

*Приложение 2.25.1к ОПОП по специальности
09.02.13 Интеграция решений с применением
технологий искусственного интеллекта*

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Магнитогорский государственный технический университет им. Г.И. Носова»

Многопрофильный колледж

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ
ДЛЯ ПРАКТИЧЕСКИХ ЗАНЯТИЙ
УЧЕБНОЙ ДИСЦИПЛИНЫ**

ОП.03. ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

для обучающихся специальности

09.02.13 Интеграция решений с применением технологий искусственного интеллекта

Магнитогорск, 2025

ОДОБРЕНО

Предметно-цикловой комиссией
«Информатики и вычислительной техники»
Председатель Т.Б. Ремез
Протокол № 5.1 от «17» февраля 2025 г.

Методической комиссией МпК

Протокол № 3 от «19» февраля 2025г

Разработчик:

преподаватель отделения №2 «Информационных технологий и транспорта»
Многопрофильного колледжа ФГБОУ ВО «МГТУ им. Г.И. Носова»

Елена Александровна Васильева

Методические указания по выполнению практических работ разработаны на основе рабочей программы учебной дисциплины «Теория вероятностей и математическая статистика».

Содержание практических работ ориентировано на подготовку обучающихся к освоению профессионального модуля программы подготовки специалистов среднего звена по специальности 09.02.13 Интеграция решений с применением технологий искусственного интеллекта и овладению общими компетенциями.

СОДЕРЖАНИЕ

1 ВВЕДЕНИЕ	4
2 МЕТОДИЧЕСКИЕ УКАЗАНИЯ	5
Практическое занятие № 1 Построение пространства элементарных исходов для заданных экспериментов	5
Практическое занятие № 2 Вычисление вероятностей событий на основе классического определения вероятности	9
Практическое занятие № 3 Вычисление условной вероятности и проверка независимости событий	11
Практическое занятие № 4 Вычисление математического ожидания и дисперсии дискретных случайных величин. Построение и анализ биномиального и нормального распределений	15
Практическое занятие № 5 Применение распределения Пуассона для моделирования редких событий	21
Практическое занятие № 6 Демонстрация центральной предельной теоремы на основе генерации выборок и построения гистограмм	23
Практическое занятие № 7 Применение центральной предельной теоремы для оценки распределения сумм случайных величин	23
Практическое занятие № 8 Моделирование закона больших чисел на основе последовательных испытаний. Оценка среднего значения выборки и математического ожидания с помощью закона больших чисел	26
Практическое занятие № 9 Построение точечных оценок параметров для различных распределений. Оценка доверительных интервалов для среднего значения и дисперсии.....	26
Практическое занятие №10 Проверка гипотез с использованием критерия Стьюдента для двух выборок. Применение критерия χ^2 для проверки гипотез о независимости признаков. Оценка ошибок первого и второго рода при тестировании гипотез.....	32
Практическое занятие №11 Вычисление коэффициента корреляции Пирсона для анализа зависимостей между признаками. Построение корреляционной матрицы для многомерных данных и её интерпретация.....	35
Практическое занятие №12 Вычисление ковариации и её применение для оценки совместной изменчивости признаков	37
Практическое занятие №13 Построение линейной регрессионной модели на основе экспериментальных данных. Интерпретация коэффициентов линейной регрессии и оценка её качества	39
Практическое занятие №14 Применение нелинейной регрессии для аппроксимации данных	42
Практическое занятие №15 Проведение однофакторного дисперсионного анализа (ANOVA) для проверки различий между группами.....	45
Практическое занятие №16 Применение дисперсионного анализа для оценки влияния различных факторов на результаты экспериментов.....	47

1 ВВЕДЕНИЕ

Важную часть теоретической и профессиональной практической подготовки студентов составляют практические занятия.

Состав и содержание практических занятий направлены на реализацию Федерального государственного образовательного стандарта среднего профессионального образования.

Ведущей дидактической целью практических занятий является формирование учебных практических умений (умений решать задачи по математике), необходимых в последующей учебной деятельности.

В соответствии с рабочей программой учебной дисциплины «Теория вероятностей и математическая статистика» предусмотрено проведение практических занятий. В рамках практического занятия обучающиеся могут выполнять одну или несколько практических работ.

В результате их выполнения, обучающийся должен:

уметь:

У 1. Применять стандартные методы и модели к решению вероятностных и статистических задач

У 2. Использовать расчетные формулы, таблицы, графики при решении статистических задач

У 3. Применять современные пакеты прикладных программ многомерного статистического анализа

Уо 01.02 анализировать задачу и/или проблему и выделять её составные части;

Уо 01.03 определять этапы решения задачи;

Уо 01.05 составлять план действий;

Уо 01.08 реализовать составленный план;

Уо 01.09 оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);

Содержание практических занятий ориентировано на подготовку обучающихся к освоению профессионального модуля программы подготовки специалистов среднего звена по специальности. А также формированию **общих компетенций**:

ОК 1. Выбирать способы решения задач профессиональной деятельности, применительно к различным контекстам.

Выполнение обучающихся практических работ по учебной дисциплине «Теория вероятностей и математическая статистика» направлено на:

- обобщение, систематизацию, углубление, закрепление, развитие и детализацию полученных теоретических знаний по конкретным темам учебной дисциплины;

- формирование умений применять полученные знания на практике, реализацию единства интеллектуальной и практической деятельности;

- развитие интеллектуальных умений у будущих специалистов: аналитических, проектировочных, конструктивных и др.;

- выработку при решении поставленных задач профессионально значимых качеств, таких как самостоятельность, ответственность, точность, творческая инициатива.

Практические занятия проводятся после соответствующей темы, которая обеспечивает наличие знаний, необходимых для ее выполнения.

2 МЕТОДИЧЕСКИЕ УКАЗАНИЯ

Практическое занятие № 1

Построение пространства элементарных исходов для заданных экспериментов

Цель: сформировать умение решать задачи с использованием формул комбинаторики: перестановок, размещений и сочетаний.

Выполнив работу, Вы будете:

уметь:

- применять стандартные методы и модели к решению комбинаторных задач;
- решать задачи с использованием формул комбинаторики;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Внимательно прочитайте теоретические сведения.

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Краткие теоретические сведения

Комбинаторика — один из разделов дискретной математики, изучающий способы подсчета числа элементов в конечных множествах, который приобрел большое значение в связи с использованием его в теории вероятностей, математической логике, теории чисел, вычислительной технике, кибернетике.

Множества без повторений

Множества элементов, состоящие из одних и тех же различных элементов и отличающиеся друг от друга только их порядком, называются **перестановками** этих элементов: $P_n = n!$, где $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$, $0! = 1$.

Размещениями называют множества, составленные из n различных элементов по m элементов, которые отличаются либо составом элементов, либо их порядком: $A_n^m = \frac{n!}{(n-m)!}$.

Сочетаниями из n различных элементов по m называют множества, содержащие m элементов из числа n заданных, и которые отличаются хотя бы одним элементом:

$$C_n^m = \frac{n!}{m!(n-m)!}, \text{ где } C_n^0 = 1, C_n^n = 1, C_n^1 = n.$$



Множества с повторениями

Перестановки с повторениями

$$\overline{P}_n(n_1, n_2, \dots, n_k) = \frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_k!}$$

Размещения с повторениями

$$\overline{A}_n^m = n^m$$

Сочетания с повторениями

$$\overline{C}_n^m = \frac{(n+m-1)!}{m!(n-1)!} = C_{n+m-1}^m$$

Правила сложения и умножения

Правило сложения. Если одно действие можно выполнить m способами, а другое n способами, и они взаимно исключают друг друга, то выполнить одно любое из этих действий можно $(n + m)$ способами.

Правило произведения. Если объект A можно выбрать из совокупности объектов m способами и после каждого такого выбора объект B можно выбрать n способами, то пара объектов (A, B) в указанном порядке может быть выбрана $m \cdot n$ способами.

Примеры решения задач

Размещения без повторений

Задача 1. Сколько можно составить телефонных номеров из 6 цифр каждый, так чтобы все цифры были различны?

Решение. Это пример задачи на размещение без повторений. Размещаются здесь 10 цифр по 6. А варианты, при которых одинаковые цифры стоят в разном порядке считаются разными.

Значит, ответ на задачу будет:

$$A_{10}^6 = \frac{10!}{(10-6)!} = 151200$$

Задача 2. Сколькими способами 4 юноши могут пригласить четырех из шести девушек на танец?

Решение. Два юноши не могут одновременно пригласить одну и ту же девушку. И варианты, при которых одни и те же девушки танцуют с разными юношами, считаются разными, поэтому:

$$A_6^4 = \frac{6!}{(6-4)!} = \frac{720}{2} = 360$$

Возможно 360 вариантов.

Перестановки без повторений

Задача 3. Сколько различных шестизначных чисел можно составить из цифр 0, 1, 2, 3, 4, 5, если цифры в числе не повторяются?

Решение.

1) Найдем количество всех перестановок из этих цифр: $P_6=6!=720$.

2) 0 не может стоять впереди числа, поэтому от этого числа необходимо отнять количество перестановок, при котором 0 стоит впереди. А это $P_5=5!=120$.

$$P_6 - P_5 = 720 - 120 = 600.$$

Сочетания без повторений

Задача 4. Сколько трехкнопочных комбинаций существует на кодовом замке (все три кнопки нажимаются одновременно), если на нем всего 10 цифр.

Решение. Так как кнопки нажимаются одновременно, то выбор этих трех кнопок – сочетание. Отсюда возможно $C_{10}^3 = \frac{10!}{(10-3)! \cdot 3!} = \frac{8 \cdot 9 \cdot 10}{6} = 120$ вариантов.

Задача 5. У одного человека 7 книг по математике, а у второго – 9. Сколькими способами они могут обменять друг у друга две книги на две книги.

Решение. Так как порядок следования книг не имеет значения, то выбор двух книг – сочетание. Первый человек может выбрать 2 книги $C_7^2 = \frac{7!}{(7-2)! \cdot 2!} = \frac{6 \cdot 7}{2} = 21$ способами.

Второй человек может выбрать 2 книги $C_9^2 = \frac{9!}{(9-2)! \cdot 2!} = \frac{8 \cdot 9}{2} = 36$ способами. Значит всего по правилу произведения возможно $21 \cdot 36 = 756$ вариантов.

Задача 6. При игре в домино 4 игрока делят поровну 28 костей. Сколькими способами они могут это сделать?

Решение. Первый игрок делает выбор из 28 костей. Второй из $28-7=21$ кости, третий из 14, а четвертый игрок забирает оставшиеся кости. Следовательно, возможно $C_{28}^7 \cdot C_{21}^7 \cdot C_{14}^7$.

Размещения и сочетания с повторениями

Задача 7. Сколько трехзначных чисел можно составить из цифр 1, 2, 3, 4, 5?

Решение. Так как порядок цифр в числе существенен, цифры могут повторяться, то это будут размещения с повторениями из пяти элементов по три, а их число равно $\overline{A}_5^3 = 5^3 = 125$.

Задача 8. В кондитерском магазине продавались 4 сорта пирожных: эклеры, песочные, наполеоны и слоеные. Сколькими способами можно купить 7 пирожных.

Решение. Покупка не зависит от того, в каком порядке укладывают купленные пирожные в коробку. Покупки будут различными, если они отличаются количеством купленных пирожных хотя бы одного сорта. Следовательно, количество различных покупок равно числу сочетаний четырех видов пирожных по семь: $\overline{C}_4^7 = \frac{(7+4-1)!}{7!(4-1)!} = \frac{10!}{7! \cdot 3!} = 120$.

Задача 9. Обезьяну посадили за пишущую машинку с 45 клавишами, определить число попыток, необходимых для того, чтобы она наверняка напечатала первую строку романа Л.Н. Толстого «Анна Каренина», если строка содержит 52 знака и повторений не будет?

