МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Тульский государственный университет»

Институт прикладной математики и компьютерных наук
Кафедра «Прикладная математика и информатика»

Утверждено на заседании кафедры
«Прикладная математика и информатика»
14 января 2020 г., протокол № 6

Заведующий кафедрой
_____ В.И. Иванов

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по выполнению практических (семинарских) занятий
по дисциплине (модулю)
«Теория вероятностей и математическая статистика»

основной профессиональной образовательной программы
высшего образования – программы бакалавриата

по направлению подготовки
01.03.03 Механика и математическое моделирование

с направленностью (профилем)
Механика деформируемого твёрдого тела

Форма обучения: очная

Идентификационный номер образовательной программы: 010303-01-20

Тула 2020 год

Разработчик методических указаний

Кочетыгов А.А., профессор каф. ПМиИ, к.т.н., доцент

(ФИО, должность, ученая степень, ученое звание)

(подпись)

Практические занятия по дисциплине «Теория вероятностей и математическая статистика» выполняются в 5–м и 6–м семестрах и имеют целью закрепление знаний по всем трём разделам (частям):

«Теория вероятностей»,

«Случайные процессы»,

«Математическая статистика».

Предлагаются индивидуальные задания для самостоятельного решения.

Для успешного освоения курса и приобретения навыков решения практических задач целесообразно использовать лекции и следующую литературу:

1. Кочетыгов А.А. Теория вероятностей: учеб. пособие / – Тула, Изд–во ТулГУ, 2016. – 234 с.

2. Кочетыгов А.А. Математическая статистика: учеб. пособие / – Тула, Изд–во ТулГУ, 2017. – 274 с.

3. Кочетыгов А.А. Математическая статистика. Решение задач с использованием пакета SPSS: учеб. пособие / Тула: Изд–во ТулГУ, 2011. – 156 с.

4. Кочетыгов А.А. Случайные процессы: учеб. пособие / – Тула, Изд–во ТулГУ, 2020. – 300 с.

Задачи, предлагаемые для решения на практических занятиях по первому разделу курса «Теория вероятностей»

Глава 1. Случайные события.

Пример 1.1. Производится три независимых выстрела с вероятностью попадания 0,4 в каждом. Найти вероятность того, что будет хотя бы 2 попадания.

Решение. Искомая вероятность $R_{2,3} = P_{2,3} + P_{3,3}$.

Воспользуемся формулой Бернулли:

$$P_{2,3} = C_3^2 0,4^2 0,6^1 = 0,288; \quad P_{3,3} = C_3^3 0,4^3 0,6^0 = 0,064; \quad R_{2,3} = 0,352.$$

Пример 1.2. Производится 4 независимых выстрела по одной цели с различных расстояний. Вероятности попаданий в этих выстрелах равны соответственно: $P_1 = 0.1$; $P_2 = 0.2$; $P_3 = 0.3$; $P_4 = 0.4$. Найти вероятность того, что: не будет попаданий; будет ровно 3 попадания.

Решение. Составляем производящую функцию

$$\varphi(z) = (0,9 + 0,1z)(0,8 + 0,2z)(0,7 + 0,3z)(0,6 + 0,4z) =$$

$$= 0,3024 + 0,4404z + 0,2144z^2 + 0,0404z^3 + 0,0024z^4.$$

Тогда искомые вероятности: $P_{0,4} = 0,3024$, $P_{3,4} = 0,0404$.

Пример 1.3. Вероятность попадания стрелком в цель равна 0,7. Сделано 25 выстрелов. Определить наивероятнейшее число попаданий.

Решение. Здесь $n = 25$, $p = 0,7$, $q = 1 - p = 0,3$.

Следовательно:

$$25 \cdot 0,7 - 0,3 < m_0 < 25 \cdot 0,7 + 0,7; \quad 17,2 < m_0 < 18,2; \quad m_0 = 18.$$

Пример 1.4. Три предприятия могут обратиться за кредитами в три банка. Каждое предприятие выбирает себе банк случайным образом и независимо от других. Клиент, обратившийся в банк, получает кредит с вероятностью P. Найти вероятность того, что ровно в двух банках из трех кредит будет выдан.

Решение. Пусть событие A означает, что ровно в двух банках из трех выдан кредит. Имеем гипотезы:

H_1 – все три банка имели посетителей;

H_2 – все предприятия обратились в один банк;

H_3 – предприятия обратились к двум банкам, а третий остался не востребованным.

Тогда $P(H_1) = 1 \cdot (2/3) \cdot (1/3) = 2/9$; $\quad P(H_2) = 1 \cdot (1/3) \cdot (1/3) = 1/9$;

$P(H_3) = 1 - 2/9 - 1/9 = 2/3$.

$$P(A|H_1) = P_{2,3} = C_3^2 P^2(1-P) = 3P^2(1-P); \;\; P(A|H_2) = 0;$$

$$P(A|H_3) = (1-(1-P)^2)P = P^2(2-P).$$

По формуле полной вероятности получаем:

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A \mid H_i) = 2P^2(1-\frac{2}{3}P).$$

Пример 1.5. Устройство состоит из трех независимо работающих элементов. Отказ любого элемента приводит к отказу всего устройства. Надежность каждого элемента равна *P*. Купили еще три аналогичных элемента. Что лучше: продублировать каждый элемент или всю схему устройства в целом?

Решение. Два варианта возможного дублирования представлены на рис.1.1.

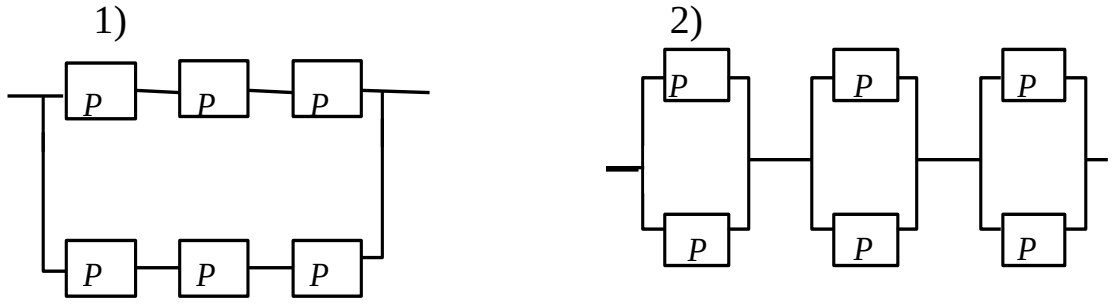


Рис. 1.1. Варианты дублирования элементов устройства

Применяя формулы сложения и умножения вероятностей рассчитаем надежность этих схем.

$$P_1 = 2P^3 - P^6; \quad P_2 = (2P - P^2)^3;$$

$$P_1 - P_2 = 2P^3 - P^6 - 8P^3 + 12P^4 - 6P^5 + P^6 =$$

$$= -6P^3 + 12P^4 - 6P^5 = -6P^3(1-2P+P^2) = -6P^3(1-P)^2 < 0.$$

Надежнее вторая схема.

Пример 1.6. Два стрелка поочередно стреляют по мишени до первого попадания. Вероятность попадания первого стрелка равна $p_1 = 0,2$, а второго – $p_2 = 0,3$. Найти вероятность того, что выиграет (попадет) первый стрелок.

Решение. Первый стрелок выигрывает соревнование (попадает) при первом выстреле с вероятностью p_1, при втором выстреле – $p_1(1-p_1)(1-p_2)$ (то есть здесь по одному промаху у первого и второго стрелка), при третьем выстреле – с вероятностью $p_1(1-p_1)^2(1-p_2)^2$ (по два промаха у каждого), при *n*–м выстреле – с вероятностью $p_1(1-p_1)^{n-1}(1-p_2)^{n-1}$.

Так как положительным исходом для первого игрока является любой из выше перечисленных исходов, то искомая вероятность будет равна сумме частных вероятностей

$$p = p_1 \sum_{n=0}^{\infty} (1-p_1)^n (1-p_2)^n .$$

С использованием формулы суммы геометрической прогрессии получаем $p = \dfrac{p_1}{1-(1-p_1)(1-p_2)} = 0.45.$

Пример 1.7. Числа натурального ряда $1,2,\ldots,n$ перемешаны и расположены в случайном порядке. Какова вероятность P_n того, что хотя бы одно число окажется на своём месте (т.е. будет равно номеру своего места)? Найти предел этой вероятности при возрастании *n*.

Решение. Пусть A_i – событие, состоящее в том, что i –е число оказалось на своём месте.

Тогда с использованием формулы вероятности суммы событий получим

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P\left(A_i A_j\right) + \sum_{i<j<k} P\left(A_i A_j A_k\right) + \ldots + (-1)^{n-1} P\left(A_1 A_2 \ldots A_n\right).$$

Из условия задачи можно заключить, что

$$P(A_i) = \frac{1}{n}; \quad P\left(A_i A_j\right) = \frac{1}{n(n-1)}; \quad P\left(A_i A_j A_k\right) = \frac{1}{n(n-1)(n-2)}; \ldots$$

Тогда

$$\sum_{i=1}^{n} P(A_i) = 1;$$

$$\sum_{1 \le i < j \le n} P(A_i A_j) = C_n^2 P(A_i A_j) = \frac{n!}{2!(n-2)!} \cdot \frac{1}{n(n-1)} = \frac{1}{2!};$$

$$\sum_{i<j<k} P(A_i A_j A_k) = C_n^3 \frac{1}{n(n-1)(n-2)} = \frac{n!}{3!(n-3)!} \cdot \frac{1}{n(n-1)(n-2)} = \frac{1}{3!}; \ldots$$

Окончательно получаем

$$P_n = 1 - \frac{1}{2!} + \frac{1}{3!} - \ldots + (-1)^{n-1} \cdot \frac{1}{n!}; \qquad \lim_{n\to\infty} P_n = 1 - \frac{1}{e} = 1 - e^{-1} = 0,632 .$$

Пример 1.8. Кусок проволоки длиной 20 см был согнут в точке, выбранной наугад (точка сгиба равномерно распределена). После этого, перегнув еще в двух местах, сделали прямоугольную рамку. Найти вероятность того, что площадь полученной рамки не превосходит 21 кв. см.

