или α = 0.05, т.е. в одном или пяти случаях из ста может быть отвергнута правильная гипотеза. Критические значения в зависимости от объема выборки и уровня значимости приведены в табл. С Приложения А.

В рассматриваемой задаче число $k = 4$, поэтому число степеней свободы $r = 4 - 1 = 3$. По указанной таблице найдем критические числа C_1 (для $\alpha_1 = 0.01$) и C_2 (для $\alpha_2 = 0.05$): ими будут $C_1 = 11{,}3$ и $C_2 = 7{,}8$.

Найдем значение χ^2. Все вычисления выполним в таблице 5.8 (n = 25, значение np_i вычислим с точностью до одного знака после запятой).

Таблица 5.8

i	x_i	m_i	np_i	$m_i - np_i$	$\dfrac{(m_i - np_i)^2}{np_i}$
1	0	13	15.5	-2.5	0.403
2	5	5	2.2	2.8	3.536
3	10	5	6.4	-1.4	0.306
4	15	2	0.9	1.1	1.344
Σ	-	25	25.0	0.0	$5.617=\chi^2$

При уровне значимости $\alpha_2 = 0.05$ событие $\{\chi^2 > C_2\}$ не произошло ($5.617 < 7.8$); полученное распределение не противоречит предполагаемому.

При менее жестких требованиях, т.е. при $\alpha = 0.01$, событие $\{\chi^2 > C_1\}$ тем более не произошло, и в этом случае можно считать, что гипотеза о распределении случайной величины с заданным законом распределения не противоречит смоделированным значениям случайной величины.

Пример 5.11. Из выборки в 15 элементов нормальной генеральной совокупности найдены оценки математического ожидания $m^* = -1.5$ и несмещенной дисперсии $s^2 = 1.21$. Найти точность оценки математического ожидания и доверительный интервал, соответствующие доверительной вероятности $\beta = 0.98$.

Определить эти же величины для выборки в 40 элементов, если оценки оказались такими же.

Решение. Истинные математическое ожидание m и дисперсия σ^2 данного нормального распределения не известны, поэтому воспользуемся формулами $\varepsilon = t_\beta \dfrac{s}{\sqrt{n}}$ и $I_\beta = (m^* - \varepsilon;\ m^* + \varepsilon) = \left(m^* - t_\beta \dfrac{s}{\sqrt{n}};\ m^* + t_\beta \dfrac{s}{\sqrt{n}}\right)$,

где ε – предельная ошибка,

I_β – доверительный интервал, соответствующий доверительной вероятности β,

t_β – значения квантиля распределения Стьюдента для числа степеней свободы $k = n-1$.

В данной задаче число степеней свободы $k = 14$, а доверительная вероятность $\beta = 0{,}98$. По таблице А приложения значение квантилей распределения

Стьюдента находится t_β=2,62449. Тогда предельная ошибка

$\varepsilon = 2.62449 \cdot \dfrac{\sqrt{1.21}}{\sqrt{15}} \approx 0.75$ и доверительный интервал $I_{0.98}$ =(-1.5-0.75;-1.5+0.75) = =(-2.25; -0.75). Полученный результат позволяет утверждать, что с вероятностью 0.98 математическое ожидание рассматриваемой случайной величины принадлежит интервалу (-2.25;-0.75).

При выборке 40 элементов в связи с тем, что с увеличением числа степеней свободы распределение Стьюдента быстро приближается к нормальному, воспользуемся формулами $\varepsilon = \dfrac{\sigma}{\sqrt{n}} z_\beta$ для вычисления предельной ошибки оценки математического ожидания и $I_\beta = (m^* - \varepsilon;\ m^* + \varepsilon) =$

$= (m^* - \dfrac{\sigma}{\sqrt{n}} z_\beta;\ m^* + \dfrac{\sigma}{\sqrt{n}} z_\beta)$ для вычисления доверительного интервала.