Решение. Порядок букв имеет значение. Буквы могут повторяться. Значит, всего есть $\overline{A}_{45}^{52} = 45^{52}$ вариантов.

Перестановки с повторениями

Задача 10. Сколькими способами можно переставить буквы слова «ананас»?

Решение. Всего букв 6. Из них одинаковы: $n_1(a)=3$, $n_2(n)=2$, $n_3(s)=1$. Следовательно,

число различных перестановок равно $P_6(3,2,1) = \frac{6!}{3! \cdot 2! \cdot 1!} = 60$.

Задачи для самостоятельного решения

1. Сколько перестановок можно сделать из букв слова «Миссисипи». Ответ: 2520
2. Имеется пять различных стульев и семь рулонов обивочной ткани различных цветов. Сколькими способами можно осуществить обивку стульев. Ответ: 16807
3. На памятные сувениры в «Поле Чудес» спонсоры предлагают кофеварки, утюги, телефонные аппараты, духи. Сколькими способами 9 участников игры могут получить эти сувениры? Сколькими способами могут быть выбраны 9 предметов для участников игры? Ответ: 4^9 , 220
4. Сколькими способами можно расставить на шахматной доске 8 ладей так, чтобы на одна из них не могла бить другую? Ответ: 40320
5. Сколько может быть случая выбора 2 карандашей и 3 ручек из пяти различных карандашей и шести различных ручек? Ответ: 200
6. В течение 30 дней сентября было 12 дождливых дней, 8 ветреных, 4 холодных, 5 дождливых и ветреных, 3 дождливых и холодных, а один день был и дождливым, и ветреным, и холодным. В течение скольких дней в сентябре стояла хорошая погода. Ответ: 15
7. На ферме есть 20 овец и 24 свиньи. Сколькими способами можно выбрать одну овцу и одну свинью? Если такой выбор уже сделан, сколькими способами можно сделать его еще раз? Ответ: 480, 437
8. Сколькими способами можно выбрать гласную и согласную буквы из слова «здание»? Ответ: 9
9. Сколько существует четных пятизначных чисел, начинающихся нечетной цифрой? Ответ: 25000
10. В книжный магазин поступили романы Ф. Купера «Прерия», «Зверобой», «Шпион», «Пионеры», «Следопыт» по одинаковой цене. Сколькими способами библиотека может закупить 17 книг на выбранный чек? Ответ: 2985

Форма представления результата:

Решение задач, оформленное в тетради.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 2

Вычисление вероятностей событий на основе классического определения вероятности

Цель: сформировать умение решать задачи на вычисление вероятности с использованием формул комбинаторики.

Выполнив работу, Вы будете:

уметь:

- применять стандартные методы и модели к решению вероятностных задач;
- решать задачи с использованием формул комбинаторики;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Краткие теоретические сведения

Вероятность события – это численная мера объективной возможности его появления.

Если имеется полная группа попарно несовместных и равновозможных событий, то вероятность $P(A)$ наступления события A вычисляется как отношение числа исходов, благоприятствующих наступлению события (m), к числу всех исходов испытания (n):

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

Свойства вероятности:

- 1) Вероятность достоверного события равна 1.
- 2) Вероятность невозможного события равна 0.
- 3) Вероятность случайного события удовлетворяет двойному неравенству $0 < P(A) < 1$.
- 4) Вероятность любого события удовлетворяет двойному неравенству $0 \leq P(A) \leq 1$.

Примеры решения задач

Задача 1. Найти вероятность того, что при бросании игральной кости выпадет число, кратное трём (событие A).

Решение: Число всех элементарных событий $n = 6$, число элементарных событий, благоприятствующих данному событию A , $m = 2$ (это выпадение 3 или 6), значит,

$$P(A) = \frac{m}{n} = \frac{1}{3}.$$

В следующих двух примерах мы используем комбинаторные понятия (сочетания и перестановки), разобранные выше (в практическом занятии № 1).

Задача 2. В урне находится 7 шаров, из которых 4 белых и 3 черных. Из урны наугад вынимается три шара. Найти вероятность того, что два шара будут белыми, а один – чёрным.

Решение: Число всех элементарных событий подсчитаем по формуле сочетаний $n = C_7^3 = 35$, число элементарных событий, благоприятствующих данному событию A , по формуле сочетаний, применив правило произведения, $m = C_4^2 C_3^1 = 18$, значит,

$$P(A) = \frac{m}{n} = \frac{C_4^2 C_3^1}{C_7^3} = \frac{18}{35}.$$

Задача 3. На книжной полке случайно размещаются 8 книг, из которых 5 – по математике и 3 – по физике. Найти вероятность того, что слева будут все книги по математике, а справа – все книги по физике.

Решение: Число всех элементарных событий $n = P_8 = 8!$, число элементарных событий, благоприятствующих данному событию A , $m = P_5 P_3 = 5!3!$, значит, $P(A) = \frac{m}{n} = \frac{5!3!}{8!} = \frac{1}{56}$.

Задачи для самостоятельного решения

Какова вероятность того, что наудачу выбранное двузначное число не содержит ни одной двойки?

2. Отряд учащихся из 25 человек участвует в военизированной игре. В отряде 5 следопытов и 4 связиста. В разведку надо направить четырех человек. Какова вероятность того, что в разведгруппу будут включены 2 связиста и 2 следопыта, если включение в разведгруппу равновероятно для любого ученика?

3. На карточках написаны целые числа от 1 до 15 включительно. Наудачу извлекаются две карточки. Какова вероятность того, что сумма чисел, написанных на этих карточках, равна десяти?

4. Для дежурства на вечере путем жеребьевки выделяются 5 человек. Вечер проводит комиссия, в составе которой 10 юношей и 2 девушки. Найдите вероятность того, что в число дежурных войдут обе девушки.

5. Имеется 6 билетов в театр, из которых 4 билета на места первого ряда. Какова вероятность того, что из трех наудачу выбранных билетов два окажутся на места первого ряда?

6. Билет в партер стоит 50 коп., на бельэтаж — 40 коп. и на ярус — 30 коп. Найдите вероятность того, что взятые наудачу два билета стоят вместе не дороже 80 коп.

7. На один ряд из семи мест случайным образом рассаживаются 7 учеников. Найдите вероятность того, что 3 определенных ученика окажутся рядом.

8. Из букв слова событие, составленного с помощью разрезной азбуки, извлекаются наудачу и складываются друг за другом в порядке их извлечения 3 карточки (буквы). Какова вероятность получить при этом слово быт?

9. Из пяти видов открыток, имеющихся в автомате, наудачу выбираются 3 открытки. Какова вероятность того, что все отобранные открытки будут разные?

10. Во время спортивной игры по команде ведущего «становись!» 10 учеников в случайном порядке образовали строй в одну шеренгу. Какова вероятность того, что ученики А и В окажутся отделенными друг от друга тремя учениками?

Форма представления результата:

Решение задач, оформленное в тетради.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;

- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 3

Вычисление условной вероятности и проверка независимости событий

Цель: сформировать умение решать задачи на вычисление вероятностей сложных событий.

Выполнив работу, Вы будете:

уметь:

- применять стандартные методы и модели к решению вероятностных задач;
- решать задачи с использованием теорем вероятности;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Краткие теоретические сведения

Теорема сложения вероятностей несовместных событий. Вероятность появления одного из двух несовместных событий, безразлично какого, равна сумме вероятностей этих событий:

$$P(A + B) = P(A) + P(B)$$

Следствие. Вероятность появления одного из нескольких попарно несовместных событий, безразлично какого, равна сумме вероятностей этих событий:

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

Сумма вероятностей попарно несовместных событий $A_1, A_2, \dots, A_n$, образующих полную группу, равна единице:

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1$$

Сумма вероятностей противоположных событий равна единице:

$$P(A) + P(\bar{A}) = 1$$

Теорема сложения вероятностей совместных событий. Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного наступления:

$$P(A + B) = P(A) + P(B) - P(A \cdot B)$$

Условной вероятностью $P(B/A)$ называется вероятность события B , вычисленная в предположении, что событие A уже наступило.

Теорема умножения вероятностей. Вероятность совместного появления двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие уже наступило:

$$P(A \cdot B) = P(A) \cdot P(B/A)$$

Два события называются *независимыми*, если появление любого из них не изменяет вероятность появления другого: $P(B) = P(B/A)$ или $P(A) = P(A/B)$.

Теорема умножения вероятностей. Вероятность совместного появления двух независимых событий равна произведению их вероятностей:

$$P(A \cdot B) = P(A) \cdot P(B)$$

Вероятность совместного наступления конечного числа событий равна произведению вероятности одного из них на условные вероятности всех остальных, причем условная вероятность каждого последующего события вычисляется в предположении, что все предыдущие уже наступили:

$$P(A_1 A_2 \dots A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 A_2) \cdot \dots \cdot P(A_n/A_1 \dots A_{n-1})$$

Вероятность совместного появления нескольких событий, независимых в совокупности, равна произведению вероятностей этих событий:

$$P(A_1 A_2 \dots A_n) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot \dots \cdot P(A_n)$$

Вероятность появления хотя бы одного из событий $A_1, A_2, A_3, \dots, A_n$ независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий:

$$P(A_1 + A_2 + \dots + A_n) = 1 - P(\bar{A}_1) \cdot P(\bar{A}_2) \cdot \dots \cdot P(\bar{A}_n)$$

Сведем все теоремы сложения и умножения вероятностей в таблицу 1.1, позволяющую, в зависимости от вида событий, выбрать формулу для нахождения их вероятностей.

Следствием двух основных теорем теории вероятностей (теоремы сложения и умножения) являются *формула полной вероятности* и *формула Байеса*.

События, в условиях которых только и может появиться событие A , называются *гипотезами* и обозначают $H_1, H_2, \dots, H_n$.

Формула полной вероятности. Вероятность события A , которое может наступить только при условии появления одного из событий $H_1, H_2, H_3, \dots, H_n$, образующих полную группу попарно несовместных событий, равна сумме произведений вероятностей каждого из событий $H_1, H_2, H_3, \dots, H_n$ на соответствующую условную вероятность события A :

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n)$$

или

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i)$$

Рассмотрим событие A , которое может наступить при условии появления одного из несовместных событий $H_1, H_2, H_3, \dots, H_n$, образующих полную группу событий. Если событие A уже произошло, то вероятность событий может быть переоценена по *формуле Байеса* (формуле вероятности гипотез):

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(A)}$$

Примеры решения задач

Задача 1. В урне находятся 6 белых и 5 чёрных шаров. Найти вероятность того, что из пяти случайно взятых шаров

- а) не больше двух белых;
- б) хотя бы один белый.

Решение. а). Событие A – не больше двух белых шаров – равно сумме трёх попарно несовместимых событий A_0, A_1, A_2 – белых шаров соответственно ни одного, один, два: $A = A_0 + A_1 + A_2$. Значит, по теореме сложения для несовместных событий

$$P(A) = P(A_0) + P(A_1) + P(A_2) = \frac{C_6^0 C_5^5}{C_{11}^5} + \frac{C_6^1 C_5^4}{C_{11}^5} + \frac{C_6^2 C_5^3}{C_{11}^5} =$$

$$\frac{1}{462} + \frac{30}{462} + \frac{150}{462} = \frac{181}{462} \quad (\text{по поводу вычисления данных вероятностей см. аналогичную}$$

задачу в примере 12).

б). Для нахождения вероятности события B – хотя бы один белый шар – неудобно использовать теорему сложения, т. к. пришлось бы подсчитывать вероятности пяти событий. Поэтому воспользуемся теоремой 5, которая сводит вычисление $P(B)$ к вычислению вероятности противоположного события $\bar{B}$ – ни одного белого шара, т. е. все шары чёрные:

$$P(B) = 1 - P(\bar{B}) = 1 - \frac{C_5^5}{C_{11}^5} = 1 - \frac{1}{462} = \frac{461}{462}.$$

Задача 2. Два стрелка стреляют по мишени. Вероятность попадания первого стрелка равна 0,6 второго – 0,8. Найти вероятности попадания

- а) обоих стрелков;
- б) ровно одного стрелка;
- в) хотя бы одного стрелка.