Решение. Пусть x – есть расстояние от точки первого сгиба до ближайшего конца, тогда площадь прямоугольника $S = x \cdot (10 - x)$.

Из условия $x \cdot (10 - x) \le 21$ находим, что либо $x \in (0; 3)$, либо $x \in (7; 10)$. Поэтому нужная рамка получится, если точка сгиба попадает в один из промежутков (0; 3), (7; 13), (17; 20) по длине стержня.

Искомая вероятность определится отношением благоприятствующей событию длины проволоки ($m = (3 - 0) + (13 - 7) + (20 - 17) = 12$) к общей длине ($n = 20$):

$$P = m/n = 0.6.$$

Пример 1.9. Из последовательности чисел 1, 2, ..., N отобраны n чисел $x_1 < x_2 < ... < x_n$. Найти вероятность того, что $x_m = M$.

Решение. C_{M-1}^{m-1} способами можно взять числа x_1, x_2, ..., x_{m-1} и C_{N-M}^{n-m} способами – остальные x_i.

Поэтому искомая вероятность

$$p = \frac{C_{M-1}^{m-1} \cdot C_{N-M}^{n-m}}{C_N^n}.$$

Пример 1.10. В первом ящике находится N_1 деталей первого сорта и N_2 деталей второго сорта, во втором ящике, соответственно, M_1 и M_2. Из первого ящика во второй переложено K деталей. Затем из второго ящика взяли одну деталь. Определить вероятность того, что она первого сорта.

Решение. Имеем две гипотезы:

H_1 – взятая деталь принадлежит к группе из K деталей, переложенных из первого ящика во второй;

H_2 – взятая деталь первоначально находилась во втором ящике.

Тогда

$$P(H_1) = \frac{K}{K + M_1 + M_2}; \quad P(H_2) = \frac{M_1 + M_2}{K + M_1 + M_2}.$$

Обозначим через A – событие, состоящее в том, что взятая деталь окажется первого сорта.

С учетом гипотез, вероятность этого события может быть определена по формуле полной вероятности:

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2).$$

Условные вероятности взятия детали первого сорта при соответствующих гипотезах будут равны:

$$P(A|H_1) = \frac{N_1}{N_1 + N_2}; \qquad P(A|H_2) = \frac{M_1}{M_1 + M_2}.$$

Искомая вероятность

$$P(A) = \frac{K}{K + M_1 + M_2} \cdot \frac{N_1}{N_1 + N_2} + \frac{M_1 + M_2}{K + M_1 + M_2} \cdot \frac{M_1}{M_1 + M_2}.$$

Пример 1.11. Вероятность того, что посетитель страховой компании заключит с ней какой–либо договор, равна 0,3. Сколько посетителей надо обслужить, чтобы с вероятностью не меньшей, чем 0,9, можно было утверждать, что будет заключен хотя бы один договор?

Решение. Обозначим событие A_i – i–й посетитель заключит договор с компанией. По условию задачи $P(A_i) = 0{,}3$. Тогда $P(\overline{A}_i) = 0{,}7$.

Вероятность того, что ни один из n посетителей не заключит договор равна $P_{0,n} = \prod_{i=1}^{n} P(\overline{A}_i) = (0{,}7)^n$.

Вероятность того, что хотя бы один посетитель заключит договор

$$R_{1,n} = 1 - P_{0,n} = 1 - (0{,}7)^n.$$

По условию задачи эта вероятность должна быть не менее 0,9. Получаем соотношение: $1 - (0{,}7)^n \geq 0{,}9$. Решая это неравенство, определяем n:

$$(0{,}7)^n \leq 1 - 0{,}9 = 0{,}1; \qquad n\ln(0{,}7) \leq \ln(0{,}1) \quad \Rightarrow \quad n \geq \frac{\ln(0{,}1)}{\ln(0{,}7)} = 6{,}4.$$

Окончательно получаем, что $n = 7$.

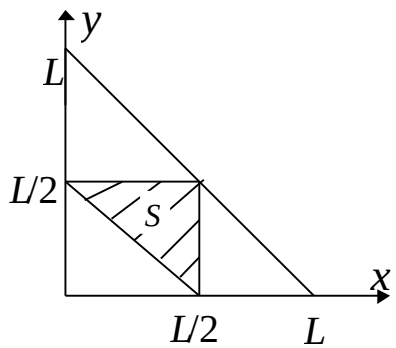
Пример 1.12. На отрезке длины L наудачу выбираются 2 точки M_1 и M_2. Определить вероятность того, что из полученных отрезков можно построить треугольник.

Решение. Пусть длины образованных отрезков равны x, y, $L - x - y$.

Эти отрезки должны удовлетворять неравенствам треугольника, т.е. сумма двух сторон больше третьей:

$$\begin{cases} x + y \geq L - x - y; \\ x + L - x - y \geq y; \\ y + L - x - y \geq x. \end{cases}$$

Сводим задачу к геометрической вероятности (рис.1.2):



$$\begin{cases} x + y \geq L/2 \\ y \leq L/2 \\ x \leq L/2 \\ x,\ y \geq 0 \end{cases}$$

$$p = \frac{\mu(S)}{\mu(Q)} = \frac{(L^2/4)\cdot 0{,}5}{L^2/2} = \frac{1}{4}.$$

Рис. 1.2. Геометрическое представление задачи

Контрольные задачи к главе 1 «Случайные события»

1.1. В лотерее 10 билетов, из которых 4 выигрышных. Какова вероятность выиграть, имея 3 билета?

1.2. Телефонный номер состоит из 5 цифр. Найти вероятность того, что все цифры различны.

1.3. Какова вероятность вытащить два разноцветных шара, если в ящике 8 белых и 12 черных шаров?

1.4. У первого акционера имеется 9 акций вида *А* и 12 акций вида *В*. У второго, соответственно, 5 и 9. В результате операции купли/продажи 7 акций первого перешли ко второму держателю акций. Найти вероятность того, что случайно выбранная акция второго акционера окажется вида *А*.

1.5. Устройство состоит из 12 независимых блоков, помеченных *Б1*, *Б2*, …, *Б12*. Вероятность того, что неисправность может произойти в одном из блоков *Б1*, *Б2*, *Б3*, *Б4* составляет 0,6. При поиске появившейся неисправности обследованы блоки *Б1*, *Б2*, *Б3*, но неисправность не обнаружена. Какова вероятность того, что неисправность будет обнаружена в блоке *Б4*?

1.6. Определить вероятность того, что трехзначный номер первой встретившейся автомашины не содержит одинаковых цифр и цифры шесть.

1.7. Из чисел 1, 2, 3, …, 15 одно за другим выбирают наугад два числа. Какова вероятность того, что разность между первым выбранным числом и вторым числом будет не меньше числа 3?

1.8. Стержень разламывается на две части в случайной точке, равномерно распределенной по всей длине стержня. Найти вероятность того, что меньший обломок имеет длину, не превосходящую одной трети длины стержня.

1.9. Для повышения надежности прибора он дублируется тремя такими же приборами. Надежность каждого прибора равна 0,6. Найти надежность системы. Сколько надо взять приборов, чтобы надежность системы стала 98%?

1.10. Из колоды в 32 карты берутся 4. Найти вероятность того, что среди них окажется хотя бы одна дама.

1.11. При переливании крови надо учитывать группы крови донора и больного. Человеку, имеющему четвертую группу крови, можно перелить кровь любой группы; человеку со второй или третьей группой крови можно перелить кровь либо той же группы, либо первой; человеку с первой группой крови можно перелить только кровь первой группы. Среди населения 30,7% имеют первую, 39,5% – вторую, 21,9% – третью и 7,9% – четвертую группы крови. Найти вероятность того, что случайно взятому больному можно перелить кровь случайно взятого донора.

1.12. Какова вероятность того, что при игре в преферанс (32 карты раздаются трем игрокам) в прикупе окажутся два туза?

1.13. Четыре поздравительные открытки случайно разложены по четырем конвертам с адресами. Найти вероятность того, что хотя бы одна открытка попала в свой конверт.

1.14. Брошены три игральные кости. Найти вероятность того, что сумма очков на выпавших гранях равна 4.

1.15. Из десяти билетов выигрышными являются два. Определить вероятность того, что среди взятых наудачу пяти билетов один выигрышный.

1.16. Устройство состоит из 5 элементов, из которых два изношены. При включении устройства включаются случайным образом два элемента. Найти вероятность того, что включенными окажутся неизношенные элементы.

1.17. Набирая номер телефона, абонент забыл последние три цифры и, помня лишь, что эти цифры различны, набрал их наудачу. Найти вероятность того, что набраны нужные цифры.

1.18. Два игрока по очереди бросают игральную кость. Каждый по одному разу. Выигравшим считается тот, кто получит большее число очков. Найти вероятность выигрыша первого игрока.

1.19. В коробке 5 одинаковых изделий, 3 из них окрашены. Наудачу извлечены два изделия. Найти вероятность того, что среди двух извлеченных изделий окажется одно окрашенное изделие.

1.20. Среди 10 электрических лампочек три нестандартные. Найти вероятность того, что две одновременно лампочки окажутся нестандартными.

1.21. Пусть вероятность того, что покупателю необходима обувь 41–го размера, равна 0,2. Найти вероятность того, что из пяти первых покупателей обувь этого размера будет необходима одному.

1.22. Вероятность того, что монета диаметром d не пересечет ни одну сторону квадратной сетки равна p. Определить размер сетки.

1.23. Два корабля должны подойти к одному и тому же причалу. Время прихода обоих кораблей независимо и равновозможно в течение данных суток. Определить вероятность того, что одному из кораблей придется ожидать освобождения причала, если время стоянки первого корабля составляет два часа, а второго – три часа.

1.24. На обслуживающее устройство в промежуток времени [0, 12] должны поступить 2 заявки. Если разность между моментами поступления заявок меньше 2, то вторая заявка теряется. Найти вероятность потери заявки.

1.25. Найти вероятность того, что сумма двух случайных чисел из отрезка [– 1, 2] больше единицы, а их произведение меньше единицы.