В этих формулах z_β находится как корень уравнения $\Phi(z_\beta) = \dfrac{1+\beta}{2}$ по таблице значений нормированной функции распределения нормального закона (табл. В приложения). z_β называется квантилью порядка $\dfrac{1+\beta}{2}$ нормированного нормального распределения.

Вычислив $\dfrac{1+\beta}{2} = \dfrac{1+0.98}{2} = 0.99$, входим с этим значением функции в табл.В Приложения и находим её аргумент, равный 2,327.

Таким образом, точность оценки $\varepsilon = \dfrac{\sqrt{1.21}}{\sqrt{40}} \cdot 2.327 \approx 0.405$, а доверительный интервал $I_{0.98}$ = (-1.5–0.405; -1.5+0.405) = (-1.905; -1.045).
Заметим, что увеличение объема выборки существенно сузило доверительный интервал.

6. БЛОК КОНТРОЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Репетиционные вопросы

Тест № 1

Вопрос 1

Два события называются *несовместными,* если …

А) появление одного из них исключает появление другого;

В) появление одного из них не влияет на вероятность появления другого;

С) появление одного из них заключается в непоявлении другого.

Вопрос 2

Два события называются *независимыми*, если …

А) появление одного из них исключает появление другого;

В) появление одного из них не влияет на вероятность появления другого;

С) появление одного из них заключается в непоявлении другого.

Вопрос 3

Два события называются *противоположными*, если …

А) появление одного из них исключает появление другого;

В) появление одного из них не влияет на вероятность появления другого;

С) появление одного из них заключается в непоявлении другого.

Вопрос 4

Событие A называется частью события B, если …

А) появление события A влечет появление события B;

В) появление события B влечет появление события A;

С) появление события A влечет появление события $B,$ а появление события В влечет появление события A.

Вопрос 5

Если появление одного события не влияет на вероятность появления другого события, то события называются …

А) несовместными; В) независимыми; С) противоположными.

Вопрос 6

Формула $P(A+B)= P(A)+P(B)$ применима в том случае, если события А и B …

A) несовместны;

B) независимы;

C) совместны;

D) зависимы.

Вопрос 7

Формула $P(A+B)= P(A)+P(B)- P(AB)$ применима …

A) только для несовместных событий;

B) для любых событий;

C) только для совместных.

Вопрос 8

Формула $P(AB)= P(A)P(B/A)$ применима . . .

A) только для независимых событий;

B) только для зависимых событий;

C) для любых событий.

Вопрос 9

Стрелок стреляет по мишени три раза. Пусть событие А – стрелок промахнулся. Укажите, какое из приведенных событий противоположно событию А.

А) только одно попадание;

В) два попадания;

С) три попадания;

Д) хотя бы одно попадание.

Вопрос 10

Игральный кубик бросается два раза. Укажите, какое из приведенных событий является эквивалентным событию: выпало максимальное число очков.

A) выпало более одного очка;

B) выпало более шести очков;

C) выпало более одиннадцати очков;

D) выпало менее одиннадцати очков.

Вопрос 11

В урне два белых и три красных шара. Наудачу извлекается 3 шара. Укажите, какое из событий является противоположным событию: извлечены три красных шара.

A) извлечены три белых шара;

B) извлечен хотя бы один белый шар;

C) извлечены два белых и один красный шар;

D) извлечены один белый и два красных шара.

Вопрос 12

Стрелок стреляет по мишени два раза с вероятностью попадания при каждом выстреле 0.6. Найти вероятность того, что оба раза мишень будет поражена.

A) 0.36; B) 0.76; C) 0.16; D) 0.84.

Вопрос 13

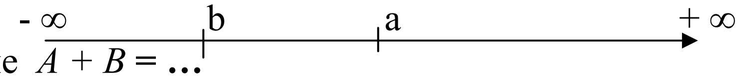
В одной урне 6 белых и 4 цветных шара, в другой – 4 белых и четыре цветных. Из наугад взятой урны наугад выбирается один шар. Найти вероятность того, что он будет белым.

A) 0.90; B) 0.45; C) $5/9 \cong 0.56$; D) 0.55.