Решение. Обозначим через A и B события – попал в мишень первый и, соответственно, второй стрелок. Заметим, что эти события *независимы*. Заметим также, что и противоположные события (промахи стрелков) – независимы. С помощью известных нам операций над событиями выразим через A и B события, вероятности которых являются искомыми.

а). Очевидно, интересующее нас событие есть AB . Тогда по теореме умножения для независимых событий $P(AB) = P(A)P(B) = 0,6 \cdot 0,8 = 0,48$.

б). Интересующее нас событие есть $A\bar{B} + \bar{A}B$. Слагаемые в этой сумме – несовместные события, а сомножители в каждом произведении – независимые. Тогда по теоремам сложения для несовместных событий и умножения для независимых событий $P(A\bar{B} + \bar{A}B) = P(A)P(\bar{B}) + P(\bar{A})P(B) = 0,6 \cdot 0,2 + 0,4 \cdot 0,8 = 0,44$.

в). Интересующее нас событие есть $A+B$. По теореме сложения (для двух произвольных событий)

$$P(A+B) = P(A) + P(B) - P(AB) = P(A) + P(B) - P(A)P(B) = 0,6 + 0,8 - 0,6 \cdot 0,8 = 0,92.$$

Заметим, что противоположным событием к $A+B$ является $\bar{A}\bar{B}$ (ни один не попал). Тогда искомую вероятность можно найти и через противоположное событие:

$$P(A+B) = 1 - P(\bar{A}\bar{B}) = 1 - P(\bar{A})P(\bar{B}) = 1 - 0,4 \cdot 0,2 = 0,92.$$

Задача 3. Телеграфное сообщение состоит из сигналов "точка" и "тире". Статистические свойства помех таковы, что искажаются в среднем 2/5 сигналов "точка" и 1/3 сигналов "тире". Известно, что среди передаваемых сигналов "точка" и "тире"

встречаются в соотношении 5:3. Определить вероятность того, что переданный сигнал искажён.

Решение: Рассмотрим следующие события: A – сигнал искажён (вероятность которого является искомой), H_1 – передан сигнал "точка", H_2 – передан сигнал "тире" (две гипотезы, образующие полную группу, т. к. они несовместны и одна из них обязательно имеет место).

По условию $P(H_1):P(H_2)=5:3$ и $P(H_1)+P(H_2)=1$, поэтому $P(H_1)=\frac{5}{8}$ и $P(H_2)=\frac{3}{8}$.

Известные вероятности того, что сигналы "точка" и "тире" искажаются, выступают как условные вероятности интересующего нас события при справедливости гипотез: $P_{H_1}(A)=2/5$, $P_{H_2}(A)=1/3$. Вероятность события A находим по формуле полной вероятности:

$$P(A)=\frac{5}{8}\cdot\frac{2}{5}+\frac{3}{8}\cdot\frac{1}{3}=\frac{3}{8}.$$

В условиях применения формулы полной вероятности рассмотрим следующую задачу: чему равна условная вероятность гипотезы H_i , если событие A произошло? Ответ даётся следующей *формулой Байеса*:

$$P_A(H_i)=\frac{P(H_i)P_{H_i}(A)}{P(A)}.$$

Продолжение задачи 3. Пусть переданный сигнал искажён. Чему равна вероятность того, что была передана "точка"?

Решение: Искомая вероятность вычисляется по формуле Байеса:

$$P_A(H_1)=\frac{P(H_1)P_{H_1}(A)}{P(A)}=\frac{\frac{5}{8}\cdot\frac{2}{5}}{\frac{3}{8}}=\frac{2}{3}.$$

Задачи для самостоятельного решения

1. Вероятность того, что стрелок при одном выстреле выбивает 10 очков, равна 0,4; 9 очков – 0,3 и 8 или меньше очков – 0,3. Найти вероятность того, что при одном выстреле стрелок выбьет не менее 9 очков.

2. В клуб принесли в корзине 9 рыжих и 11 серых котят. Сначала вынимают одного котенка, затем другого, не возвращая первого обратно. Какова вероятность того, что котята разного цвета?

3. В одном ящике 6 синих и 11 зеленых шаров, а в другом – 7 синих и 9 зеленых шаров. Из каждого ящика взяли по одному шару. Какова вероятность того, что оба шара синие?

4. В коробке 4 белых и 5 черных футболок. Наугад вытаскивают две футболки. Найти вероятность того, что одна футболка белая, другая – черная?

5. В урне находятся 10 шаров: 3 белых и 7 черных. Извлекают один шар, а затем другой. Найти вероятность того, что первый шар черный, а второй – синий.

6. В специализированную больницу поступает в среднем 50% больных с заболеванием А, 30% - с заболеванием В, 20% - с заболеванием С. Вероятность полного излечения болезни А равна 0,7, для болезней В и С эти вероятности равны соответственно 0,8 и 0,9. Какой процент пациентов, полностью излечившихся по окончании курса лечения?

7. На трех дочерей Нину, Еву и Айгуль в семье возложена обязанность мыть посуду. Нина выполняет 40% всей работы, а остальные 60% работы Ева и Айгуль делят поровну. С вероятностью 0,02 Нина может разбить по крайней мере одну тарелку; для Евы и Айгуль эта вероятность равна соответственно 0,03 и 0,04. Родители не знают, кто мыл посуду вечером, но они слышали звон разбитой посуды. Чья очередь мыть посуду в этот вечер наиболее вероятна?

8. Один властелин, которому его звездочет наскучил своими ложными предсказаниями, решил казнить его. Однако, будучи очень добрым повелителем, он решил дать звездочету шанс и предложить ему распределить по двум урнам четыре шара: два белых и два черных. Палач выбирает одну из урн и извлекает из нее один шар. Если шар будет черным, то звездочета казнят, в противном случае его жизнь будет спасена. Каким образом звездочет должен разместить шары в урнах, чтобы обеспечить себе максимальную вероятность спасения?

9. Устройство состоит из трех элементов, работающих независимо. Вероятности безотказной работы для них равны соответственно 0.6, 0.7, 0.8. Найти вероятность того, что отказал один элемент, а два другие - исправны.

10. В электрическую цепь последовательно включены три элемента, работающие независимо. Вероятности отказа каждого из элементов равны соответственно 0.1, 0.15, 0.2. Найти вероятность того, что тока в цепи не будет.

Форма представления результата:

Решение задач, оформленное в тетради.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 4

Вычисление математического ожидания и дисперсии дискретных случайных величин.

Построение и анализ биномиального и нормального распределений

Цель работы: изучение законов распределения дискретной случайной величины; научиться решать задачи на нахождение числовых характеристик случайной величины.

Выполнив работу, Вы будете:

уметь:

- решать задачи на нахождение числовых характеристик случайной величины;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;

- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Краткие теоретические сведения

Рядом *распределения ДСВ* называют совокупность всех возможных значений x_i и соответствующих им вероятностей p_i . Удобно ряд распределения представлять в виде таблицы.

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

Ряд распределения обладает следующим свойством: сумма всех вероятностей случайной величины равна 1, т.к. события x_i образуют полную группу.

$$\sum_{i=1}^n p_i = p_1 + p_2 + \dots + p_n = 1 \quad (1)$$

Многоугольником (полигоном) распределения называют графическое представление ряда распределения ДСВ (рис.1).

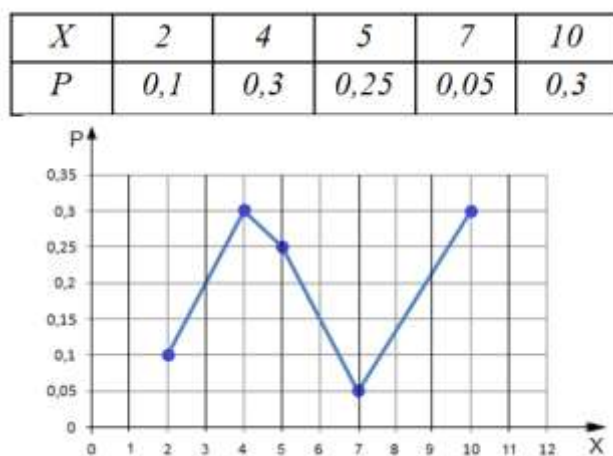


Рисунок 1 – Многоугольник распределения для данного ряда распределения ДСВ

Функцией распределения случайной величины называют вероятность того, что случайная величина X в результате испытания примет значение, меньшее x , т.е. $F(x) = P(X < x)$.

X	x_1	x_2	x_3	...	x_n
P	p_1	p_2	p_3	...	p_n

$$F(x) = \begin{cases} 0 & \text{при } x < x_1 \\ p_1 & \text{при } x_1 < x < x_2 \\ p_1 + p_2 & \text{при } x_2 < x < x_3 \\ p_1 + p_2 + p_3 & \text{при } x_3 < x < x_4 \\ \dots & \\ 1 & \text{при } x > x_n \end{cases} \quad (2)$$

Функция распределения случайной величины является неубывающей функцией, значения которой лежат в интервале $[0; 1]$. График функции распределения произвольной ДСВ представляет собой «возрастающую ступеньку».

Модой $M_o(X)$ ДСВ X называется такое значение случайной величины, вероятность которого наибольшая.

Медианой $M_e(X)$ ДСВ X называется среднее по положению в пространстве событий значение случайной величины.

Математическое ожидание $M(X)$ ДСВ X приближенно равно среднему значению случайной величины.

Дисперсия $D(X)$ ДСВ X характеризует разброс случайной величины относительно среднего значения (математического ожидания).

Среднее квадратическое отклонение $\sigma(X)$ ДСВ X характеризует ширину диапазона значений случайной величины.

Математическое ожидание $M(X)$ ДСВ X равно сумме произведений всех возможных значений величины X на соответствующие вероятности

$$M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n$$

или

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

Дисперсия $D(X)$ ДСВ X равна разности между математическим ожиданием квадрата величины X и квадратом ее математического ожидания

$$D(X) = M(X^2) - (M(X))^2$$

Среднее квадратическое отклонение $\sigma(X)$ ДСВ X равно корню квадратному из ее дисперсии

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

Примеры решения задач

Задача 1. Задают ли закон распределения ДСВ каждая из следующих таблиц?

а)

X	0	1	2	3	4
P	0,05	0,15	0,20	0,25	0,35

б)

X	5	6	7	8	9
P	0,1	0,2	0,3	0,4	0,15

Решение. а) Данная таблица задает закон распределения ДСВ, поскольку выполняется равенство (1): $0,05 + 0,15 + 0,20 + 0,25 + 0,35 = 1$.

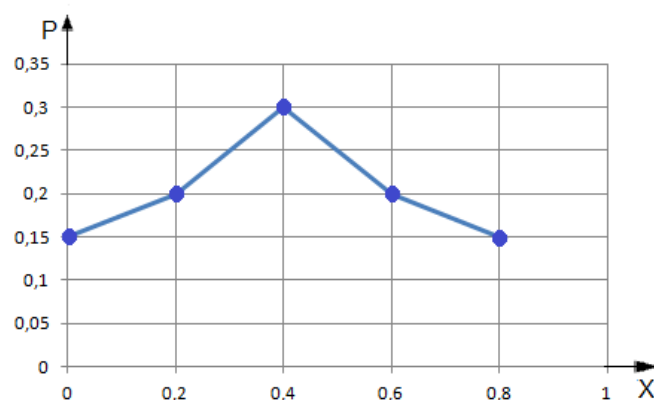
б) Данная таблица не задает закон распределения ДСВ, поскольку не выполняется равенство (1): $0,1 + 0,2 + 0,3 + 0,4 + 0,15 = 1,15 \neq 1$.