1.26. На пол, разграфленный параллельными прямыми на полосы шириной a, бросается наугад игла длины l ($l < a$). Найти вероятность того, что игла пересечет какую–нибудь прямую.

1.27. На плоскости отрезок длиной 10 *см* закреплен в одним концом и вращается вокруг точки закрепления так, что все направления отрезка равновероятны. Найти среднюю проекцию отрезка на заданную ось.

1.28. Вероятность того, что посетитель страховой компании заключит с ней какой–либо договор, равна 0.4. Сколько посетителей надо обслужить, чтобы с вероятностью не меньшей, чем 0.9, можно было утверждать, что будет заключен договор?

1.29. Стрельба заканчивается после третьего попадания по мишени. Найти вероятность того, что при этом будет 5 промахов, если вероятность попадания при одном выстреле равна 0,3.

1.30. Производится залп из шести орудий по некоторому объекту. Вероятность попадания в объект из каждого орудия равна 0,6. Найти вероятность ликвидации объекта, если для этого необходимо не менее четырех попаданий.

1.31. Вероятность хотя бы одного попадания стрелком в мишень при трех выстрелах равна 0,875. Найти вероятность попадания при одном выстреле.

1.32. Сотрудники отдела маркетинга полагают, что в ближайшее время ожидается рост спроса на продукцию фирмы. Вероятность этого они оценивают в 80%. Консультационная фирма, занимающаяся прогнозом рыночной ситуации, подтвердила предположение о росте спроса. Положительные прогнозы консультационной фирмы сбываются с вероятностью 95%, а отрицательные – с вероятностью 99%. Какова вероятность того, что рост спроса действительно произойдет?

1.33. 2 автомата производят одинаковые детали, которые поступают на общий конвейер. Производительность 1–го автомата вдвое больше производительности 2–го. 1–й автомат производит в среднем 60% деталей отличного качества, а 2–й – 84% деталей отличного качества. Наудачу взятая с конвейера деталь оказалась отличного качества. Найти вероятность того, что эта деталь изготовлена 1–м автоматом.

1.34. Два равносильных противника играют в шахматы. Что более вероятно: выиграть 2 партии из 4 или 3 партии из 6? Ничьи не рассматриваются.

1.35. Два баскетболиста делают по 3 броска мячом в корзину. Вероятности попадания мяча в корзину при каждом броске равны соответственно 0,6 и 0,7. Начти вероятность того, что у первого баскетболиста будет больше попаданий, чем у второго.

1.36. Известно, что в среднем 60% всего числа изготовляемых заводом телефонных аппаратов является продукцией первого сорта. Чему равна вероятность того, что в изготовленной партии окажется 6 аппаратов первого сорта, если партия содержит 10 аппаратов?

1.37. Рабочий обслуживает три станка. Каждый из станков может выходить из строя независимо друг от друга с вероятностями соответственно $p_1 = 0{,}3,\ p_2 = 0{,}4,\ p_3 = 0{,}5.$ Стало известно, что вышел из строя один станок. Какова вероятность того, что это первый станок.

1.38. Прибор может собираться из высококачественных деталей и из деталей обычного качества. 40% приборов собирается из высококачественных деталей. Если прибор собран из высококачественных деталей, то его надежность за время t равна 0.95, если из деталей обычного качества – его надежность 0.7. Прибор испытывали в течение времени t и он отработал безотказно. Какова вероятность того, что он был собран из высококачественных деталей?

1.39. Найти вероятность того, что при случайной расстановке трех ладей

на шахматной доске они не будут угрожать друг другу.

1.40. Разрыв электрической цепи происходит в том случае, когда выходит из строя хотя бы одно из трех последовательно соединенных элементов. Определить вероятность того, что не будет разрыва цепи, если элементы выходят из строя соответственно с вероятностями 0,3; 0,4 и 0,6.

1.41. Студент пришел на экзамен, зная 20 из 25 вопросов программы. Экзаменатор задал студенту 3 вопроса. Найти вероятность того, что студент знает эти вопросы.

1.42. В электрическую цепь последовательно включены три элемента, работающие независимо один от другого. Вероятности отказов первого, второго и третьего элементов соответственно равны: $p_1 = 0{,}1$; $p_2 = 0{,}15$; $p_3 = 0{,}2$. Найти вероятность того, что тока в цепи не будет.

1.43. Оптовая база обслуживает 12 магазинов. От каждого из них заявка на товары на следующий день может поступить с вероятностью 0,3. Найти наивероятнейшее число заявок на следующий день.

1.44. Пассажир может обратиться за получением билета в одну из трех касс. Вероятности обращения в каждую кассу зависят от их местоположения и равны соответственно 0,2, 0,3 и 0,5. Вероятности того, что к моменту прихода пассажира имеющиеся в кассе билеты будут распроданы, равны: для первой кассы 0,2, для второй 0,3, для третьей 0,4. Пассажир направился за билетом в одну из касс и приобрел билет. Какова вероятность того, что это была первая касса?

1.45. Имеется две партии деталей, причем известно, что в одной партии все детали удовлетворяют техническим условиям, а в другой партии 1/4 деталей недоброкачественные. Деталь, взятая из наудачу выбранной партии, оказалась доброкачественной. Определить вероятность того, что деталь взята из второй партии.

1.46. Вероятность возникновения опасной для прибора перегрузки в каждом опыте равна 0,4. Проводится испытание прибора в трех независимых опытах. Вероятность отказа прибора при одной опасной перегрузке равна 0,2; при двух перегруз как равна 0,5; при трех перегрузках равна 0,8. Определить вероятность отказа прибора в испытании.

1.47. В квадрат с вершинами (0;0), (0;1), (1;0), (1;1) наудачу брошена точка $M(a;\ b)$. Найти вероятность того, что корни уравнения $x^2 + ax + b = 0$ будут действительными.

1.48. Станок–автомат штампует детали. Вероятность того, что за смену не будет выпущено ни одной нестандартной детали, равна 0,93. Найти вероятность того, что не будет выпущено ни одной нестандартной детали: а) за две смены; б) за три смены.

1.49. На факультете обучаются 650 студентов. Найти вероятность того, что ровно 4 студента имеют день рождения 1 апреля.

1.50. Вероятность наступления события в каждом опыте равна 0,2. Опыты производятся последовательно до наступления события. Найти вероятность

того, что придется производить четвертый опыт.

1.51. В круг радиуса R вписан квадрат. Чему равна вероятность того, что поставленные наудачу внутри круга две точки окажутся внутри квадрата?

1.52. Имеются две партии изделий по 15 и 10 штук, причем в каждой партии есть одно бракованное изделие. Изделие, взятое наудачу из первой партии, переложено во вторую, после чего выбирается изделие из второй партии. Определить вероятность извлечения бракованного изделия из второй партии.

1.53. В ящике содержится 16 деталей завода № 1, 20 деталей завода № 2, 24 деталей завода № 3. Вероятность того, что детали завода № 1 отличного качества, равна 0,9; для деталей заводов № 2 и № 3 она соответственно равна 0,65 и 0,92. Найти вероятность того, что извлеченная наудачу деталь окажется отличного качества.

1.54. Вероятность того, что во время работы ЭВМ возникают сбои в процес–соре, оперативной памяти и в остальных устройствах, относятся как 3:2:5. Вероятность обнаружения сбоя в процессоре, оперативной памяти и в остальных устройствах соответственно равны 0,84, 0,78 и 0,93. Найти вероятность того, что возникший в ЭВМ сбой будет обнаружен.

1.55. На сборку поступают детали с трех автоматов. Первый выпускает 22%, второй – 30%, третий – 48% деталей данного типа. Первый автомат дает 0,26% брака, второй – 0,13%, третий – 0,18%. Найти вероятность поступления на сборку бракованной детали.

1.56. Шар $x^2+y^2+z^2=9$ помещен внутри эллипсоида $\frac{x^2}{25}+\frac{y^2}{16}+\frac{z^2}{9}=1$. Найти вероятность того, что брошенная наудачу внутрь эллипсоида точка окажется внутри шара.

1.57. Пятнадцать экзаменационных билетов содержат по два вопроса, которые не повторяются. Студент может ответить только на 25 вопросов. Определить вероятность того, что экзамен будет сдан, если для этого достаточно ответить на два вопроса из одного билета или на один вопрос из первого билета и на дополнительный вопрос из другого билета.

1.58. Разрыв электрической цепи может произойти вследствие выхода из строя элемента A_1 или двух элементов A_2 и A_3, которые выходят из строя независимо друг от друга соответственно с вероятностями 0,35, 0,24 и 0,12. Найти вероятность разрыва электрической цепи.

1.59. Найти вероятность того, что наудачу взятое двузначное число окажется кратным 2, либо 5, либо тому и другому одновременно.

1.60. Из последовательности чисел 1,2,3,…,10 наудачу выбираются два числа. Какова вероятность, что одно из них меньше 6, а другое больше 7?

1.61. Прибор может работать в 2 режимах: нормальном и в ненормальном. Нормальный режим наблюдается в 85% случаев, ненормальный – в 15% случаев. Вероятность выхода прибора из строя за время t в нормальном режиме равна 0,1, в ненормальном – 0,7. Найти вероятность выхода прибора из строя за время t.

Глава 2. Случайные величины.

Пример 2.1. Вероятность того, что произвольный посетитель страховой компании заключит с ней какой–либо договор, равна 0,4. Определить математическое ожидание, дисперсию, среднее квадратическое отклонение числа клиентов (из трех посетителей), заключивших договор со страховой компанией.

Решение. Возможные значения СВ X (число клиентов (из трех посетителей), заключивших договор со страховой компанией): 0, 1, 2, 3.