Вопрос 14

Событие А = {переменная *X* принадлежит промежутку (- ∞ **,** *a*)}

Событие *B* = { переменная *X* принадлежит промежутку *(* *b* , + ∞)}.



Событие *A* + *B* **= …**

А) {переменная *X* принадлежит промежутку (- ∞ , + ∞)};

В) {переменная *X* принадлежит промежутку (*b,* *a*)};

С) {переменная *X* принадлежит промежутку (a, ∞)};

D) {переменная *X* принадлежит промежутку (- ∞ , *a*).

Вопрос 15

Подбрасываются две различные монеты. Пространство элементарных событий: Ω={ГГ,ГЦ,ЦГ,ЦЦ}.

Множество элементов пространства элементарных событий, благоприятствующих событию А={на первой монете появился герб} …

А) { ГГ, ГЦ}; В) {ГГ, ЦЦ}; С) { ГГ, ЦГ}; D) {ГЦ, ЦГ}.

Вопрос 16

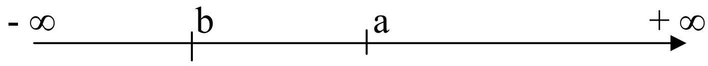
В цехе 4 станка. Вероятность того, что каждый из станков работает в данный момент, равна 0,9. Найти с точностью до сотых вероятность того, что в данный момент включены все станки.

А) 0,66; В) 0,5; С) 1**;** **D)** 0,81.

Вопрос 17

Событие *A* = {переменная *X* принадлежит промежутку (- ∞ **,** *a*)}.

Событие *B* = { переменная *X* принадлежит промежутку (*b* , + ∞)}.



Событие *AB* **= …**

А) {переменная **Х** принадлежит промежутку (*b,* *a*)};

В) {переменная **Х** принадлежит промежутку (- ∞ , + ∞)};

С) {переменная **Х** принадлежит промежутку (*a,* ∞)};

D) {переменная **Х** принадлежит промежутку (- ∞ , *a*).

Вопрос 18

В группе 20 человек. На студенческую конференцию надо выбрать <u>двух</u> человек. Сколькими способами это можно сделать?

А) 380; В) 190; С) 15; D) 400.

Вопрос 19

Подбрасываются две различные монеты. Пространство элементарных событий: Ω={ГГ,ГЦ,ЦГ,ЦЦ}.

Множество элементов пространства элементарных событий, благоприятствующих событию *A* = {на первой монете появился герб} …

А) {ГГ, ГЦ}; В) {ГГ, ЦЦ}; С) {ГГ, ЦГ}; D) {ГЦ, ЦГ}.

Вопрос 20

Для участия в олимпиаде выделено из первой группы 5 студентов, из второй – 2 студента, из третьей – 3 студента. Вероятность того, что студент станет участником олимпиады для первой группы равна **0,9**, для второй – **0,7**, для третьей – **0,8**. Найти вероятность того, что наудачу выбранный студент попал на олимпиаду.

A) 0,83; B) 0,35; C) 0,24; D) 0,29

Тест № 2

Вопрос 1

Случайная величина задана рядом распределения:

i	-3	-1	0	3	5
p_i	0.2	0.2	0.4	0.1	0.1

Математическое ожидание этой случайной величины равно 0. Найти её дисперсию. A) 4,8; B) 5,4; C) 1; D) 2,4.

Вопрос 2

Случайная величина задана плотностью вероятности

$$f(x) = \begin{cases} 0 \text{ при } x < 1, \\ 0.25 \text{ при } 1 \le x \le 5, \\ 0 \text{ при } x > 5. \end{cases}$$

Найти математическое ожидание этой случайной величины.

A) 3; B) 4/3; C) $2/\sqrt{3}$; D) 0.25.

Вопрос 3

Случайная величина задана плотностью вероятности

$$f(x) = \begin{cases} 0 \text{ при } x < 1, \\ 0.25 \text{ при } 1 \le x \le 5, \\ 0 \text{ при } x > 5. \end{cases}$$

Найти дисперсию данной случайной величины.