Задача 2. ДСВ X имеет закон распределения, заданный таблицей. Чему равна вероятность p_4 ? Постройте многоугольник распределения.

X	0	0,2	0,4	0,6	0,8
P	0,15	0,2	0,3	p_4	0,15

Решение. Поскольку должно выполняться равенство (1), то

$$p_4 = 1 - (0,15 + 0,20 + 0,3 + 0,15) = 0,2$$



Задача 3. В партии из 10 деталей имеется 8 стандартных. Из этой партии наудачу взято 2 детали. Найдите закон распределения ДСВ, равной числу стандартных деталей в выборке.

Решение. Пусть случайная величина X – число стандартных деталей в выборке. Случайная величина X может принимать значения $k=\{0,1,2\}$. Найдём вероятности этих значений.

X	0	1	2
P	p_1	p_2	p_3

$$p_1 = P(X = 0) = \frac{m}{n}$$

Вычислим число всех исходов (n): $C_{10}^2 = \frac{10!}{2!8!} = \frac{9 \cdot 10}{2} = 45$
(из 10 деталей берём 2)

Вычислим число благоприятствующих исходов (m): $C_8^0 \cdot C_2^2 = 1 \cdot 1 = 1$
(из 2 выбранных деталей 0 стандартных)

$$p_1 = P(X = 0) = \frac{1}{45}$$

$$p_2 = P(X = 1) = \frac{m}{n}$$

Вычислим число всех исходов (n): $C_{10}^2 = 45$
(из 10 деталей берём 2)

Вычислим число благоприятствующих исходов (m): $C_8^1 \cdot C_2^1 = 8 \cdot 2 = 16$
(из 2 выбранных деталей 1 стандартная)

$$p_2 = P(X = 1) = \frac{16}{45}$$

$$p_3 = P(X = 2) = \frac{m}{n}$$

Вычислим число всех исходов (n): $C_{10}^2 = 45$
(из 10 деталей берём 2)

Вычислим число благоприятствующих исходов (m): $C_8^2 \cdot C_2^0 = \frac{8!}{2!6!} \cdot 1 = \frac{7 \cdot 8}{2} = 28$
(из 2 выбранных деталей 2 стандартные)

$$p_3 = P(X = 2) = \frac{28}{45}$$

X	0	1	2
P	1/45	16/45	28/45

Данная таблица задает закон распределения ДСВ, поскольку выполняется равенство (1):

$$\sum_{i=1}^3 p_i = p_1 + p_2 + p_3 = \frac{1}{45} + \frac{16}{45} + \frac{28}{45} = 1$$

Задача 4. ДСВ X имеет закон распределения, заданный таблицей. Найти функцию распределения этой ДСВ и построить ее график.

X	2	4	5	7	10
P	0,1	0,3	0,25	0,05	0,3

Решение. Для построения функции распределения $F(x)$ дискретной случайной величины X воспользуемся формулой (2):

при $x < 2$ $F(x) = 0$;

при $2 < x < 4$ $F(x) = p_1 = 0,1$;

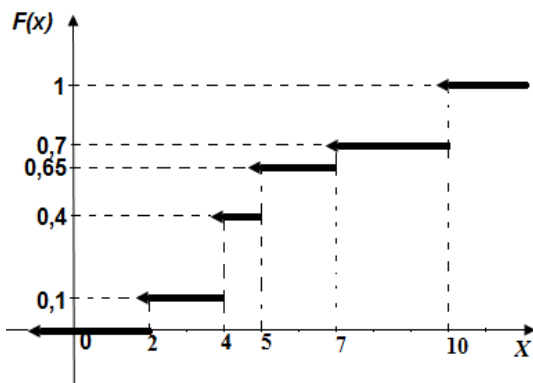
при $4 < x < 5$ $F(x) = p_1 + p_2 = 0,1 + 0,3 = 0,4$;

при $5 < x < 7$ $F(x) = p_1 + p_2 + p_3 = 0,1 + 0,3 + 0,25 = 0,65$;

при $7 < x < 10$ $F(x) = p_1 + p_2 + p_3 + p_4 = 0,1 + 0,3 + 0,25 + 0,05 = 0,7$;

при $x > 10$ $F(x) = p_1 + p_2 + p_3 + p_4 + p_5 = 0,1 + 0,3 + 0,25 + 0,05 + 0,3 = 1$

$$F(x) = \begin{cases} 0 & \text{при } x < 2 \\ 0,1 & \text{при } 2 < x < 4 \\ 0,4 & \text{при } 4 < x < 5 \\ 0,65 & \text{при } 5 < x < 7 \\ 0,7 & \text{при } 7 < x < 10 \\ 1 & \text{при } x > 10 \end{cases}$$



Задача 5. ДСВ X задана законом распределения. Вычислить значения моды, медианы, математического ожидания, дисперсии и среднего квадратического отклонения.

X	2	4	5	7	10
P	0,1	0,3	0,25	0,05	0,3

Решение.

Моды будет две, т.к. наибольшее значение вероятности соответствует двум значениям случайной величины $M_o(X)=4$ и $M_o(X)=10$.

Медиана, то значение случайной величины, которое находится посередине $M_e(X)=5$.

Математическое ожидание вычислим по формуле $M(X) = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n$
 $M(X) = 2 \cdot 0,1 + 4 \cdot 0,3 + 5 \cdot 0,25 + 7 \cdot 0,05 + 10 \cdot 0,3 = 6$

Дисперсию вычислим по формуле $D(X) = M(X^2) - (M(X))^2$

$M(X^2) = 2^2 \cdot 0,1 + 4^2 \cdot 0,3 + 5^2 \cdot 0,25 + 7^2 \cdot 0,05 + 10^2 \cdot 0,3 = 43,90$

$D(X) = 43,90 - (6)^2 = 7,90$

Среднее квадратическое вычислим по формуле $\sigma(X) = \sqrt{D(X)}$

$$\sigma(X) = \sqrt{7,90} \approx 2,81$$

Задачи для самостоятельного решения

1. Найти математическое ожидание, дисперсию, среднее квадратическое отклонение, моду и медиану дискретной случайной величины, представленной законом распределения:

X	10	13	17	20	25
P	0,4	0,3	0,1	0,15	0,05

Составить функцию распределения $F(x)$.

2. Найти математическое ожидание, дисперсию, среднее квадратическое отклонение, моду и медиану дискретной случайной величины, представленной законом распределения:

X	8	14	17	20	23
P	0,2	0,1	0,2	0,4	0,1

Составить функцию распределения $F(x)$.

3. Найти математическое ожидание, дисперсию, среднее квадратическое отклонение, моду и медиану дискретной случайной величины, представленной законом распределения:

X	14	15	17	25	26
P	0,1	0,35	0,3	0,2	0,05

Составить функцию распределения $F(x)$.

4. Найти математическое ожидание, дисперсию, среднее квадратическое отклонение, моду и медиану дискретной случайной величины, представленной законом распределения:

X	16	20	25	30	35
P	0,2	0,15	0,15	0,3	0,2

Составить функцию распределения $F(x)$.

Форма представления результата:

Решение задач, оформленное в тетради.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 5

Применение распределения Пуассона для моделирования редких событий

Цель работы: изучение законов распределения дискретной случайной величины; научиться решать задачи на нахождение числовых характеристик случайной величины.

Выполнив работу, Вы будете:

уметь:

- решать задачи на применение законов распределения случайной величины;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Краткие теоретические сведения

Биномиальным называется закон распределения ДСВ X – числа появления события A в n независимых испытаниях, в каждом из которых вероятность появления события p – величина постоянная (вероятность не появления события $q=1-p$).

Следовательно, при $p=\text{const}$ ($0 < p < 1$) в n независимых испытаниях вероятность возможного значения $X=k$ вычисляется по формуле *Бернулли*:

$$P_n(X = k) = C_n^k p^k q^{n-k}, \text{ где } C_n^k = \frac{n!}{k!(n-k)!}$$

Числовые характеристики биномиального распределения: $M(X)=np$, $D(X)=npq$, $\sigma(X) = \sqrt{D(X)}$.

При $p=\text{const}$ ($p \leq 0,1$) в n независимых испытаниях вероятность возможного значения $X=k$ вычисляется по первой формуле

$$P_n(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ где } \lambda = np$$

При условии $p \rightarrow 0$, $n \rightarrow \infty$, $np \rightarrow \lambda = \text{const}$ закон распределения Пуассона является предельным случаем биномиального закона. Так как при этом вероятность p события A в каждом испытании мала, то закон распределения Пуассона называют часто *законом редких явлений*.

Закон распределения Пуассона применяют в теории массового обслуживания, в которой одним из главных является понятие *потока событий*.

Потоком событий называют последовательность событий, которые наступают в случайные моменты времени. Например: поступление вызовов на АТС, на пункт неотложной

медицинской помощи, прибытие самолетов в аэропорт, последовательность отказов элементов и др.

Интенсивностью потока λ называют среднее число событий, которые появляются в единицу времени.

Если постоянная интенсивность потока известна, то вероятность появления k событий простейшего потока за время длительностью t определяется второй формулой

$$P_t(X = k) = \frac{(\lambda \cdot t)^k}{k!} e^{-\lambda \cdot t}$$

Числовые характеристики распределения Пуассона: $M(X) = \lambda$, $D(X) = \lambda$, $\sigma(X) = \sqrt{D(X)}$.

Примеры решения задач

Задача 1. Вероятность попадания в мишень при одном выстреле для данного стрелка 0,7 и не зависит от номера выстрела. Найти вероятность того, что при 5 выстрелах произойдет: а) ровно два попадания в мишень; б) не менее четырех попаданий в мишень.

Решение. Пусть случайная величина X – число попаданий в мишень.

а) По условию задачи $n=5$, $k=2$, $p=0,7$, следовательно, $q=1-p=1-0,7=0,3$. Тогда по формуле Бернулли

$$P_5(X = 2) = C_5^2 p^2 q^{5-2} = \frac{5!}{2!3!} \cdot (0,7)^2 \cdot (0,3)^3 = 10 \cdot 0,49 \cdot 0,027 = 0,1323$$

б) По условию задачи $n=5$, $p=0,7$, $q=0,3$, не менее четырех попаданий означает либо четыре, либо пять попаданий, т.е. $k \geq 4$. Тогда по формуле Бернулли

$$P_5(X \geq 4) = P_5(X = 4) + P_5(X = 5) = C_5^4 p^4 q^{5-4} + C_5^5 p^5 q^{5-5} = 5 \cdot (0,7)^4 \cdot (0,3)^1 + 1 \cdot (0,7)^5 \cdot (0,3)^0 = 0,36015 + 0,16807 = 0,52822 \approx 0,528$$

Задача 2. Вероятность изготовления нестандартной детали 0,004. Найти вероятность того, что среди 1000 деталей окажется 5 нестандартных.

Решение. Пусть случайная величина X – число нестандартных деталей.

По условию задачи $n=1000$, $k=5$, $p=0,004$. Тогда $\lambda=np=1000 \cdot 0,004=4$ и по первой формуле пуассоновского распределения

$$P_{1000}(X = 5) = \frac{4^5}{5!} e^{-4} = \frac{4 \cdot 4 \cdot 4 \cdot 4 \cdot 4}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} \cdot \frac{1}{e^4} = \frac{128}{15} \cdot \frac{1}{e^4} \approx 0,1562$$

Задача 3. Среднее число вызовов поступающих на АТС в одну минуту, равно двум. Найти вероятность того, что за 5 минут поступит: а) два вызова; б) менее двух вызовов.

Решение. Пусть случайная величина X – число вызовов за 5 минут.