Используя формулу Бернулли (из теории вероятностей), вычислим вероятности различного числа клиентов (из трех), заключивших договор со страховой компанией:

$$P_{m,n} = C_n^m P^m q^{m-n}, \qquad C_n^m = \frac{n(n-1)\ldots(n-m+1)}{m!}.$$

Ряд распределения СВ X (число клиентов, заключивших договор со страховой компанией) имеет вид:

Значение СВ x_i	0	1	2	3
Вероятность P_i	0,216	0,432	0,288	0,064

Вычислим числовые характеристики величины X.

$$m_x = \sum_i x_i P_i = 0 \cdot 0.216 + 1 \cdot 0.432 + 2 \cdot 0.288 + 3 \cdot 0.064 = 1.2;$$

$$D_x = \sum_i (x_i - m_x)^2 P_i = (0-1.2)^2 0.216 + (1-1.2)^2 0.432 + (2-1.2)^2 0.288 +$$

$$+ (3-1.2)^2 0.064 = 0.72;$$

$$\sigma_x = \sqrt{D_x} = \sqrt{0.72} = 0.848.$$

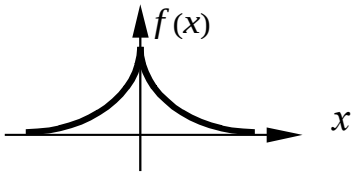
$$\mu_3 = \sum_i (x_i - m_x)^3 P_i = (0-1.2)^3 0.216 + (1-1.2)^3 0.432 + (2-1.2)^3 0.288 +$$

$$+ (3-1.2)^3 0.064 = 0.144;$$

$$A_s = \frac{\mu_3}{\sigma^3} = \frac{0.144}{0.72 \cdot 0.848} = 0.236.$$

Пример 2.2. Непрерывная случайная величина X подчинена закону распределения с плотностью.

$$f(x) = ae^{-|x|}$$



Определить математическое ожидание, дисперсию, среднее квадратическое отклонение, асимметрию, эксцесс величины X.

Решение. Определим коэффициент a.

Для этого воспользуемся свойством плотности распределения $\int_{-\infty}^{\infty} f(x)dx = 2a\int_{0}^{\infty} e^{-x}dx = 2a = 1$; отсюда a = 0,5.

Так как функция $xe^{-|x|}$ нечетная, то математическое ожидание величины X равно нулю:

$$m_x = \int_{-\infty}^{\infty} 0.5xe^{-|x|}dx = 0.$$

Дисперсия и среднее квадратичное отклонение, соответственно, равны:

$$D_x = 2\cdot\int_{0}^{\infty} 0.5e^{-x}x^2dx = 2; \quad \sigma_x = \sqrt{D_x} = \sqrt{2}.$$

Так как распределение симметрично, то $A_s = 0$.

Для вычисления эксцесса находим $\mu_4 = 2\int_{0}^{\infty} 0.5x^4e^{-x}dx = 24$.

Отсюда $E_x = \dfrac{\mu_4}{\sigma_x^4} - 3 = 3$.

Пример 2.3. Вероятность наступления события A в одном опыте равна p. За каждую удачную серию из n опытов, в которых происходит событие A, игрок получает y^n рублей; если же $n = 0$, то он платит 1 руб. Найти величину y так, чтобы игра была безобидной (математическое ожидание проигрыша равнялось бы нулю).

Решение. Пусть X – выигрыш. Возможные значения величины X в серии: $-1, y, y^2, \ldots, y^n \ldots$

Тогда $M_x = -1(1-p) + y\cdot p(1-p) + \ldots + p^n\cdot y^n(1-p) + \ldots =$

$= (1-p)(-2+1/(1-py)) = (1-p)(2py-1)/(1-py) = 0.$

Следовательно, $y = 1/(2p)$.

Пример 2.4. Вероятность приема сигнала одной радиостанции другой радиостанцией равна 0.2 при каждой попытке. Позывные сигналы подаются каждые 5 секунд до тех пор, пока не будет получен ответный сигнал. Общее время прохождения позывного и ответного сигналов равно 16 секунд. Найти среднее число подаваемых сигналов до установления двусторонней связи.

Решение. Вероятность того, что ответный сигнал будет получен на первый позывной равна $p\cdot p$, на второй – $(1-p^2)p^2$, на третий – $(1-p^2)^2p^2$ …, на n–й позывной сигнал – $(1-p^2)^{n-1}p^2$.

Вероятность $(1-p^2)^{n-1}p^2$ означает, что первые $n-1$ позывных сигналов приемной радиостанцией не приняты, а принят n–й позывной сигнал и получен ответный сигнал.

В этом случае средний номер подаваемого сигнала до установления связи будет равен величине

$$M = 1\cdot p^2 + 2\cdot(1-p^2)p^2 + 3\cdot(1-p^2)^2p^2 + \ldots + n(1-p^2)^{n-1}p^2 + \ldots.$$

С учетом задержки в 3 сигнала среднее число подаваемых сигналов до установления двусторонней связи будет равно

$$N = 3 + M = 3 + p^2\sum_{i=1}^{\infty} i(1-p^2)^{i-1}.$$

Найдем сумму этого ряда.

Воспользовавшись формулой суммы членов геометрической прогрессии, имеем: $1+q+q^2+q^3+\ldots=\dfrac{1}{1-q}$,

продифференцировав его, получим $1+2q+3q^2+4q^3+\ldots = \dfrac{1}{(1-q)^2}$.

Отсюда для случая $q = 1-p^2;\ p = 0.2$ получаем

$$N = 3 + \frac{p^2}{(1-1+p^2)^2} = 3 + \frac{1}{p^2} = 28.$$

Контрольные задачи к главе 2 «Случайные величины»

2.1. Случайная величина X задана функцией распределения

$$F(x) = \begin{cases} 0 & \text{при } x \le 2, \\ 0,5x-1 & \text{при } 2 < x \le 4, \\ 1 & \text{при } x > 4. \end{cases}$$

Найти вероятность того, что в результате испытания X примет значения, заключенные в промежутке (2,5; 3,6).

2.2. При каких значениях параметров A и B функция $F(x) = A + Be^{-x}$ может быть функцией распределения для неотрицательных значений случайной величины X.

2.3. Случайная величина X задана функцией распределения

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

Найти вероятность того, что в результате четырех независимых испытаний величина X ровно 3 раза примет значение, принадлежащее интервалу (0,25;0,75).

2.4. Вероятность попадания мячом в корзину при одном броске равна 0,3. Составить закон распределения числа попаданий при трех бросках.

2.5. Два стрелка делают по одному выстрелу в мишень. Вероятность попадания в нее первым стрелком равна 0,5, вторым – 0,4. Составить закон распределения числа попаданий в мишень.

2.6. Испытываются 3 элемента, работающих независимо друг от друга. Длительности времени (в часах) безотказной работы элементов имеют функции плотности распределения: для первого: $F_1(t) = 1 - e^{-0,1t}$, для второго: $F_2(t) = 1 - e^{-0,2t}$, для третьего: $F_3(t) = 1 - e^{-0,3t}$. Найти вероятность того, что в интервале времени от 0 до 5 часов: откажет только один элемент; откажут только два элемента; откажут все три элемента.

2.7. Дана плотность вероятности $f(x)$ случайной величины X:

$$f(x) = \begin{cases} 0 & \text{при } x \le 0, \\ \cos x & \text{при } 0 < x \le \dfrac{\pi}{2}, \\ 0 & \text{при } x > \dfrac{\pi}{2}. \end{cases}$$

Найти функцию распределения F(x).

2.8. Устройство состоит из трех независимо работающих элементов. Вероятность отказа каждого элемента в одном опыте равна 0,1. Составить закон распределения числа отказавших элементов в одном опыте.

2.9. В партии из 6 деталей имеется 4 стандартных. Наудачу отобраны 3 детали. Составить закон распределения числа стандартных деталей среди отобранных.

2.10. Случайная величина X имеет плотность распределения

$$f(x) = \begin{cases} 0, & x \notin (a,b); \\ \dfrac{2}{b-a} - \dfrac{2}{(b-a)^2} \left| x - \dfrac{b+a}{2} \right| & x \in (a,b). \end{cases}$$

причем a и b не известны, но $b > a$, а $m_x = 5$ и $D_x = 6$. Найдите a и b.

2.11. В среднем по 10% договоров страховая компания выплачивает страховые суммы в связи с наступлением страхового случая. Вычислить математическое ожидание и дисперсию числа таких договоров среди наудачу выбранных четырех.

2.12. Из пяти роз две белые. Составить закон распределения случайной величины, выражающей число белых роз среди двух одновременно взятых.

2.13. Среди 15 собранных агрегатов 6 нуждаются в дополнительной смазке. Составить закон распределения числа агрегатов, нуждающихся в дополнительной смазке, среди пяти наудачу выбранных из общего числа.

2.14. Из поступивших в ремонт 10 часов 7 нуждаются в общей чистке механизма. Часы не рассортированы по виду ремонта. Мастер, желая найти часы,

нуждающиеся в чистке, рассматривает их поочередно и, найдя такие часы, прекращает дальнейший просмотр. Найти математическое ожидание и дисперсию числа просмотренных часов.

2.15. Абонент забыл последнюю цифру нужного ему номера телефона, однако помнит, что она нечетная. Найти математическое ожидание и дисперсию числа сделанных им наборов номера телефона до попадания на нужный номер, если последнюю цифру он набирает наудачу, а набранную цифру в дальнейшем не набирает.

2.16. Вероятность отказа за время испытаний на надежность для каждого прибора серии равна p. Определить математическое ожидание числа приборов, давших отказ, если испытанию подверглись N приборов.

2.17. Дискретная случайная величина X принимает 3 возможных значения: $X_1 = 4$ с вероятностью $p_1 = 0,5$; $X_2 = 6$ с вероятностью $p_2 = 0,3$ и X_3 с вероятностью p_3. Найти X_3 и p_3, зная, что M(X) = 8.

2.18. Отдел технического контроля проверяет изделия на стандартность. Вероятность того, что изделие стандартно, равна 0,9. В каждой партии содержится 5 изделий. Найти математическое ожидание случайной величины X – числа партий, в каждой из которых содержится ровно 4 стандартных изделия, если проверке подлежат 50 партий.

2.19. Найти дисперсию случайной величины X – числа появлений события A в двух независимых испытаниях, если вероятности появления события в этих испытаниях одинаковы и известно, что $M(X) = 0,9$.