A) 4/3; B) 1.3; C) $2/\sqrt{3}$; D) 0.25.

Вопрос 4

Дана плотность вероятности случайной величины η, распределенной по нормальному закону.

$$f(x) = \frac{1}{\sqrt{18\pi}} e^{\frac{-x^2 - 14x - 49}{18}}$$

Математическое ожидание и дисперсия этой случайной величины равны …

A) M[η] = -7; D[η] = 3; B) M[η] = 7; D[η] = 3;

C) M[η] = -7; D[η] = 9; D) M[η] = 7 ; D[η] = 9.

Вопрос 5

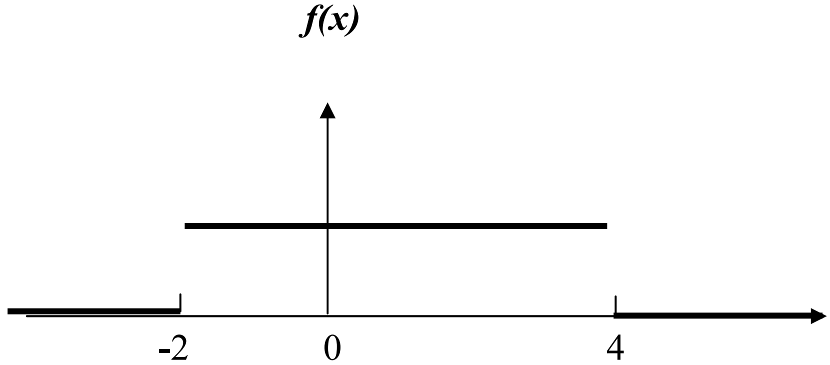
Случайная величина задана функцией распределения

$$F(x) = \frac{1}{\sqrt{18\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{18}} dx.$$

Найти вероятность того, что случайная величина примет положительные значения. A) 0,8413 B) 0,1587; C) 0,3413; D) 0,5;

Вопрос 6

Дан график плотности вероятности *f(x)* случайной величины, равномерно распределенной на интервале [**-2, 4**].



Математическое ожидание и дисперсия этой случайной величины равны соответственно …

A) M[η] = 1; D[η] = 3; B) M[η]=2; D[η]=3;

C) M[η]=1; D[η]=9; **D**. M[η]=2 ; D[η]=9.

Вопрос 7

Случайная величина задана рядом распределения:

x_i	-3	-1	0	1	3
p_i	0.2	0.2	0.4	0.1	0.1

Найти вероятность того, что случайная величина будет принимать положительные значения.

A) 0,2; B) 0,4; C) 0,6; D) 0,8.

Таблица ответов к тесту №1

№ воп-роса	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
Ответ	**A**	**B**	**C**	**A**	**B**	**A**	**B**	**C**	**D**	**C**	**D**	**A**	**D**	**A**	**A**	**A**	**A**	**A**	**A**	**A**

Таблица ответов к тесту №2

№ Вопроса	1	2	3	4	5	6	7
Ответ	B	A	B	C	D	A	C

Вопросы для экзамена по курсу « Математика.Часть 2. Теория вероятностей и элементы математической статистики»