а) По условию задачи $\lambda=2$, $t=5$, $k=2$. Тогда по второй формуле пуассоновского распределения

$$P_5(X = 2) = \frac{(2 \cdot 5)^2}{2!} e^{-2 \cdot 5} = \frac{10^2}{2} \cdot \frac{1}{e^{10}} = \frac{50}{e^{10}}$$

б) По условию задачи $\lambda=2$, $t=5$, $k < 2$. Тогда по второй формуле пуассоновского распределения

$$P_5(X < 2) = P_5(X = 0) + P_5(X = 1) = \frac{(2 \cdot 5)^0}{0!} e^{-2 \cdot 5} + \frac{(2 \cdot 5)^1}{1!} e^{-2 \cdot 5} = \frac{10^0}{1} \cdot \frac{1}{e^{10}} + \frac{10^1}{1} \cdot \frac{1}{e^{10}} = \frac{1}{e^{10}} + \frac{10}{e^{10}} = \frac{11}{e^{10}}$$

Задачи для самостоятельного решения

1. На цель сбрасывается 6 бомб, вероятность попадания каждой в цель составляет 0,3. Найти вероятность поражения цели 4 бомбами.

2. Фирма снабжает своей продукцией пять магазинов. От каждого магазина может поступить заявка на очередной день с вероятностью 0,4 независимо от заявок других магазинов. Какова вероятность того, что поступит не менее четырех заявок?

3. Вероятность попадания стрелком в мишень при каждом выстреле не зависит от результатов предыдущих выстрелов и равна 0,8. Стрелок сделал 5 выстрелов. Найти вероятность следующего события: мишень поражена одной пулей.

4. Вероятность попадания бомбы в цель составляет 0,25. Сбрасывается 8 бомб. Найти вероятность того, что будет не менее 7 попаданий.

5. Магазин получил 1000 стеклянных бутылок минеральной воды. Вероятность того, что при перевозке бутылка будет разбита, равна 0,003. Найти вероятность того, что при перевозке будут разбиты не более двух бутылок.

6. На АТС поступают в среднем 12 заказов в минуту. Найти вероятность того, что за 20 мин поступят ровно 10 заказов.

7. Обувной магазин продал 200 пар обуви. Вероятность того, что в магазин будет возвращена бракованная пара равна 0,01. Найти вероятность того, что из проданных пар обуви будет возвращено не более двух пар.

8. В страховую компанию в среднем поступает 2 иска в час. Определите вероятность того, что в течение 1,5 часов не поступит ни одного иска.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 6

Демонстрация центральной предельной теоремы на основе генерации выборок и построения гистограмм

Практическое занятие № 7

Применение центральной предельной теоремы для оценки распределения сумм случайных величин

Цель работы: изучение законов распределения непрерывной случайной величины; научиться решать задачи на нахождение числовых характеристик случайной величины.

Выполнив работу, Вы будете:

уметь:

- решать задачи на нахождение числовых характеристик случайной величины;
- анализировать задачу и выделять её составные части;

- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Задача 1. Задана плотность распределения $f(x)$ непрерывной случайной величины ξ .

1. Построить график $f(x)$.
2. Найти интегральную функцию $F(x)$.
3. Вычислить математическое ожидание, дисперсию, средеквадратическое отклонение, моду, медиану.
4. Найти вероятность попадания случайной величины ξ в интервал (a, b) .

$$f(x) = \begin{cases} 0 & \text{при } x \leq -1 \\ x^2 + A \cdot |x| & \text{при } -1 < x \leq 2 \\ 0 & \text{при } x > 2 \end{cases} \quad a=0, b=1$$

Решение.

1. Найдем неизвестный параметр A плотности распределения вероятности из соотношения

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

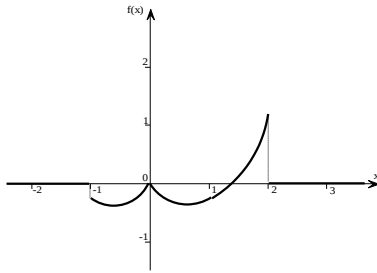
Поскольку в нашем примере плотность $f(x)$ на $(-\infty; -1]$ и на $(2; +\infty)$ равна нулю, то можно записать:

$$\int_{-1}^2 x^2 + A \cdot |x| dx = 1$$

Решив данный интеграл и полученной уравнение получим, что $A = -4/3$, тогда плотность распределения примет вид:

$$f(x) = \begin{cases} 0 & \text{при } x \leq -1 \\ x^2 - \frac{4}{3}|x| & \text{при } -1 < x \leq 2 \\ 0 & \text{при } x > 2 \end{cases}$$

График функции $f(x)$ изображен на рисунке.



2. Найдем интегральную функцию $F(x)$:

$$x \in (-\infty; -1], F(x) = P\{\xi < x\} = 0$$

$$x \in (-1; 2], F(x) = P\{\xi < x\} = \int_{-\infty}^x f(t) \cdot dt = \int_{-\infty}^{-1} 0 \cdot dt + \int_{-1}^x t^2 - \frac{4}{3} \cdot |t| \cdot dt = \frac{x^3}{3} - \frac{2}{3}|x|^2 + 1$$

$$x \in (-1; +\infty), F(x) = P\{\xi < x\} = \int_{-\infty}^{-1} 0 \cdot dt + \int_{-1}^2 t^2 - \frac{4}{3} \cdot |t| \cdot dt + \int_2^x 0 \cdot dt = 1$$

Тогда можно записать:

$$F(x) = \begin{cases} 0 & \text{при } x \leq -1 \\ \frac{x^3}{3} - \frac{2}{3}|x|^2 + 1 & \text{при } -1 < x \leq 2 \\ 1 & \text{при } x > 2 \end{cases}$$

3. Числовые характеристики искомой случайной величины.

Математическое ожидание:

$$M_{\xi} = \int_{-\infty}^{+\infty} x \cdot f(x) dx = \int_{-\infty}^{-1} x \cdot 0 dx + \int_{-1}^2 x \cdot (x^2 - \frac{4}{3} \cdot |x|) dx + \int_2^{+\infty} x \cdot 0 dx = 0.63(8)$$

Дисперсия:

$$D_{\xi} = \int_{-\infty}^{+\infty} x^2 \cdot f(x) dx - M_{\xi}^2 = \int_{-\infty}^{-1} x^2 \cdot 0 dx + \int_{-1}^2 x^2 \cdot (x^2 - \frac{4}{3} \cdot |x|) dx + \int_2^{+\infty} x^2 \cdot 0 dx - 0.63(8)^2 = 2.8500$$

Среднее квадратичное отклонение:

$$\sigma_{\xi} = \sqrt{D_{\xi}} = 1.688194301613$$

Мода равна 2.

4. Вероятность того, что $0 \leq \xi \leq 1$, вычислим по формуле:

$$P\{0 < \xi < 1\} = \int_0^1 x^2 - \frac{4}{3} |x| dx = 0.333.$$

Задачи для самостоятельного решения

1.	$F(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{x^2}{\pi^2} & \text{при } 0 < x \leq \pi; \\ 1 & \text{при } x > \pi; \end{cases}$	6.	$F(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{x^2}{e^2} & \text{при } 0 < x \leq e; \\ 1 & \text{при } x > e; \end{cases}$
2.	$F(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{x^4}{16} & \text{при } 0 < x \leq 2; \\ 1 & \text{при } x > 2; \end{cases}$	7.	$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{x^2}{1+x^2} & \text{при } x > 0; \end{cases}$
3.	$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{(x^2-x)}{2} & \text{при } 1 \leq x \leq 2; \\ 1 & \text{при } x > 2; \end{cases}$	8.	$F(x) = \begin{cases} 0 & \text{при } x \leq 0; \\ \frac{16}{25}x^2 & \text{при } 0 < x \leq \frac{5}{4}; \\ 1 & \text{при } x > \frac{5}{4}; \end{cases}$
4.	$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ x^2 & \text{при } 0 \leq x \leq 1; \\ 0 & \text{при } x > 1; \end{cases}$	9.	$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{x^3}{125} & \text{при } 0 < x \leq 5; \\ 1 & \text{при } x > 5; \end{cases}$
5.	$F(x) = \begin{cases} 0 & \text{при } x < 0; \\ \frac{x^3}{8} & \text{при } 0 < x < 2; \\ 1 & \text{при } x > 2; \end{cases}$	10.	$f(x) = \begin{cases} 0 & \text{при } x < 0; \\ x & \text{при } 0 \leq x < 1; \\ -x + 2 & \text{при } 1 \leq x < 2; \\ 0 & \text{при } x > 2; \end{cases}$

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие № 8

Моделирование закона больших чисел на основе последовательных испытаний.

Оценка среднего значения выборки и математического ожидания с помощью закона больших чисел

Практическое занятие № 9

Построение точечных оценок параметров для различных распределений. Оценка доверительных интервалов для среднего значения и дисперсии

Цель работы: научиться осуществлять расчет по заданной выборке числовых характеристик; научиться осуществлять расчет по заданной выборке числовых характеристик.

Выполнив работу, Вы будете:

уметь:

- осуществлять расчет по заданной выборке числовых характеристик;
- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

1. Пятьюдесятью абитуриентами на вступительных экзаменах в МПК получены следующие количества баллов:

12, 14, 19, 15, 14, 18, 13, 16, 17, 12,
20, 17, 15, 13, 17, 16, 20, 14, 14, 13,
17, 16, 15, 19, 16, 15, 18, 17, 15, 14,
16, 15, 15, 18, 15, 15, 19, 14, 16, 18,
18, 15, 15, 17, 15, 16, 16, 14, 14, 17.

Требуется: а) составить вариационный ряд; б) составить таблицу частот; в) построить полигон частот.

α_1									
μ_1									

2. Обследование оплаты труда 60 преподавателей дало следующие результаты (в усл.ед.):

114, 104, 112, 101, 90, 122, 126, 116, 128, 140, 124, 120, 160, 104, 140, 90, 118, 132, 154, 124, 104, 121, 156, 160, 128, 132, 104, 82, 130, 114, 142, 122, 160, 98, 116, 98, 132, 142, 116, 126, 108, 121, 102, 104, 122, 96, 122, 138, 124, 123, 108, 121, 102, 104, 122, 96, 122, 138, 124, 123.

а) Составьте интервальную таблицу частот с шириной интервала 10 (у.е.) начиная с 80 (у.е.).

б) Постройте гистограмму.

α_1									
μ_1									

3. Построить эмпирическое распределение веса студентов в килограммах для следующей выборки:

64, 57, 63, 62, 58, 61, 63, 60, 60, 61, 65, 62, 62, 60, 64, 61, 59, 59, 63, 61, 62, 58, 58, 63, 61, 59, 62, 60, 60, 58, 61, 60, 63, 63, 58, 60, 59, 60, 59, 61, 62, 62, 63, 57, 61, 58, 60, 64, 60, 59, 61, 64, 62, 59, 65.

Воспользуемся процедурой Гистограмма.

1. В ячейку A1 введем слово Наблюдения, а в диапазон A2:E12 — значения веса студентов.
2. Для вызова процедуры Гистограмма выберем из меню Сервис подпункт Анализ данных и в открывшемся окне в поле Инструменты анализа укажем процедуру Гистограмма.
3. В появившемся окне Гистограмма заполним рабочие поля: во Входной диапазон введем диапазон исследуемых данных (A2:E12); в Выходной диапазон — ссылку на левую верхнюю ячейку выходного диапазона (F1). Установим переключатели в положение Интегральный процент и Вывод графика;
4. После этого нажмем кнопку ОК.

В результате получим таблицу и диаграмму (рис.1).