2.20. Математическое ожидание и среднее квадратическое отклонение нормально распределенной случайной величины X соответственно равны 20 и 5. Найти вероятность того, что в результате испытания X примет значение, заключенное в интервале (15; 25).

2.21. Дана функция: $\varphi(x) = \begin{cases} 0 & \text{при } x < 0, \\ Cxe^{-x} & \text{при } x \geq 0. \end{cases}$

При каком значении параметра C эта функция является плотностью распределения некоторой непрерывной случайной величины X? Найти математическое ожиданий и дисперсию случайной величины X.

2.22. Величина X задана плотностью вероятности $f(x) = C(x^2 + 2x)$ в интервале (0;1), вне этого интервала $f(x) = 0$. Найти: постоянный параметр C; математическое ожидание величины X.

2.23. Случайная величина X в интервале (0;1) задана плотностью вероятности $f(x) = 0,5\sin x$; вне этого интервала $f(x) = 0$. Найти дисперсию X.

2.24. Устройство состоит из N элементов, работающих независимо один от другого. Вероятность отказа любого элемента в течение времени t равна p. Необходимо найти математическое ожидание и дисперсию этой случайной величины.

2.25. Производится три независимых выстрела по мишени. Вероятность попадания при каждом выстреле равна 0.25. Определить среднее квадратиче-

ское отклонение числа попаданий при трех выстрелах.

2.26. Среднее число клиентов, посещающих страховую компанию за 10 мин., равно трем. Найти вероятность того, что в ближайшие 5 минут придет хотя бы один клиент.

2.27. Среднее время настройки прибора составляет 5 минут и подчинено показательному распределению. Мастер уже потратил 5 минут на настройку очередного прибора. Найти вероятность того, что он затратит еще не менее трех минут на настройку этого прибора.

2.28. В отдел заказов в среднем приходит 18 клиентов в час. Определить вероятность того, что за две текущие минуты в отдел заказов придет хотя бы один клиент.

2.29. Время ожидания заявки в очереди на процессор подчиняется показательному закону распределения со средним значением 20 секунд. Найти вероятность того, что очередная (произвольная) заявка будет ожидать процессор более 35 секунд.

2.30. Группа студентов в количестве 15 человек проводит собрание в зале, в котором 20 рядов по 10 мест в каждом. Каждый студент занимает место в зале случайным образом. Какова вероятность того, что не более трех человек будут находиться на седьмом месте ряда?

2.31. В партии из 10 деталей имеется 3 нестандартных. Составить закон распределения дискретной случайной величины X – числа нестандартных деталей среди трех отобранных и найти ее математическое ожидание.

2.32. Непрерывная случайная величина X распределена по показательному закону с плотностью $f(x) = 0{,}7e^{-0{,}7x}$ $(x \geq 0)$. Найти вероятность того, что она примет значения из интервала (2; 6).

2.33. Найти математическое ожидание и дисперсию числа выигрышных лотерейных билетов, если их приобрели 20 штук, а вероятность выигрыша по одному билету равна 0,3.

2.34. Найти дисперсию дискретной случайной величины X – числа отказов элемента некоторого устройства в десяти независимых опытах, если вероятность отказа элемента в каждом опыте равна 0,15.

2.35. Для какого из указанных интервалов функция $\sin x$ может являться функцией распределения случайной величины X: $(0; \pi/2)$, $(0; \pi)$? Найти вероятность попадания случайной величины X в интервал $(\pi/6; \pi/2)$.

2.36. Функция распределения величины X имеет вид: $F(x) = A + B\operatorname{arctg} 0{,}5x$; $-\infty < X < +\infty$. Определить: *а*) постоянные A и B; *б*) плотность распределения вероятностей $f(x)$; в) $P(2/\sqrt{3} \leq X \leq 2\sqrt{3})$.

2.37. Длина изготавливаемой станком–автоматом детали представляет собой случайную величину, распределенную по нормальному закону с параметрами a =15 см, σ = 0,2 см. Найти вероятность брака, если допустимые размеры детали должны быть 15 ± 0,3 см. Какую точность длины изготовляемой детали можно гарантировать с вероятностью 0,97?

2.38. Клиенты банка, не связанные друг с другом, не возвращают кредиты

в срок с вероятностью 0,1. Составить закон распределения числа возвращенных в срок кредитов из 5 выданных. Найти математическое ожидание и этой случайной величины.

2.39. Плотность распределения вероятностей случайной величины X имеет вид $f(x) = Ae^{-x^2+2x+3}$. Найти значение параметра A, $M(X)$, $D(X)$, функцию распределения $F(X)$, вероятность $P(-1/3 \le X \le 4/3)$.

2.40. Найти дисперсию диаметра втулок, представляющего собой случайную величину, распределенную по нормальному закону, если отклонение от математического ожидания, не превосходящее 0,15 см, имеет место с вероятностью 0,8.

2.41. Имеются 4 ключа, из которых только один подходит к замку. Найти математическое ожидание и дисперсию числа попыток открывания замка, если испробованный ключ в последующих попытках не участвует.

2.42. В отдел заказов в среднем приходит 12 клиентов в час. Определить вероятность того, что за три текущие минуты в отдел заказов придет хотя бы один клиент.

2.43. Дискретная случайная величина X принимает значение $x_1 = 1$, $x_2 = 2$, $x_3 = 3$. Известно также $M(X) = 2{,}3$ и $M(X^2) = 5{,}9$. Найти вероятности, соответствующие возможным значениям величины X.

2.44. При установившемся технологическом процессе 70 % всей производимой продукции станок–автомат выпускает первым сортом, 30 % – вторым. Составить закон распределения числа изделий первого сорта среди 4 изделий, отобранных случайным образом. Найти математическое ожидание и дисперсию рассматриваемой случайной величины.

Глава 3. Системы случайных величин.

Пример 3.1. Система двух случайных величин подчинена закону распределения с плотностью: $f(x,y) = \dfrac{1}{\pi^2(1+x^2)(1+y^2)}$. Найти функцию распределения $F(x,y)$. Определить вероятность попадания случайной точки (X, Y) в квадрат R с координатами: (1, 1), (0, 1), (1, 0), (0, 0).

Решение. $$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(x,y)dxdy = \frac{1}{\pi^2}\int_{-\infty}^{x}\int_{-\infty}^{y}\frac{dxdy}{(1+x^2)(1+y^2)} =$$

$$= \left[\frac{1}{\pi}arctg(x) + \frac{1}{2}\right]\cdot\left[\frac{1}{\pi}arctg(y) + \frac{1}{2}\right].$$

Тогда вероятность попадания в прямоугольник R:

$$P((X,Y)\subset R) = \int_0^1\int_0^1 f(x,y)dxdy = \frac{1}{\pi^2}\int_0^1\int_0^1\frac{dxdy}{(1+x^2)(1+y^2)} =$$

$$= \frac{1}{\pi^2}\int_0^1 \frac{dx}{1+x^2}\int_0^1 \frac{dy}{1+y^2} = \frac{1}{16}.$$

Пример 3.2. Плотность распределения системы (X,Y) имеет вид: $f(x,y) = \dfrac{1}{\pi^2(x^2+y^2+x^2y^2+1)}$. Определить, зависимы или независимы величины X и Y.

Решение. Разлагая знаменатель на множители, имеем:

$$f(x,y) = \frac{1}{\pi(1+x^2)}\frac{1}{\pi(1+y^2)}.$$

Из того, что функция $f(x,y)$ распалась на произведение двух функций, из которых одна зависит только от X, другая - только от Y, заключаем, что величины X и Y должны быть независимы.

Действительно,

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy = \frac{1}{\pi(1+x^2)}\int_{-\infty}^{\infty}\frac{dy}{\pi(1+y^2)} = \frac{1}{\pi(1+x^2)}.$$

Аналогично, $f_2(x) = \dfrac{1}{\pi(1+y^2)}$.

Отсюда убеждаемся, что $f(x,y) = f(x)f(y)$, и, следовательно, величины X и Y независимы.

Пример 3.3. Система случайных величин (X,Y) подчинена закону распределения с плотностью $f(x,y) = \begin{cases} a(x+y), & 0 \le x \le 3; \quad 0 \le y \le 3; \\ 0, & x < 0; \quad y < 0; \quad y > 3. \end{cases}$

Требуется: определить параметр a; вычислить вероятность попадания точки (X,Y) в квадрант, ограниченный прямыми $x=1$; $x=2$; $y=1$; $y=2$; найти m_x, m_y, σ_x, σ_y.

Решение. Найдем параметр a из уравнения $a\int_0^3\int_0^3 (x+y)dxdy = 1$.

Отсюда:

$$a\int_0^3\int_0^3 (x+y)dxdy = a\int_0^3\left[xy+\frac{y^2}{2}\right]_0^3 dx = a\int_0^3\left[3x+\frac{9}{2}\right]dx = a\left[\frac{3}{2}x^2+\frac{9}{2}x\right]_0^3 =$$

$$= a\left[\frac{27}{2}+\frac{27}{2}\right] = 27\cdot a; \quad 27\cdot a = 1; \quad a = \frac{1}{27}.$$

Искомая вероятность

$$P[(X,Y)\subset Q]=\frac{1}{27}\int_1^2\int_1^2(x+y)dxdy=\frac{1}{27}\int_1^2\left[xy+\frac{y^2}{2}\right]\Bigg|_1^2 dx=$$

$$=\frac{1}{27}\int_1^2\left[2x+2-x-\frac{1}{2}\right]\Bigg|_1^2 dx=\frac{1}{27}\int_1^2\left[x+\frac{3}{2}\right]dx=\frac{1}{27}\left[\frac{x^2}{2}+\frac{3x}{2}\right]\Bigg|_1^2 dx=$$

$$=\frac{1}{27}\left[2+3-\frac{1}{2}-\frac{3}{2}\right]=\frac{1}{9}.$$

По соответствующим формулам вычисляем другие числовые характеристики:

$$m_x=\frac{1}{27}\int_0^3\int_0^3 x(x+y)dxdy=\frac{1}{27}\int_0^3\left[x^2y+\frac{xy^2}{2}\right]\Bigg|_0^3 dx=\frac{1}{27}\int_0^3\left[3x^2+\frac{9}{2}x\right]dx=$$

$$=\frac{1}{27}\left[x^3+\frac{9}{4}x^2\right]\Bigg|_0^3 dx=\frac{1}{27}\left[27+\frac{81}{4}\right]=1\frac{3}{4};\qquad m_x=m_y=1\frac{3}{4};$$

$$\sigma_x^2=\sigma_y^2=\iint_Q(x-m_x)^2 f(x,y)dxdy=\frac{1}{27}\int_0^3\int_0^3\left[x-\frac{7}{4}\right]^2(x+y)dxdy=$$

$$=\frac{1}{27}\int\int\left[x-\frac{7}{4}\right]\left[x-\frac{7}{4}+y+\frac{7}{4}\right]dxdy=$$

$$=\frac{1}{27}\int_0^3\int_0^3\left[x-\frac{7}{4}\right]^3 dxdy+\frac{1}{27}\int_0^3\int_0^3\left[x-\frac{7}{4}\right]^2\left[y+\frac{7}{4}\right]dxdy=$$

$$=\frac{1}{27}\int_0^3 y\left[x-\frac{7}{4}\right]^3\Bigg|_0^3 dx+\frac{1}{27\cdot 2}\int_0^3\left[x-\frac{7}{4}\right]^2\left[y+\frac{7}{4}\right]^2\Bigg|_0^3 dx=\frac{11}{16};$$

$$\sigma_x=\sigma_y=\frac{\sqrt{11}}{4}.$$

Пример 3.4.