1. Пространство элементарных событий.
2. Сумма, произведение, разность событий. Несовместные события.
3. Относительные частоты событий, закон устойчивости относительных частот.
4. Аксиомы теории вероятностей.
5. Противоположные события. Вероятность суммы событий.
6. Элементы комбинаторики (размещения, сочетания из n элементов по k элементов).
7. Классическое определение вероятности, примеры.
8. Гипергеометрическое распределение.
9. Геометрическое определение вероятности, примеры.
10. Условная вероятность, примеры.
11. Теорема умножения вероятностей для нескольких событий.
12. Попарно независимые события. Привести примеры.
13. Независимые события в совокупности.
14. Формула полной вероятности.
15. Формула Байеса.
16. Последовательность независимых испытаний (схема Бернулли).
17. Дискретная случайная величина и ряд распределения.
18. Функция распределения и ее свойства.
19. Плотность вероятности распределения случайной величины.
20. Плотность вероятности равномерного распределения.
21. Математическое ожидание случайной величины.
22. Дисперсия случайной величины.
23. Математическое ожидание и дисперсия равномерно распределенной случайной величины.
24. Биномиальный закон распределения.
25. Нормальное распределение случайной величины и его свойства.
26. Закон больших чисел, центральная предельная теорема.
27. Генеральная совокупность, числовые характеристики генеральной совокупности.
28. Выборка, выборочное среднее.
29. Состоятельные и несмещенные оценки (дисперсия, стандартное отклонение).
30. Интервальный вариационный ряд, гистограмма.
31. Доверительный интервал для математического ожидания при известной дисперсии (большая выборка).
32. Доверительный интервал для математического ожидания при неизвестной дисперсии (малая выборка).
33. Проверка гипотез по критерию Пирсона.

ПРИЛОЖЕНИЕ

Таблица А

Значения квантилей распределения Стьюдента

Число степеней свободы	Доверительные вероятности				
	0,9	0,95	0,98	0,99	0,999
1	6,31375	12,7062	31,821	63,6559	636,578
2	2,91999	4,30266	6,96455	9,924988	31,5998
3	2,35336	3,18245	4,54071	5,840848	12,9244
4	2,13185	2,77645	3,74694	4,60408	8,61008
5	2,01505	2,57058	3,36493	4,032117	6,8685
6	1,94318	2,44691	3,14267	3,707428	5,95872
7	1,89458	2,36462	2,99795	3,499481	5,40807
8	1,85955	2,30601	2,89647	3,355381	5,04137
9	1,83311	2,26216	2,82143	3,249843	4,78089
10	1,81246	2,22814	2,76377	3,169262	4,58676
11	1,79588	2,20099	2,71808	3,105815	4,43688
12	1,78229	2,17881	2,68099	3,054538	4,31784
13	1,77093	2,16037	2,6503	3,012283	4,22093
14	1,76131	2,14479	2,62449	2,976849	4,14031
15	1,75305	2,13145	2,60248	2,946726	4,07279
16	1,74588	2,1199	2,58349	2,920788	4,01487
17	1,73961	2,10982	2,56694	2,898232	3,96511
18	1,73406	2,10092	2,55238	2,878442	3,92174
19	1,72913	2,09302	2,53948	2,860943	3,88332
20	1,72472	2,08596	2,52798	2,845336	3,84956
21	1,72074	2,07961	2,51765	2,831366	3,8193
22	1,71714	2,07388	2,50832	2,818761	3,79223
23	1,71387	2,06865	2,49987	2,807337	3,76764
24	1,71088	2,0639	2,49216	2,796951	3,74537
25	1,70814	2,05954	2,4851	2,787438	3,72514
26	1,70562	2,05553	2,47863	2,778725	3,70666
27	1,70329	2,05183	2,47266	2,770685	3,68949
28	1,70113	2,04841	2,46714	2,763263	3,67392
29	1,69913	2,04523	2,46202	2,756387	3,65952
30	1,69726	2,04227	2,45726	2,749985	3,64598
60	1,67065	2,0003	2,39012	2,660272	3,46015
120	1,65765	1,97993	2,35783	2,617417	3,37342