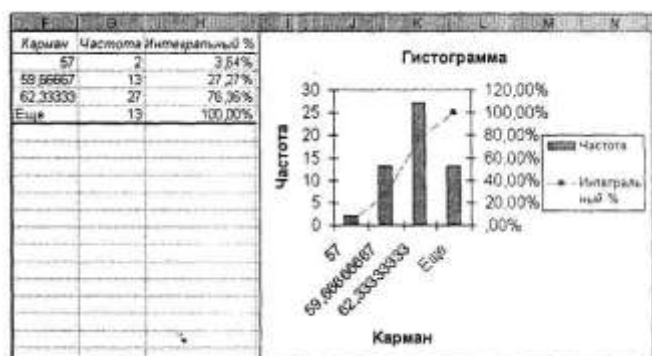


Рис.1. Результаты процедуры Гистограмма пакета Анализа.

Порядок выполнения работы:

1. Решить задания в тетради и на компьютере.
2. Получить у преподавателя задания для самостоятельной работы и решить их.

Задание

Решить задачу. На заводе железобетонных изделий N для создания марки бетона высокого качества проводилось исследование 100 различных пробных сортов бетона, для которых подсчитывался процент прочности на сжатие (случайная величина X). Получен следующий результат (таблица из 100 чисел). Найти эмпирическое распределение признака X, построить графическое отображение распределения. Найти исправленные оценки (статистики) генеральных параметров (выборочное среднее; исправленная дисперсия; исправленное среднеквадратичное отклонение; исправленная асимметрия; исправленный эксцесс. Найти моду и медиану по сгруппированным данным. Проверить гипотезу о том, что генеральная совокупность измеримого признака X, из которой извлечена выборка, распределена по нормальному закону при уровне значимости $\alpha = 0,05$.

42.7		
32.9	35.1	
37.0	37.1	44.2
28.8	42.3	43.5
34.0	44.3	41.0
	37.8	43.7
39.7		27.3
44.9	37.9	
38.5	41.7	38.7
31.5	23.2	30.8
36.6	47.7	38.1
	32.4	44.3
30.2		36.9
38.4	43.8	
35.5	31.7	36.2
45.9	44.5	42.8
37.7	37.0	37.6
	40.3	29.3
37.4		30.2
38.8	40.5	
42.8	41.1	42.5
39.4	32.9	39.6
54.3	43.0	37.1
	32.7	16.6
		43.1
37.4		
37.1	49.9	
29.6	36.7	
31.4	42.5	
47.1	26.4	
	34.2	
31.1		
43.2	37.2	
29.6	33.3	
31.4	38.8	
47.1	25.2	
	33.7	
28.2		
31.3	49.2	
29.0	35.2	
32.4	32.9	
26.6	37.5	
	37.9	
34.6		
36.1	47.0	
41.2	38.6	
32.3	29.3	
48.7	41.0	
	44.5	

Пример. Пусть для изучения количественного признака X из генеральной совокупности извлечена выборка объема $n=100$, имеющая следующее статистическое распределение:

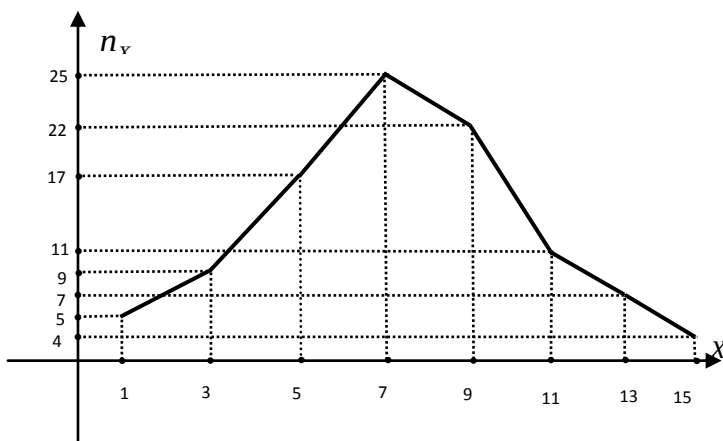
			5	7	9	11	13	5
			7	5	2	1		

Требуется:

- построить полигон частот по данному распределению выборки;
- найти выборочное среднее $\bar{X}$, выборочное среднее квадратичное отклонение $\bar{\sigma}$, «исправленное» выборочное среднее квадратичное отклонение S ;
- при данном уровне значимости α проверить по критерию Пирсона гипотезу о нормальном распределении генеральной совокупности;
- в случае принятия гипотезы о нормальном распределении генеральной совокупности найти доверительные интервалы для математического ожидания m и среднего квадратичного отклонения σ , при уровне надежности $\gamma = 1 - \alpha$.

Решение

а). Отложим на оси абсцисс варианты $x_1, \dots, x_8$, а на оси ординат – соответствующие частоты $n_1, \dots, n_8$. Соединив полученные точки ломаной, получим полигон частот.



б). Выборочное среднее $\bar{X}$, выборочное среднее квадратичное отклонение $\bar{\sigma}$, «исправленное» выборочное среднее квадратичное отклонение S находятся соответственно по формулам (см. п.13):

$$\bar{X} = \frac{1}{n} \sum_{i=1}^8 x_i n_i, \quad \bar{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^8 (x_i - \bar{X})^2 n_i}, \quad S = \sqrt{\frac{1}{n-1} \sum_{i=1}^8 (x_i - \bar{X})^2 n_i}.$$

Результаты вычислений удобно оформлять в виде таблицы.

i	x_i	n_i	$x_i n_i$	$\bar{X}$	$x_i - \bar{X}$	$(x_i - \bar{X})^2$	$(x_i - \bar{X})^2 n_i$	$\bar{\sigma}^2$	$\bar{\sigma}$	S^2
1	1	5	5		-6,62	43,82	219,10			
2	3	9	27		-4,62	21,34	192,06			
3	5	17	85		-2,62	6,86	116,62			
4	7	25	175		-0,62	0,38	9,50			
5	9	22	198		1,38	1,90	41,80			

6	11	11	121		3,38	11,42	125,62			
7	13	7	91		5,38	28,94	202,58			
8	15	4	60		7,38	54,46	217,84			
Σ		100	762	7,62			1125,12	11,25	3,36	11,36

Итак, $\bar{X} = 7,62$, $\bar{\sigma} = 3,36$, $S = 3,37$. ($S = \sqrt{S^2} = \sqrt{11,36}$)

в). Для проверки гипотезы о нормальном распределении генеральной совокупности применим критерий Пирсона. Учитывая, что варианты равноотстоят друг от друга (с шагом $h = x_i - x_{i-1} = 2$), это можно сделать следующим образом.

Концы интервалов, серединами которых являются x_i , вычисляются по формулам:

$\alpha_{i-1} = x_{i-1} - \frac{h}{2}$, $\alpha_i = x_i + \frac{h}{2}$, теоретические вероятности попадания X в интервал (α_{i-1}, α_i) ,

согласно формуле п.14, равны $p_i = \Phi\left(\frac{\alpha_i - \bar{X}}{S}\right) - \Phi\left(\frac{\alpha_{i-1} - \bar{X}}{S}\right)$, теоретические частоты равны

$n'_i = np_i$ и, согласно п.14, $\chi^2 = \sum_{i=1}^8 \frac{(n_i - n'_i)^2}{n'_i}$.

Оформим вычисления в виде таблицы (см. ниже).

Из приложения 3 находим при данном уровне значимости $\alpha = 0,05$ и при числе степеней свободы $k = 8 - 3 = 5$ критическое значение $\chi^2_{крит} = 11,1$. Поскольку $\chi^2 = 2,856 < \chi^2_{крит} = 11,1$, гипотеза о нормальном распределении генеральной совокупности принимается.

i	x_i	α_i	$\frac{\alpha_i - \bar{X}}{S}$	$\Phi\left(\frac{\alpha_i - \bar{X}}{S}\right)$	p_i	n'_i	n_i	$\frac{(n_i - n'_i)^2}{n'_i}$
0		0	-2,26	-0,4881				
1	1	2	-1,67	-0,4525	0,0356	3,56	5	0,582
2	3	4	-1,07	-0,3577	0,0948	9,48	9	0,024
3	5	6	-0,48	-0,1844	0,1733	17,33	17	0,006
4	7	8	0,11	0,0438	0,2282	22,82	25	0,208
5	9	10	0,71	0,2611	0,2173	21,73	22	0,003
6	11	12	1,30	0,4032	0,1421	14,21	11	0,725
7	13	14	1,89	0,4706	0,0674	6,74	7	0,010
8	15	16	2,48	0,4934	0,0228	2,28	4	1,298
Σ								$\chi^2 = 2,856$

г). Теперь в предположении, что случайная величина X распределена нормально, найдем доверительные интервалы для математического ожидания μ и среднего квадратичного отклонения σ по формулам п.13:

$$\mu \in \left(\bar{X} - t_\gamma \frac{S}{\sqrt{n}}, \bar{X} + t_\gamma \frac{S}{\sqrt{n}} \right), \quad \sigma \in (S(1 - q_\gamma), S(1 + q_\gamma)),$$

где $\bar{X} = 7,62$, $S = 3,37$ – выборочное среднее и исправленное выборочное среднее квадратичное отклонение, найденные в п.б), t_γ, q_γ – коэффициенты, зависящие от уровня надежности $\gamma = 1 - \alpha = 0,95$ и объема выборки $n = 100$, которые находятся из приложений 5, 6:

$$t_\gamma = t(0,95; 100) = 1,984, \quad q_\gamma = q(0,95; 100) = 0,143.$$

Итак, с надежностью $\gamma = 0,95$ μ и σ принадлежат следующим интервалам: $\mu \in (6,95; 8,29)$, $\sigma \in (2,89; 3,85)$.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №10

Проверка гипотез с использованием критерия Стьюдента для двух выборок.

Применение критерия χ^2 для проверки гипотез о независимости признаков. Оценка ошибок первого и второго рода при тестировании гипотез

Цель работы: ознакомление студентов с методами оценки различий между двумя независимыми выборками с использованием t-критерия Стьюдента и проверки гипотезы о независимости признаков с помощью критерия χ^2 . Формирование понимания природы возможных ошибок при принятии решения относительно истинности гипотез (ошибки I и II рода)..

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Задача №1. Выберите две группы испытуемых, отличающиеся значением одного количественного признака (например, уровень IQ). Определите наличие значимых различий между этими группами методом t-теста Стьюдента для независимых выборок.

Постановка задачи. Выбрать две случайные выборки (например, первая группа — студенты гуманитарных специальностей, вторая — технических), оценить различия среднего значения показателя уровня интеллекта (IQ).

Формулировка нулевой и альтернативной гипотез. Нулевая гипотеза (H₀): Среднее значение IQ одинаковое для обеих групп. Альтернативная гипотеза (H_a): Среднее значение IQ различается между группами.

Выбор необходимого метода анализа Поскольку мы имеем дело с двумя независимыми выборками, используем критерий Стьюдента для независимых выборок.

Проведение расчётов вручную либо в программе (Python/R/Excel)

```
import numpy as np
from scipy import stats
```

```
# Примеры данных (случайные выборки)
```

```
iq_group1 = np.random.normal(loc=100, scale=15, size=30)
```

```
iq_group2 = np.random.normal(loc=110, scale=15, size=30)
```

```
# Проведение теста Стьюдента
```

```
t_statistic, p_value = stats.ttest_ind(iq_group1, iq_group2)
```

```
print("Т-статистика:", t_statistic)
```

```
print("Р-значение:", p_value)
```

Интерпретация результатов. Если p-value меньше выбранного уровня значимости (α , например, 0.05), отвергнуть нулевую гипотезу и сделать вывод о наличии значимых различий.

Задача №2. Оценить независимость распределения признака (например, успеваемость студентов по математике и информатике) с помощью критерия χ^2 .

Сбор исходных данных. Создать таблицу сопряжённости для выбранных переменных (успеваемость по двум предметам).

Успеваемость	Отличники математики	Хорошисты математики
Отличники информатика	20	10
Хорошисты информатика	8	12

Расчёт ожидаемых частот для каждой ячейки таблицы. Рассчитать ожидаемое количество наблюдений для каждой категории.