События A и B имеют одинаковую вероятность p.

Какова должна быть условная вероятность $P(A|B)$, чтобы коэффициент корреляции между A и B был равен числу r.

Решение.

Пусть $X = 1$ или 0 в зависимости от того, произошло событие A или нет.

Пусть $Y = 1$ или 0 в зависимости от того, произошло событие B или нет. Обозначим $P(A|B) = P(B|A) = p_1$.

Тогда $P(AB) = pp_1$, $P(\bar{A}B) = p_1(1 - p_1)$, $P(A\bar{B}) = p(1-p_1)$, $P(\bar{A}\bar{B}) = 1 - 2p + pp_1$, т.е. совместное распределение (X,Y) задается таблицей:

X	Y	
	1	0
1	pp_1	$p(1 - p_1)$
0	$p(1 - p_1)$	$1 - 2p + pp_1$

Следовательно, $M[X] = M[Y] = p$, $D[X] = D[Y] = p(1-p)$, $M[XY] = pp_1$, $R_{xy} = p(p_1 - p)$, $r_{xy} = \dfrac{p(p_1 - p)}{p(1-p)} = r$, т.е. $p_1 = P(A|B) = p + r(1 - p)$.

Пример 3.5. В продукции завода брак вследствие дефекта инструмента составляет 3%, а вследствие дефекта сырья (материала) - 4.5%. Годная продукция составляет 95%. Найти коэффициент корреляции дефектов инструмента и сырья.

Решение. Пусть X - случайная величина, принимающая значение 1, если данное изделие обладает браком вследствие дефекта инструмента, и $X = 0$ - в противном случае.

Аналогично, $Y = 1$, если данное изделие обладает браком вследствие дефекта сырья, и $Y = 0$, если дефекта сырья не проявился.

По условию $P(X = 0, Y = 0) = 0,95$.

Так как

$P(X = 0) = 0,97 = P(X = 0, Y = 0) + P(X = 0, Y = 1) = 0,95 + P(X = 0, Y = 1)$,

то $P(X = 0, Y = 1) = 0,02$.

Аналогично находим

$$P(X = 1, Y = 0) = 0,005; \qquad P(X = 1, Y = 1) = 0,025.$$

Совместное распределение величин (X, Y) задается таблицей:

X	Y	
	1	0
1	0,025	0,005
0	0,020	0,950

Тогда $M[X] = 0.03$; $M[X^2] = 0.03$; $D[X] = 0.0291$; $M[Y] = 0.045$;

$M[Y^2] = 0.045$; $D[Y] = 0.0042975$; $R_{xy} = 0.02365$; $r_{xy} = 0.67$.

То есть, коэффициент корреляции дефектов инструмента и сырья равен 0,67.

Контрольные задачи к главе 3 «Системы случайных величин»

3.1. Условная плотность распределения случайной величины X при условии, что случайная величина Y приняла значение y, равна $f(x \mid y) = \dfrac{2}{\pi} e^{-(4x+y)^2/4}$. Найти условную дисперсию $D_{x|y}$.

3.2. Непрерывная двумерная случайная величина (X,Y) распределена равномерно в квадрате с вершина $(0,0),(0,1),(1,0)$ и $(1,1)$. Найдите вероятность попадания случайного вектора (X,Y) в круг $(x-1)^2+(y-1)^2 \le 1/2$.

3.3. Распределение двумерного вектора (X,Y) задано в таблице:

X	Y		
	0,1	0,15	0,2
0,3	0,25	0,15	0,32
0,6	0,1	0,05	0,13

Найти условное распределение случайной величины *Y* при условии, что случайная величина *X* приняла значение x_i, $i=1,2$.

3.4. Совместная плотность распределения непрерывной двумерной случайной величины (X,Y) имеет вид

$$f(x,y)=\begin{cases}0, (x-3)^2+(y-2)^2>4;\\ C\left(2-\sqrt{(x-3)^2+(y-2)^2}\right), (x-3)^2+(y-2)^2\le 4.\end{cases}$$

Найдите постоянную C.

3.5. Двумерный случайный вектор (X,Y) распределен равномерно в треугольнике с вершинами в точках $(0;0)$, $(0;2)$, $(1;0)$. Найти условную плотность распределения случайной величины *X* при условии, что случайная величина *Y* приняла значение *y*.

3.6. Задана функция распределения двумерной СВ *(X,Y)*:

$$F(x,y)=\begin{cases}1-3^{-x}-3^{-y}+3^{-x-y},\ x\ge 0,\ y\ge 0;\\ 0,\quad x<0 \text{ или } y<0.\end{cases}$$

Найти вероятность попадания случайной точки *(X,Y)* в прямоугольник, ограниченный прямыми $x = 0$, $x = \pi/4$, $y = \pi/6$, $y = \pi/3$.

3.7. Совместная плотность распределения непрерывной двумерной случайной величины (X,Y) имеет вид:

$$f(x,y)=\begin{cases}0, & x\le 0 \quad \text{или} \quad y\le 0;\\ Ce^{-4x-2y}, & x>0 \quad \text{и} \quad y>0.\end{cases}$$

Найти постоянную *C*.

3.8. Условная плотность распределения случайной величины *X* при условии, что случайная величина *Y* приняла значение *y*, равна

$$f\left(x\,|\,y\right)=\begin{cases}1/6, & |x|\le 3 \quad \text{и} \quad |y|\le 10;\\ 0, & |x|>3 \quad \text{и} \quad |y|\le 10.\end{cases}$$

Найти условное математическое ожидание $m_{x|y}$.

3.9. Двумерная случайная величина (X,Y) имеет совместную функцию распределения

$$F(x,y)=\frac{1}{4\pi^2}\left(4\operatorname{arctg} x\ \operatorname{arctg} y+2\pi\operatorname{arctg} x+2\pi\operatorname{arctg} y+\pi^2\right).$$

Найти вероятность события $\left\{-1\le X\le 1,\quad 1\le Y<\sqrt{3}\right\}$.

3.10. Непрерывный двумерный случайный вектор (X,Y) имеет плотность распределения

$$f(x,y)=\begin{cases}Cy, & (x,y)\in D;\\ 0, & (x,y)\notin D,\end{cases}$$

где D – область, определения, ограниченная линиями $y=x^2$ и $y=1$. Найдите условную плотность распределения случайной величины X, при условии, что случайная величина Y приняла значение y.

3.11. Двумерная случайная величина (X,Y) имеет совместную функцию распределения

$$F(x,y)=\begin{cases}0, & x\le 0 \quad \text{или} \quad y\le 0;\\ \dfrac{\sin x+\sin y+\sin(x+y)}{2}, & 0<x\le\dfrac{\pi}{2} \quad \text{и} \quad 0<y\le\dfrac{\pi}{2};\\ \dfrac{\sin x-\cos x+1}{2}, & 0<x\le\dfrac{\pi}{2} \quad \text{и} \quad y>\dfrac{\pi}{2};\\ \dfrac{\sin y-\cos y+1}{2}, & x>\dfrac{\pi}{2} \quad \text{и} \quad 0<y\le\dfrac{\pi}{2};\\ 1, & x>\dfrac{\pi}{2} \quad \text{и} \quad y>\dfrac{\pi}{2}.\end{cases}$$

Проверьте, являются ли случайные величины X и Y независимыми.

3.12. Задана дискретная двумерная СВ *(X,Y)*:

Y	X		
	$x_1 = 2$	$x_2 = 5$	$x_3 = 8$
$y_1 = 0.4$	0.15	0.30	0.35
$y_2 = 0.8$	0.05	0.12	0.03

Найти условное значение математического ожидания составляющей *Y* при условии, что составляющая *X* приняла значение $X = x_2 = 5$.

3.13. Дискретные независимые СВ заданы своими распределениями:

X	5	8	Y	1	7
P	0.2	0.8	P	0.56	0.44

Найти коэффициент вариации величины $Z = X + Y$.

3.14. Двумерный случайный вектор (X,Y) имеет нормальное распределение с плотностью распределения $f(x,y) = \frac{\sqrt{3}}{\pi} e^{-4x^2-2xy-y^2}$. Найти условную плотность распределения случайной величины *Y* при условии, что случайная величина *X* приняла значение *x*.

3.15. В круге $x^2+y^2 \le R^2$ двумерная плотность вероятности $f(x,y) = C\left(R - \sqrt{x^2+y^2}\right)$ вне круга $f(x,y) = 0$. Найти постоянную *C*.

3.16. Задана плотность совместного распределения непрерывной двумерной СВ (*X*,*Y*): $f(x,y) = \frac{1}{\pi} e^{-(1/2)(x^2+2xy+5y^2)}$. Найти плотность распределения величины *X*.