Таблица В

Значения нормальной стандартной функции распределения

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
0	0,5	0,3	0,6179	0,6	0,7257	0,9	0,8159
0,01	0,504	0,31	0,6217	0,61	0,7291	0,91	0,8186
0,02	0,508	0,32	0,6255	0,62	0,7324	0,92	0,8212
0,03	0,512	0,33	0,6293	0,63	0,7357	0,93	0,8238
0,04	0,516	0,34	0,6331	0,64	0,7389	0,94	0,8264
0,05	0,5199	0,35	0,6368	0,65	0,7422	0,95	0,8289
0,06	0,5239	0,36	0,6406	0,66	0,7454	0,96	0,8315
0,07	0,5279	0,37	0,6443	0,67	0,7486	0,97	0,834
0,08	0,5319	0,38	0,648	0,68	0,7517	0,98	0,8365
0,09	0,5359	0,39	0,6517	0,69	0,7549	0,99	0,8389
0,1	0,5398	0,4	0,6554	0,7	0,758	1	0,8413
0,11	0,5438	0,41	0,6591	0,71	0,7611	1,01	0,8438
0,12	0,5478	0,42	0,6628	0,72	0,7642	1,02	0,8461
0,13	0,5517	0,43	0,6664	0,73	0,7673	1,03	0,8485
0,14	0,5557	0,44	0,67	0,74	0,7704	1,04	0,8508
0,15	0,5596	0,45	0,6736	0,75	0,7734	1,05	0,8531
0,16	0,5636	0,46	0,6772	0,76	0,7764	1,06	0,8554
0,17	0,5675	0,47	0,6808	0,77	0,7794	1,07	0,8577
0,18	0,5714	0,48	0,6844	0,78	0,7823	1,08	0,8599
0,19	0,5753	0,49	0,6879	0,79	0,7852	1,09	0,8621
0,2	0,5793	0,5	0,6915	0,8	0,7881	1,1	0,8643
0,21	0,5832	0,51	0,695	0,81	0,791	1,11	0,8665
0,22	0,5871	0,52	0,6985	0,82	0,7939	1,12	0,8686
0,23	0,591	0,53	0,7019	0,83	0,7967	1,13	0,8708
0,24	0,5948	0,54	0,7054	0,84	0,7995	1,14	0,8729
0,25	0,5987	0,55	0,7088	0,85	0,8023	1,15	0,8749
0,26	0,6026	0,56	0,7123	0,86	0,8051	1,16	0,877
0,27	0,6064	0,57	0,7157	0,87	0,8078	1,17	0,879
0,28	0,6103	0,58	0,719	0,88	0,8106	1,18	0,881
0,29	0,6141	0,59	0,7224	0,89	0,8133	1,19	0,883

Продолжение таблицы В

X	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
1,2	0,8849	1,5	0,9332	1,8	0,9641	2,1	0,9821
1,21	0,8869	1,51	0,9345	1,81	0,9649	2,11	0,9826
1,22	0,8888	1,52	0,9357	1,82	0,9656	2,12	0,983
1,23	0,8907	1,53	0,937	1,83	0,9664	2,13	0,9834
1,24	0,8925	1,54	0,9382	1,84	0,9671	2,14	0,9838
1,25	0,8944	1,55	0,9394	1,85	0,9678	2,15	0,9842
1,26	0,8962	1,56	0,9406	1,86	0,9686	2,16	0,9846
1,27	0,898	1,57	0,9418	1,87	0,9693	2,17	0,985
1,28	0,8997	1,58	0,9429	1,88	0,9699	2,18	0,9854
1,29	0,9015	1,59	0,9441	1,89	0,9706	2,19	0,9857
1,3	0,9032	1,6	0,9452	1,9	0,9713	2,2	0,9861
1,31	0,9049	1,61	0,9463	1,91	0,9719	2,21	0,9864
1,32	0,9066	1,62	0,9474	1,92	0,9726	2,22	0,9868
1,33	0,9082	1,63	0,9484	1,93	0,9732	2,23	0,9871
1,34	0,9099	1,64	0,9495	1,94	0,9738	2,24	0,9875
1,35	0,9115	1,65	0,9505	1,95	0,9744	2,25	0,9878
1,36	0,9131	1,66	0,9515	1,96	0,975	2,26	0,9881
1,37	0,9147	1,67	0,9525	1,97	0,9756	2,27	0,9884
1,38	0,9162	1,68	0,9535	1,98	0,9761	2,28	0,9887
1,39	0,9177	1,69	0,9545	1,99	0,9767	2,29	0,989
1,4	0,9192	1,7	0,9554	2	0,9772	2,3	0,9893
1,41	0,9207	1,71	0,9564	2,01	0,9778	2,31	0,9896
1,42	0,9222	1,72	0,9573	2,02	0,9783	2,32	0,9898
1,43	0,9236	1,73	0,9582	2,03	0,9788	2,33	0,9901
1,44	0,9251	1,74	0,9591	2,04	0,9793	2,34	0,9904
1,45	0,9265	1,75	0,9599	2,05	0,9798	2,35	0,9906
1,46	0,9279	1,76	0,9608	2,06	0,9803	2,36	0,9909
1,47	0,9292	1,77	0,9616	2,07	0,9808	2,37	0,9911
1,48	0,9306	1,78	0,9625	2,08	0,9812	2,38	0,9913
1,49	0,9319	1,79	0,9633	2,09	0,9817	2,39	0,9916