Проведение вычислений критерия χ^2

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

где O_{ij} — наблюдаемые частоты,

E_{ij} — ожидаемые частоты.

Рассчитайте сумму отклонений и сравните её с критическим значением по таблице распределения χ^2 с числом степеней свободы $(R-1)(C-1)$, где R — число строк, C — число столбцов.

Или воспользуйтесь функцией Python:

```
import pandas as pd
from scipy.stats import chi2_contingency
```

```
data = [[20, 10], [8, 12]]
chi2, p_val, dof, expected = chi2_contingency(data)
print(f'Хи-квадрат статистика: {chi2}')
print(f'p-value: {p_val}')
```

Интерпретация результата. Сравнить p-value с уровнем значимости. Если оно ниже установленного порога, признать признаки зависимыми.

Задача №3. Проанализируйте возможные ошибки I и II рода, возникающие при проверке статистических гипотез.

Определение понятий ошибок. Ошибка I рода — ошибка отказа от верной нулевой гипотезы (ложноположительный результат). Ошибка II рода — принятие неверной нулевой гипотезы (ложноотрицательный результат).

Пример ошибки I рода. Например, исследование показало, что новая методика повышает успеваемость студентов, хотя в реальности никакой разницы нет.

Пример ошибки II рода. Методика действительно эффективна, однако исследование ошибочно заключило, что изменений нет.

Способы минимизации ошибок. Увеличивать размер выборки, снижать уровень значимости (α) для уменьшения вероятности ошибки I рода, повышать мощность исследования (уменьшать вероятность ошибки II рода).

Вопрос для обсуждения:

Какие факторы влияют на вероятность возникновения ошибок I и II рода?

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №11

Вычисление коэффициента корреляции Пирсона для анализа зависимостей между признаками. Построение корреляционной матрицы для многомерных данных и её интерпретация

Цель работы: освоение методов расчета коэффициентов корреляции Пирсона для определения линейных связей между численными признаками, а также изучение способов визуализации и интерпретации корреляций в многомерных наборах данных.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Основные понятия и методы

Коэффициент корреляции Пирсона. Статистический показатель, измеряющий силу и направление линейной связи между двумя непрерывными переменными. Значения варьируются от -1 до $+1$:

$|r|=1$ — сильная прямая связь (линейная зависимость);

$|r| \approx 0$ — отсутствие линейной связи.

Корреляционная матрица. Таблица, отображающая взаимные корреляционные связи всех пар признаков в наборе данных.

Задача №1: Расчет коэффициента корреляции Пирсона

Цель задания: рассчитать коэффициент корреляции Пирсона для пары признаков.

Данные: два набора измерений некоторых характеристик (например, возраст и доход):

$x_i = \{25, 30, 35, 40, 45\}$, $y_i = \{30\,000, 35\,000, 40\,000, 45\,000, 50\,000\}$

Порядок действий:

Подготовка данных: введите данные в виде массивов в Python.

```
import numpy as np
```

```
age = np.array([25, 30, 35, 40, 45])
```

```
income = np.array([30000, 35000, 40000, 45000, 50000])
```

Выполнение расчетов:

```
corr_coef = np.corrcoef(age, income)[0, 1]
```

```
print(f"Корреляционный коэффициент Пирсона: {corr_coef:.3f}")
```

Интерпретация результата: определите характер связи между возрастом и доходом на основе полученного коэффициента.

Задача №2: Построение корреляционной матрицы

Цель задания: построить корреляционную матрицу для множества признаков и интерпретировать полученные взаимосвязи.

Данные: предположим, имеется таблица с несколькими характеристиками респондентов (возраст, доход, образование, стаж работы).

Возраст	Доход	Образование	Стаж работы
25	30k	Высшее	3
30	35k	Среднее	7
35	40k	Высшее	10
40	45k	Высокое	15
45	50k	Среднее	20

Подготовьте данные в форме массива NumPy или DataFrame Pandas.

```
import pandas as pd
```

```
df = pd.DataFrame({
    'Возраст': [25, 30, 35, 40, 45],
    'Доход': [30000, 35000, 40000, 45000, 50000],
    'Образование': ['Высшее', 'Среднее', 'Высшее', 'Высшее', 'Среднее'],
    'Стаж работы': [3, 7, 10, 15, 20]
})
```

Постройте корреляционную матрицу для численных признаков.

```
correlation_matrix = df.select_dtypes(include=['number']).corr()
```

```
print(correlation_matrix)
```

Отобразите корреляционную матрицу графически (тепловая карта):

```
import seaborn as sns
```

```
import matplotlib.pyplot as plt
```

```
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.show()
```

Проинтерпретируйте выявленные зависимости: какая существует связь между возрастом и доходом, стажем работы и образованием?

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;

- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);

- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №12

Вычисление ковариации и её применение для оценки совместной изменчивости признаков

Цель работы: Научиться рассчитывать ковариацию между двумя признаками, понимать смысл и назначение данной меры, уметь интерпретировать результаты для принятия обоснованных выводов о совместном поведении величин.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Понятие ковариации

Ковариация характеризует совместное изменение двух случайных величин. Положительная ковариация означает, что обе величины изменяются в одном направлении (одновременно увеличиваются или уменьшаются); отрицательная показывает обратную тенденцию.

Формула ковариации для двух показателей x и y :

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n - 1},$$

где x_i, y_i — наблюдения по каждому показателю;
 $\bar{x}, \bar{y}$ — средние арифметические каждого ряда;
 n — объем выборки.

Задача №1: Вычислить ковариацию вручную

Используя следующий набор данных, рассчитайте ковариацию доходов сотрудников и стажа работы:

Сотрудник	Доход (тыс. руб.)	Стаж работы (лет)
A	30	2
B	40	5
C	50	8
D	60	10
E	70	12

Шаг 1. Запишите формулу для расчета ковариации.

Шаг 2. Найдите среднее значение доходов и средний стаж работы.

Шаг 3. Используя формулу, вычислите ковариацию и сделайте вывод о взаимосвязи дохода и стажа.

Решение задачи №1:

Средние значения:

Средний доход: $\bar{X} = (30 + 40 + 50 + 60 + 70) / 5 = 50$

Средний стаж: $\bar{Y} = (2 + 5 + 8 + 10 + 12) / 5 = 7.4$

Подставляем в формулу:

$\text{Cov}(X, Y) = ((30 - 50) \cdot (2 - 7.4) + (40 - 50) \cdot (5 - 7.4) + \dots + (70 - 50) \cdot (12 - 7.4)) / (5 - 1) = 120$

Таким образом, положительная ковариация (+120) свидетельствует о прямой связи между стажем и доходом.

Задача №2: Автоматизированный расчет ковариации средствами Python

Даны следующие показатели потребления электроэнергии и температуры воздуха:

Потребление (кВт/ч)	Температура (°C)
150	10
180	15
200	20
220	25
250	30

Напишите программу на Python для автоматического расчета ковариации.

Алгоритм:

Загрузите данные в списки или массивы.

Используйте библиотеку numpy для быстрого расчета ковариации.

`import numpy as np`

`consumption = np.array([150, 180, 200, 220, 250])`

`temperature = np.array([10, 15, 20, 25, 30])`

```
cov_xy = np.cov(consumption, temperature)[0][1]  
print(f"Ковариация: {cov_xy:.2f}")
```

Результат:

Ковариация равна примерно 450, что подтверждает прямую связь между потреблением энергии и температурой.

Выводы по занятию:

По окончании занятия студент сможет:

Правильно вычислять ковариацию как меру совместного изменения двух показателей.

Использовать готовые инструменты (библиотеки Python) для автоматизации процесса расчета.

Понимать роль ковариации в анализе поведения признаков и оценивать природу их взаимоотношений.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №13

Построение линейной регрессионной модели на основе экспериментальных данных.

Интерпретация коэффициентов линейной регрессии и оценка её качества

Цель работы: Научиться строить линейную регрессию, используя реальные данные, интерпретировать коэффициенты уравнения регрессии и оценивать качество полученной модели.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);

- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Понятия и термины

Линейная регрессия — статистический метод, позволяющий моделировать линейную зависимость одной зависимой переменной от одной или нескольких независимых переменных.

Общее уравнение простой линейной регрессии выглядит следующим образом:

$$Y=a+bX+e,$$

где Y — зависимая переменная;

a — свободный член (пересечение оси ординат);

b — коэффициент наклона линии регрессии;

X — независимая переменная;

e — случайная ошибка.

Постановка задачи

Предположим, у вас имеются экспериментальные данные о расходовании топлива автомобилем в зависимости от пройденного расстояния.

Необходимо построить простую линейную регрессионную модель, позволяющую предсказывать расход топлива автомобиля в зависимости от километража.

Исходные данные:

Пробег (км)	Расход топлива (литры)
100	8
200	15
300	22
400	28
500	35

Последовательность выполнения задания

Шаг 1: Импорт необходимых модулей

Для реализации будем использовать библиотеки Python:

```
import numpy as np
```

```
import pandas as pd
```

```
import matplotlib.pyplot as plt
```

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import mean_squared_error, r2_score
```

Шаг 2: Подготовка данных

Загружаем данные в виде списков или таблиц:

```
data = {  
    'пробег': [100, 200, 300, 400, 500],  
    'расход': [8, 15, 22, 28, 35]  
}
```

```
df = pd.DataFrame(data)  
X = df['пробег'].values.reshape(-1, 1) # Независимая переменная  
Y = df['расход'] # Зависимая переменная
```

Шаг 3: Создание и обучение модели

Используем класс LinearRegression из библиотеки Scikit-Learn:

```
model = LinearRegression() # Создаем объект класса  
model.fit(X, Y) # Тренируем модель
```

Шаг 4: Получение коэффициентов уравнения регрессии

Выведем свободные члены и коэффициенты наклона:

```
intercept = model.intercept_  
coefficient = model.coef_[0]
```

```
print(f"Уравнение регрессии: Y = {intercept:.2f} + {coefficient:.2f}X")
```

Полученное уравнение может выглядеть так:

$Y = 3.5 + 0.06X$

Шаг 5: Прогнозирование и проверка точности модели

Сделаем прогнозы и оценим точность с помощью метрик MSE и R^2 :

```
predictions = model.predict(X)  
mse = mean_squared_error(Y, predictions)  
r2 = r2_score(Y, predictions)
```

```
print(f"Средняя квадратичная ошибка (MSE): {mse:.2f}")
```

```
print(f"Коэффициент детерминации ( $R^2$ ): {r2:.2f}")
```

Если MSE близка к нулю, а R^2 близко к единице, значит модель хорошо описывает данные.

Шаг 6: Графическое представление

Отобразим исходные точки и линию регрессии:

```
plt.scatter(df['пробег'], df['расход'], color="blue") # Данные  
plt.plot(df['пробег'], predictions, color="green") # Линия регрессии  
plt.xlabel('Пробег (км)')  
plt.ylabel('Расход топлива (литры)')  
plt.title('Линия регрессии расхода топлива')  
plt.grid(True)  
plt.show()
```

Интерпретация результатов:

Коэффициент b (наклон) показывает, насколько изменяется расход топлива при увеличении пробега на один километр.

Свободный член a отражает базовую величину расхода при нулевом километраже.

Метрика R^2 оценивает долю вариабельности зависимой переменной, объясняемой моделью. Чем ближе R^2 к 1, тем лучше модель объясняет поведение данных.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные

вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №14

Применение нелинейной регрессии для аппроксимации данных

Цель работы: Овладеть техникой подбора наилучшей нелинейной модели для описания данных, научиться выбирать оптимальную форму нелинейной зависимости и интерпретировать результаты.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Нелинейная регрессия используется тогда, когда отношения между независимой и зависимой переменными нельзя адекватно описать линейной моделью. Она позволяет подобрать более точную модель, учитывая специфику реальных процессов.

Примером нелинейной зависимости является экспоненциальная, полиномиальная или логистическая форма.

Задача:

Пусть нам даны экспериментальные данные о скорости роста бактерий в культуре среды в зависимости от времени культивирования.

Время (часы)	Число клеток (млрд)
0	0.1
1	0.3
2	0.7
3	1.5
4	3.2
5	6.5
6	12.5

Требуется подобрать подходящую нелинейную модель, объяснить ход выбора формы модели и провести оценку качества подобранной модели.

Этапы выполнения задания:

Шаг 1: Предварительное рассмотрение данных

Анализ графика зависимости числа клеток от времени покажет рост численности бактерий в геометрической прогрессии. Поэтому имеет смысл попробовать следующие типы нелинейных функций:

Экспоненциальная модель: $N(t) = ae^{bt}$

Полиномиальная модель: $N(t) = a + bt + ct^2 + \dots$

Шаг 2: Выбор оптимальной модели

Для начала попробуем применить экспоненциальную модель, так как скорость размножения бактерий часто соответствует экспоненциальному росту.

Шаг 3: Реализация нелинейной регрессии на Python

Импортируем необходимые модули:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.optimize import curve_fit
```

Подготовка данных:

```
data = {'Время': [0, 1, 2, 3, 4, 5, 6],
        'Количество клеток': [0.1, 0.3, 0.7, 1.5, 3.2, 6.5, 12.5]}
```

```
df = pd.DataFrame(data)
```

```
X = df['Время']
```

```
Y = df['Количество клеток']
```

Подбор экспоненциальной модели:

```
def exponential(x, a, b):
```

```
    return a * np.exp(b*x)
```

```
popt, pcov = curve_fit(exponential, X, Y)
predicted_values = exponential(X, *popt)
```

Шаг 4: Оценка качества модели

Можно воспользоваться средней квадратичной ошибкой (MSE) и визуализировать график:
`from sklearn.metrics import mean_squared_error`

```
mse = mean_squared_error(Y, predicted_values)
print(f"Модель экспоненциального роста:\nПараметры: a={popt[0]:.2f}, b={popt[1]:.2f}\nОшибка (MSE): {mse:.2f}")
```

```
plt.figure(figsize=(8, 6))
plt.scatter(X, Y, label='Экспериментальные данные')
plt.plot(X, predicted_values, color='red', label=f'Аппроксимирующая кривая\n$a*e^{(bx)}$')
plt.legend()
plt.xlabel('Время (часы)')
plt.ylabel('Количество клеток (млрд)')
plt.title('Нелинейная регрессия экспоненциальной моделью')
plt.grid(True)
plt.show()
```

Шаг 5: Интерпретация полученных результатов

Оценив графики и ошибку MSE, делаем вывод о пригодности выбранной модели. Обычно хорошую модель отличает гладкость кривой и малая средняя квадратичная ошибка.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №15

Проведение однофакторного дисперсионного анализа (ANOVA) для проверки различий между группами

Цель работы: Научиться проводить однофакторный ANOVA-анализ для выявления существенных различий между средними значениями трех или более групповых выборок.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Понятие однофакторного дисперсионного анализа (ANOVA)

Однофакторный дисперсионный анализ применяется для сравнения средних значений трёх и более независимых групп. Основная идея заключается в сравнении внутригрупповой дисперсии (различий между элементами внутри каждой группы) и межгрупповой дисперсии (различий между самими группами).

Если различия между группами существенно больше внутрииндивидуальной изменчивости, делается вывод о существенности различий между группами.

Задача:

Исследователи изучают влияние различных удобрений на урожайность пшеницы. Имеются три типа удобрений, каждая группа состоит из пяти участков земли. Необходимо проверить, различаются ли удобрения по эффективности.

Удобрение А	Удобрение Б	Удобрение В
100 кг/га	110 кг/га	120 кг/га
105 кг/га	115 кг/га	125 кг/га
110 кг/га	120 кг/га	130 кг/га
115 кг/га	125 кг/га	135 кг/га
120 кг/га	130 кг/га	140 кг/га

Ход выполнения задания:

Шаг 1: Формулирование гипотез

Нулевая гипотеза (H₀): Все группы имеют одинаковые средние значения урожая.

Альтернативная гипотеза (H_A): Хотя бы одна группа отличается средним значением урожая от остальных.

Шаг 2: Подготовка данных

Реализуем пример на Python:

```
import numpy as np
import pandas as pd
from scipy import stats
```

```
# Данные
```

```
fertilizerA = np.array([100, 105, 110, 115, 120])
fertilizerB = np.array([110, 115, 120, 125, 130])
fertilizerC = np.array([120, 125, 130, 135, 140])
```

```
groups = [fertilizerA, fertilizerB, fertilizerC]
```

Шаг 3: Проведение ANOVA-тестирования

Применяем встроенную функцию ANOVA из модуля SciPy:

```
statistic, p_value = stats.f_oneway(*groups)
```

```
print(f"F-статистика: {statistic:.2f}")
```

```
print(f"P-значение: {p_value:.2f}")
```

Шаг 4: Интерпретация результатов

При уровне значимости $\alpha=0.05$ сравниваем полученное р-значение с пороговым уровнем:

Если $p < \alpha$, отвергаем нулевую гипотезу и считаем, что существуют значимые различия между группами.

Если $p > \alpha$, принимаем нулевую гипотезу и утверждаем, что разница между группами незначительна.

Например, если получено р-значение около 0.001, то мы можем уверенно утверждать, что удобрения оказывают значительное влияние на урожайность.

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.

Практическое занятие №16

Применение дисперсионного анализа для оценки влияния различных факторов на результаты экспериментов

Цель работы: Научиться применять однофакторный и многофакторный дисперсионный анализ (ANOVA) для выявления значимых различий между группами и оценки вклада отдельных факторов в общую изменчивость результатов эксперимента.

Выполнив работу, Вы будете:

уметь:

- анализировать задачу и выделять её составные части;
- составлять план действия;
- реализовать составленный план;
- оценивать результат и последствия своих действий (самостоятельно или с помощью наставника);
- определять необходимые источники информации;
- структурировать получаемую информацию;
- выделять наиболее значимое в перечне информации;
- грамотно излагать свои мысли и оформлять документы по профессиональной тематике на государственном языке;
- понимать общий смысл четко произнесенных высказываний на известные темы (профессиональные и бытовые);
- участвовать в диалогах на знакомые общие и профессиональные темы.

Материальное обеспечение:

Раздаточный материал (карточки с заданиями).

Технические средства обучения: компьютер с лицензионным программным обеспечением и мультимедиа проектор.

Задание:

Разберите решение задач, приведенных ниже, и запишите их в тетрадь.

Решите задачи из раздела для самостоятельного решения.

Примеры решения задач

Дисперсионный анализ (ANOVA) — это статистический метод, предназначенный для оценки влияния одного или нескольких факторов на изменение какого-то показателя путем разложения общей дисперсии на компоненты, обусловленные влиянием этих факторов и остаточной дисперсией.

F-тест — основной инструмент в рамках ANOVA, позволяющий определить значимость различий между средними значениями групп.

Постановка задачи:

Допустим, компания проводит маркетинговый эксперимент, проверяя эффективность трёх рекламных каналов ("Реклама на телевидении", "Контекстная реклама", "Социальные сети"). Для каждого канала фиксируют количество привлечённых клиентов за месяц.

Также дополнительно проводится анализ с учётом региона распространения рекламы ("Москва", "Санкт-Петербург", "Другие регионы").

Исходные данные представлены в следующей таблице:

Канал рекламы	Москва	СПб	Другие регионы
Телевизионная	100	120	80
Контекстная	150	160	130
Социальные сети	180	190	170

Требуется выяснить:

Есть ли существенные различия в привлечении клиентов между разными каналами рекламы?

Оказывают ли различные регионы воздействие на итоговый результат?

Последовательность выполнения задания:

Шаг 1: Формулировка гипотез

Общая гипотеза (H_0): Нет значительных различий между эффектами разных каналов рекламы и регионов.

Альтернативная гипотеза (H_1): Существуют значительные различия между результатами каналов и регионов.

Шаг 2: Подготовка данных

Запишем данные в удобной форме для обработки на Python:

```
import pandas as pd
import numpy as np
from scipy import stats

# Данные
channels = ["Телепередача", "Контекстная", "Социальные сети"]
regions = ["Москва", "СПб", "Другие регионы"]
```

Матрица данных

```
data = [
    [100, 120, 80], # Телевизионная реклама
    [150, 160, 130], # Контекстная реклама
    [180, 190, 170] # Реклама в социальных сетях
]
```

Преобразуем данные в удобный формат

```
df = pd.DataFrame(data, index=channels, columns=regions)
```

Шаг 3: Одномерный ANOVA (однофакторный анализ)

Сначала рассмотрим каналы отдельно от регионов.

Приводим данные в вид, пригодный для анализа

```
values_tv = data[0]
values_context = data[1]
values_social = data[2]
```

ANOVA-тест

```
f_value, p_value = stats.f_oneway(values_tv, values_context, values_social)
```

```
print(f"Одномерный ANOVA:")
print(f"F-статистика: {f_value:.2f}")
print(f"P-значение: {p_value:.2f}")
```

Если p-значение меньше 0.05, то можно считать, что различия между рекламными каналами значимы.

Шаг 4: Двухмерный ANOVA (многофакторный анализ)

Теперь проведём двухфакторный анализ с учетом обоих факторов одновременно.

```
import statsmodels.api as sm
```

```
from statsmodels.formula.api import ols
```

```
# Готовим фрейм данных
```

```
data_flat = []
```

```
for ch in channels:
```

```
    for reg in regions:
```

```
        data_flat.append((ch, reg, df.loc[ch, reg]))
```

```
df_flat = pd.DataFrame(data_flat, columns=["Channel", "Region", "Clients"])
```

```
# Проводим ANOVA-анализ
```

```
model = ols('Clients ~ Channel + Region', data=df_flat).fit()
```

```
anova_table = sm.stats.anova_lm(model, typ=2)
```

```
print("\nДвухмерный ANOVA:")
```

```
print(anova_table)
```

Посмотрим, какое влияние оказывает каждый фактор и взаимодействие между ними.

Шаг 5: Интерпретация результатов

Изучим таблицу ANOVA:

F-statistics и соответствующие p-values покажут, являются ли различия значимыми.

Сумма квадратов SS скажет нам, какой вклад вносят отдельные факторы и их взаимодействие.

Если p-значения будут малыми (<0.05), это будет свидетельствовать о существовании значительной зависимости количества привлечённых клиентов от фактора (канала рекламы или региона).

Форма представления результата:

Представить выполненные задания в тетради для практических работ преподавателю.

Критерии оценки:

Решение задачи считается безупречным, если правильно выбран способ решения, само решение сопровождается необходимыми объяснениями, верно выполнены нужные вычисления и преобразования, получен верный ответ, последовательно и аккуратно записано решение.

Отметка «5» ставится, если:

- работа выполнена полностью;
- в логических рассуждениях и обосновании решения нет пробелов и ошибок;
- в решении нет математических ошибок (возможна одна неточность, описка, не являющаяся следствием незнания или непонимания учебного материала).

Отметка «4» ставится, если:

- работа выполнена полностью, но обоснования шагов решения недостаточны (если умение обосновывать рассуждения не являлось специальным объектом проверки);
- допущена одна ошибка или два-три недочета в выкладках, рисунках, чертежах или графиках (если эти виды работы не являлись специальным объектом проверки).

Отметка «3» ставится, если:

- допущены более одной ошибки или более двух-трех недочетов в выкладках, чертежах или графиках, но учащийся владеет обязательными умениями по проверяемой теме.

Отметка «2» ставится, если:

- допущены существенные ошибки, показавшие, что учащийся не владеет обязательными умениями по данной теме в полной мере.