3.17. Функция распределения СВ *X* имеет вид $F(x) = a - b\ arctg\ x$. Найти плотность распределения вероятностей, определив постоянные *a* и *b*.

3.18. В продукции завода брак вследствие дефекта *A* составляет 12 %, причём среди забракованной по признаку *A* продукции в 4 % случаев встречается дефект *B*, а в продукции свободной от дефекта *A* дефект *B* встречается в 2% случаев. Найти коэффициент корреляции дефектов *A* и *B*.

3.19. Система двух случайных величин подчинена закону распределения с плотностью: $f(x,y) = \frac{1}{\pi^2(1+x^2)(1+y^2)}$. Найти функцию распределения $F(x,y)$:

3.20.

3.20. Случайные величины *X*, *У* независимы и нормально распределены с параметрами $M(X) = M(У) = 0$, $D(X) = D(У) = 1$. Найти вероятность того, что случайная точка (*X*, *У*) попадет в кольцо $2 \le \sqrt{x^2+y^2} \le 3$.

3.21. Найти плотность распределения случайной величины $(X,У)$, если функция распределения $F(x, y) = (1 - e^{-2x})(1 - e^{-3y})$ при $x \geq 0$, $y \geq 0$.

Глава 4. «Функции случайных величин».

Пример 4.1. Случайная величина X подчинена закону Коши с плотностью распределения $f(x) = \dfrac{1}{\pi(1+x^2)}$, величина Y связана с X зависимостью $Y = 1 - X^3$. Найти плотность распределения величины Y.

Решение. Так как функция $Y = 1 - X^3$ монотонная на участке $(-\infty, +\infty)$, то

$$X = \psi(Y) = (1-Y)^{\frac{1}{3}}, \qquad \psi'(y) = -\frac{1}{3(1-y)^{2/3}}.$$

Следовательно $f(y) = \left\{ 3\pi \cdot \left(1 + (1-y)^{\frac{2}{3}} \right) \cdot (1-y)^{\frac{2}{3}} \right\}^{-1}$.

Пример 4.2. Ошибка прибора выражается функцией $U = 3Z + 2X - Y - 4$, где X, Y, Z - так называемые первичные ошибки, представляющие собой систему случайных величин, которая характеризуется математическими ожиданиями $m_x = -4$, $m_y = 1$, $m_z = 1$ и корреляционной матрицей

$$\begin{Vmatrix} D_x & R_{xy} & R_{xz} \\ & D_y & R_{yz} \\ & & D_z \end{Vmatrix} = \begin{Vmatrix} 2 & 1 & -1 \\ & 3 & 1 \\ & & 4 \end{Vmatrix}.$$

Определить математическое ожидание и дисперсию ошибки прибора.

Решение. Так как функция U линейна, то

$$m_u = 3m_z + 2m_x - m_y - 4 = -10,$$

$$D_u = 3^2 D_z + 2^2 D_x + (-1)^2 D_y + 2(6R_{xz} - 3R_{yz} - 2R_{xy}) = 25.$$

Пример 4.3. Абсцисса точки попадания снаряда выражается формулой $X = X_н + 1.8 \cdot \omega \cdot H + X_б$, где $X_н$ (ошибка наводки, м), ω (угловая скорость, рад/сек), H (дальность стрельбы, м), $X_б$ (баллистическая ошибка, м) представляют собой случайные величины с математическими ожиданиями $m_{x_н} = 0$, $m_\omega = 0.1$, $m_H = 1000$, $m_{x_б} = 0$ и средними квадратическими отклонениями $\sigma_{x_н} = 4$, $\sigma_\omega = 0.005$, $\sigma_H = 50$, $\sigma_{x_б} = 3$.

Нормированная корреляционная матрица системы имеет вид:

$$\begin{Vmatrix} 1 & 0.5 & 0.3 & 0 \\ & 1 & 0.4 & 0 \\ & & 1 & 0 \\ & & & 1 \end{Vmatrix}.$$

Найти математическое ожидание и среднее квадратическое отклонение величины X.

Решение. Применяем формулы для нелинейной функции. Подставляя в исходную формулу математические ожидания аргументов, имеем математическое ожидание величины X $m_x = 1.8 \cdot 0.1 \cdot 1000 = 180 м$.

Для определения среднего квадратического отклонения величины X найдем частные производные:

$$\left[\frac{\partial X}{\partial X_н}\right]_m = 1, \quad \left[\frac{\partial X}{\partial \omega}\right]_m = 1.8 \cdot m_H = 1800, \quad \left[\frac{\partial X}{\partial H}\right]_m = 1.8 \cdot m_\omega = 0.18, \quad \left[\frac{\partial X}{\partial X_б}\right]_m = 1.$$

$$\sigma_x^2 = (1 \cdot 4)^2 + (180 \cdot 0.005)^2 + (0.18 \cdot 50)^2 + (13)^2 + 2 \cdot 1 \cdot 1800 \cdot 0.5 \cdot 0.4 \cdot 0.05 +$$
$$+ 2 \cdot 1 \cdot 0.18 \cdot 0.3 \cdot 4 \cdot 50 + 24 \cdot 1800 \cdot 40.18 \cdot 0.4 \cdot 0.0005 \cdot 50 =$$
$$= 16 + 81 + 81 + 9 + 36 + 21.6 + 64.8 = 307.4 м^2 \ ; \ \sigma_x = 17.6 м.$$

Пример 4.4. На окружности радиуса r случайным образом располагаются две точки, которые затем соединяются между собой и с центром окружности. Найти математическое ожидание площади полученного треугольника.

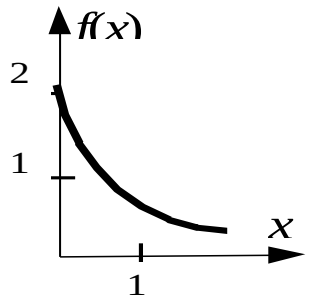
Решение. Если зафиксировать одну точку, то другая полностью определяется равномерно распределенным случайным углом φ, образованным радиусами, проведенными к этим точкам из центра круга.

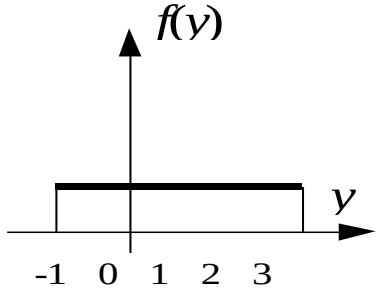
Имеем $f(\varphi) = \dfrac{1}{2\pi}, \quad \varphi \in [0, 2\pi], \quad S = \dfrac{r^2}{2}|\sin\varphi|$.

Тогда математическое ожидание площади треугольника

$$M[S] = \frac{1}{2} r^2 \int_0^{2\pi} |\sin\varphi| \frac{1}{2\pi} d\varphi = \frac{r^2}{2\pi} \int_0^{\pi} \sin\varphi d\varphi = \frac{r^2}{\pi}.$$

Пример 4.5. По виду законов распределения двух независимых случайных величин найти математическое ожидание, дисперсию и закон распределения их суммы.





Решение. Имеем $f(x) = 2e^{-2x}$, $x \ge 0$; $m_x = \frac{1}{2}$; $D_x = \frac{1}{4}$;

$$f(y) = \frac{1}{4}; \quad -1 \le y \le 3; \quad m_y = 1; \quad D_y = \frac{4}{3}; \quad Z = X + Y.$$

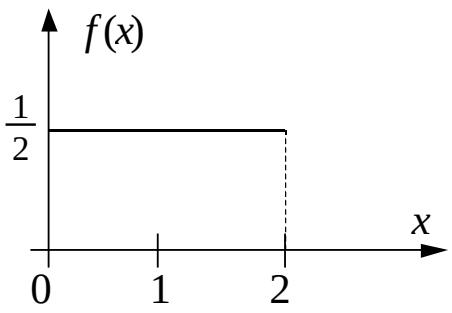
Используем композицию законов распределения

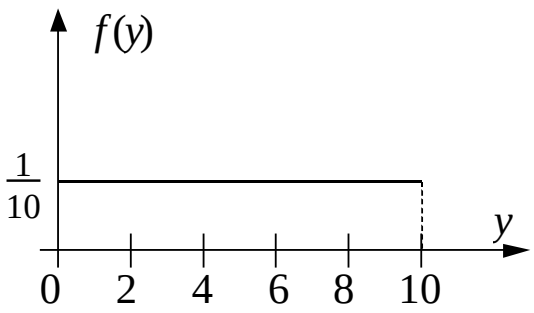
$$f(z) = \int_{-\infty}^{\infty} f(x) f_y(z-x)dx = 0{,}5\int_{0}^{z+1} e^{-2x}dx + 0{,}5\int_{z-3}^{z+1} e^{-2x}dx = -0{,}25e^{-2x}\,|_0^{z+1} - 0{,}25e^{-2x}\,|_{z-3}^{z+1}$$

$$f(z) = \begin{cases} -0{,}25(e^{-2(z+1)} - 1), & -1 \le z \le 3 \\ -0{,}25(e^{-2(z+1)} - e^{-2(z-3)}), & z \ge 3 \\ 0, & z \le -1 \end{cases}$$

$m_z = m_x + m_y = 3/2$; $D_z = 19/12$.

Пример 4.6. По известным законам распределения X и Y найти закон распределения их суммы





Решение. По формуле свертки имеем $f(z) = \int_{-\infty}^{\infty} f(x) f(z-x)dx.$

$$f(x) = \frac{1}{2}; \quad 0 \le x \le 2; \quad f(y) = \frac{1}{10}; \quad 0 \le y \le 10; \quad Z = X + Y.$$

Для нахождения этого интеграла разобьем интервал на 5 частей (рис.4.3).