Окончание таблицы В

X	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
2,4	0,9918	2,7	0,9965	3	0,9987	3,3	0,9995
2,41	0,992	2,71	0,9966	3,01	0,9987	3,31	0,9995
2,42	0,9922	2,72	0,9967	3,02	0,9987	3,32	0,9995
2,43	0,9925	2,73	0,9968	3,03	0,9988	3,33	0,9996
2,44	0,9927	2,74	0,9969	3,04	0,9988	3,34	0,9996
2,45	0,9929	2,75	0,997	3,05	0,9989	3,35	0,9996
2,46	0,9931	2,76	0,9971	3,06	0,9989	3,36	0,9996
2,47	0,9932	2,77	0,9972	3,07	0,9989	3,37	0,9996
2,48	0,9934	2,78	0,9973	3,08	0,999	3,38	0,9996
2,49	0,9936	2,79	0,9974	3,09	0,999	3,39	0,9997
2,5	0,9938	2,8	0,9974	3,1	0,999	3,4	0,9997
2,51	0,994	2,81	0,9975	3,11	0,9991	3,41	0,9997
2,52	0,9941	2,82	0,9976	3,12	0,9991	3,42	0,9997
2,53	0,9943	2,83	0,9977	3,13	0,9991	3,43	0,9997
2,54	0,9945	2,84	0,9977	3,14	0,9992	3,44	0,9997
2,55	0,9946	2,85	0,9978	3,15	0,9992	3,45	0,9997
2,56	0,9948	2,86	0,9979	3,16	0,9992	3,46	0,9997
2,57	0,9949	2,87	0,9979	3,17	0,9992	3,47	0,9997
2,58	0,9951	2,88	0,998	3,18	0,9993	3,48	0,9997
2,59	0,9952	2,89	0,9981	3,19	0,9993	3,49	0,9998
2,6	0,9953	2,9	0,9981	3,2	0,9993	3,5	0,9998
2,61	0,9955	2,91	0,9982	3,21	0,9993	3,6	0,9998
2,62	0,9956	2,92	0,9982	3,22	0,9994	3,7	0,9998
2,63	0,9957	2,93	0,9983	3,23	0,9994	3,8	0,9999
2,64	0,9959	2,94	0,9984	3,24	0,9994	3,9	1
2,65	0,996	2,95	0,9984	3,25	0,9994	4	1
2,66	0,9961	2,96	0,9985	3,26	0,9994		
2,67	0,9962	2,97	0,9985	3,27	0,9995		
2,68	0,9963	2,98	0,9986	3,28	0,9995		
2,69	0,9964	2,99	0,9986	3,29	0,9995		

Таблица С

Значения квантилей распределения Стьюдента

Число степеней свободы	Уровень значимости					
	0,001	0,005	0,01	0,025	0,05	0,1
1	10,83	7,879	6,635	5,024	3,841	2,706
2	13,82	10,6	9,21	7,378	5,991	4,605
3	16,27	12,84	11,34	9,348	7,815	6,251
4	18,47	14,86	13,28	11,14	9,488	7,779
5	20,51	16,75	15,09	12,83	11,07	9,236
6	22,46	18,55	16,81	14,45	12,59	10,64
7	24,32	20,28	18,48	16,01	14,07	12,02
8	26,12	21,95	20,09	17,53	15,51	13,36
9	27,88	23,59	21,67	19,02	16,92	14,68
10	29,59	25,19	23,21	20,48	18,31	15,99
11	31,26	26,76	24,73	21,92	19,68	17,28
12	32,91	28,3	26,22	23,34	21,03	18,55
13	34,53	29,82	27,69	24,74	22,36	19,81
14	36,12	31,32	29,14	26,12	23,68	21,06
15	37,7	32,8	30,58	27,49	25	22,31
16	39,25	34,27	32	28,85	26,3	23,54
17	40,79	35,72	33,41	30,19	27,59	24,77
18	42,31	37,16	34,81	31,53	28,87	25,99
19	43,82	38,58	36,19	32,85	30,14	27,2
20	45,31	40	37,57	34,17	31,41	28,41
21	46,8	41,4	38,93	35,48	32,67	29,62
22	48,27	42,8	40,29	36,78	33,92	30,81
23	49,73	44,18	41,64	38,08	35,17	32,01
24	51,18	45,56	42,98	39,36	36,42	33,2
25	52,62	46,93	44,31	40,65	37,65	34,38
26	54,05	48,29	45,64	41,92	38,89	35,56
27	55,48	49,65	46,96	43,19	40,11	36,74
28	56,89	50,99	48,28	44,46	41,34	37,92
29	58,3	52,34	49,59	45,72	42,56	39,09
30	59,7	53,67	50,89	46,98	43,77	40,26

Равномерно распределенные случайные числа

Таблица D

10	09	73	25	33	76	52	01	35	86	34	67	35	48	76
37	54	20	48	05	64	69	47	42	96	24	80	52	40	37
08	42	26	89	53	19	64	50	93	03	23	20	90	25	60
99	01	90	25	29	09	37	67	07	15	38	31	13	11	65
12	80	79	99	70	80	15	73	61	47	64	03	23	66	53
80	95	90	91	17	39	29	27	49	45	66	06	57	47	17
20	63	61	04	02	00	82	29	16	65	31	06	01	08	05
15	95	33	47	64	35	08	03	36	06	85	26	97	76	02
88	67	67	43	97	04	43	62	76	59	63	57	33	21	35
96	95	11	68	77	12	17	17	68	33	73	79	64	57	53
34	07	27	68	50	36	69	73	61	70	65	81	33	98	85
45	57	18	24	06	35	30	34	26	14	86	79	90	74	39
02	05	16	56	92	68	66	57	48	18	73	05	38	52	47
05	32	54	70	48	90	55	35	75	48	28	46	82	87	09
03	53	96	47	78	35	80	83	42	82	60	93	52	03	44
11	19	92	91	70	98	52	01	77	67	14	90	56	86	07
23	40	30	97	32	11	80	50	54	31	39	80	82	77	32
18	62	38	85	79	83	45	29	96	34	06	28	89	80	83
83	49	12	56	24	86	68	54	02	00	86	50	75	84	01
35	27	38	84	35	99	59	46	73	48	87	51	76	49	69
22	10	94	05	58	60	97	09	34	33	50	50	07	39	98
50	72	56	82	48	29	40	52	42	01	52	77	56	78	61
13	74	67	00	78	18	47	54	06	10	68	71	17	78	17
36	76	66	73	51	90	36	47	64	93	29	60	91	10	62
91	82	60	89	28	93	78	56	13	68	23	47	83	41	13
65	48	11	76	74	17	46	85	09	50	58	04	77	69	74
80	12	43	56	35	17	72	70	80	15	45	31	82	23	74
74	35	09	98	17	77	40	27	72	14	43	23	60	02	10
69	91	62	68	03	66	25	22	91	48	36	93	68	72	03
09	89	32	05	05	14	22	56	85	14	46	42	75	67	88

Содержание