1. При $z < 0$ $f(z-x) = 0$ $\Rightarrow$ $f(z) = 0.$

2. При $0 \le z \le 2$ $f(z) = \frac{1}{2}\int_0^z f(z-x)dx = \frac{z}{20}.$

3. При $2 \le z \le 10$ $f(z) = \frac{1}{2}\int_0^2 f(z-x)dx = \frac{1}{10}.$

4. При $10 \le z \le 12$ $f(z) = \frac{1}{2}\int_{z-10}^{2} f(z-x)dx = \frac{12-z}{20}.$

5. При $z > 12$ $f(z) = 0.$

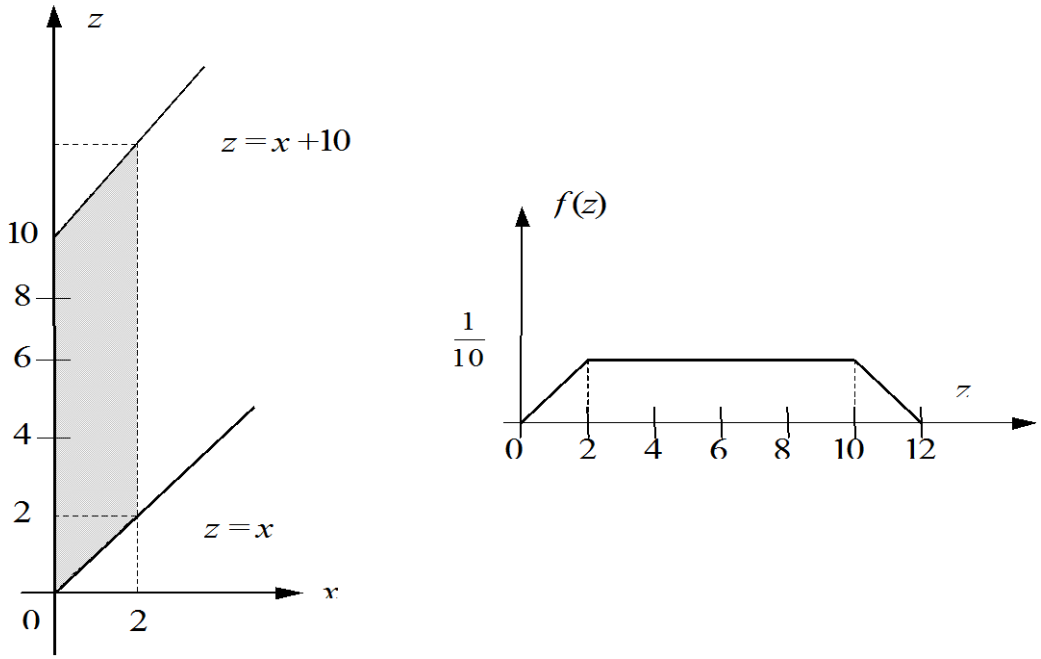


Рис.4.1. Графическое представление области интегрирования и результата свертки

Окончательно получаем $f(z)=\begin{cases} 0, & z<0 \\ \dfrac{z}{20}, & 0\le z\le 2 \\ \dfrac{1}{10}, & 2\le z\le 10 \\ \dfrac{12-z}{20}, & 10\le z\le 12 \\ 0, & z>12. \end{cases}$.

Контрольные задачи к главе 4 «Функции случайных величин»

4.1. Из накопителя перед первой технологической операцией детали забираются на обработку регулярно через каждые 10 минут. Из накопителя перед второй технологической операцией детали забираются на обработку регулярно через каждые 30 минут. Найти в процентах коэффициент вариации суммарного времени ожидания детали в накопителях (при случайном ее попадании туда).

4.2. Известно, что $X \sim N\left(a,\sigma^2\right)$, $Y=e^X$. Найти закон распределения *Y*.

4.3. На окружности радиуса r случайным образом располагаются две точки, которые затем соединяются между собой и с центром окружности. Найти математическое ожидание площади полученного треугольника.

4.4. Пусть *X*, *Y*, *Z* – случайные величины: *X* – выручка фирмы, *Y* – ее затраты, *Z* = *X* – *Y* – прибыль.

X:

x_i	3	4	5
p_i	1/3	1/3	1/3

Y:

y_j	1	2
p_j	1/2	1/2

Найти распределения прибыли Z.

4.5. Вес гайки и болта являются нормально распределенными величинами с математическими ожиданиями 15 и 40 *гр.*и средними квадратическими отклонениями 2 и 5 *гр.*, соответственно. Ковариационный момент этих величин равен 7 *гр.*^{2}. Найти среднее квадратическое отклонение веса всего узла «гайка + болт».

4.6. Дискретные независимые СВ заданы своими распределениями:

X	2	3	Y	4	5
P	0.3	0.7	P	0.6	0.4

Найти коэффициент вариации величины $Z = X + Y$.

4.7. Независимые СВ X и Y заданы плотностью распределения:

$$f_1(x) = e^{-x}\ (0 \le x < \infty); \quad f_2(y) = \frac{1}{2}e^{-y/2}\ (0 \le y < \infty).$$

Найти плотность распределения СВ $Z = X + Y$.

4.8. Ошибка прибора выражается функцией $U = 3X - 2Y + 4Z - 5$, где X, Y, Z – так называемые первичные ошибки, представляющие собой систему случайных величин, которая характеризуется математическими ожиданиями $m_x = -1,\ m_y = 7,\ m_z = 6$ и корреляционной матрицей $\begin{Vmatrix} 2 & 1 & -1 \\ & 3 & 2 \\ & & 4 \end{Vmatrix}$. Найти среднее квадратическое отклонение ошибки прибора.

4.9. Дискретная случайная величина X имеет ряд распределения, представленный в таблице.

X	1	2	3	4
P	0,1	0,4	0,3	0,2

Найдите математическое ожидание и дисперсию величины $Y = X^2 + 1$.

4.10. Связь между нормально распределенными показателями X и Y выражается зависимостью Y = 1,3x + 0,5. При этом дисперсия Y в 4 раза выше дисперсии X. Найти степень тесноты связи величин X и Y.

4.11. Показатель Y выражается формулой $Y = 6X_1 + 3.5X_2X_3 + 5$, где X_1, X_2, X_3 представляют собой величины с математическими ожиданиями $m_{x_1} = 60,\ m_{x_2} = 2,\ m_{x_3} = 10$ и средними квадратическими отклонениями $\sigma_{x_1} = 3,\ \sigma_{x_2} = 0.1,\ \sigma_{x_3} = 2$ Нормированная корреляционная матрица системы имеет вид: $\begin{Vmatrix} 1 & 0.4 & 0.5 \\ & 1 & 0.3 \\ & & 1 \end{Vmatrix}$. Найти среднее квадратическое отклонение СВ Y.

4.12. Случайные величины X_1 и X_2 имеют математические ожидания

$m_{x_1} = 2$, $m_{x_2} = -1$, дисперсия $D_{x_1} = 3, D_{x_2} = 4$ и ковариацию $\operatorname{cov}(X_1, X_2) = -1$. Найдите математическое ожидание и дисперсию случайной величины $Y = 2X_1 - 3X_2 - 5$.

4.13. Независимые случайные величины X_1 и X_2 имеют экспоненциальное распределение с параметрами λ_1 и λ_2 соответственно. Найти математическое ожидание случайной величины $Y = X_1 X_2$.

4.14. Два стрелка независимо один от другого производят по одному выстрелу в мишень. Вероятность попадания для каждого из стрелков равна 0,6. Пусть случайные величины X и Y означают число попаданий в мишень для первого и для второго стрелка соответственно. Построить закон распределения и найти математическое ожидание для случайной величины: $Z_2 = X^Y$.

4.15. СВ X задана плотностью распределения $f(x) = (1/2)sin(x)$ в интервале $(0,\pi)$; вне интервала $f(x) = 0$. Найти дисперсию величины $Y=X^2$.

4.16. Случайная величина X задана плотностью вероятности $f(x) = x + 0,5$ в интервале (0;1); вне этого интервала $f(x) = 0$. Найти математическое ожидание функции $Y = X^3$.

4.17. Две случайные величины (X и Y) имеют характеристики: $m_x = 2;\ m_y = 3;\ \sigma_x^2 = 1;\ \sigma_y^2 = 4;\ r_{xy} = 0,5$. Определить дисперсию суммы и разности этих величин.

4.18. Две независимые случайные величины (X и Y) распределены равномерно: X на интервале[– 5; 1], Y на интервале [3; 6]. Найти математическое ожидание и дисперсию случайной величины $7X - 9Y + 8$.

4.19. Ребро куба измерено приближенно, причем $a \le x \le b$. Рассматривая ребро куба как случайную величину X, распределенную равномерно в интервале (a, b), найти математическое ожидание объема куба.

4.20. Дискретные независимые СВ заданы своими распределениями:

X	1	3	Y	2	4
P	0.3	0.7	P	0.6	0.4

Найти коэффициент вариации величины $Z = 2X + 3Y$.

4.21. Предприятие состоит из двух подразделений. Месячная прибыль каждого подразделения является нормально распределенной величиной с математическими ожиданиями 550 и 400 тыс. руб. и средними квадратическими отклонениями 60 и 50 тыс. руб., соответственно. Ковариационный момент этих величин равен 70 тыс. руб.^{2} Найти коэффициент вариации прибыли всего предприятия.

4.22. Один станок дает в среднем 3% брака, другой – 5%. Производительности станков одинаковы. Каков коэффициент вариации числа бракованных изделий в общей продукции из 100 деталей?

4.23. Две СВ (X и Y) имеют характеристики:

$$m_x = 2;\ m_y = 3;\ \sigma_x^2 = 1;\ \sigma_y^2 = 4;\ r_{xy} = 0,5.$$

Определить дисперсию суммы этих величин.

4.24. Вес гайки и болта являются нормально распределенными величинами с математическими ожиданиями 10 и 40 *гр.* и средними квадратическими отклонениями 2 и 5 гр., соответственно. Ковариационный момент этих величин равен 7 гр.^{2}. Найти среднее квадратическое отклонение веса всего узла «гайка + болт».

4.25. Погрешность в изготовлении детали образуются в результате суммарного воздействия трех факторов *A*, *B* и *C*. Их характеристики известны:

$$\sigma_A = \sigma_B = 0,1; \quad \sigma_C = 0,2; \quad m_A = 0,9; \quad m_B = 2,0; \quad m_C = 3;$$

$$r_{AB} = 0,4; \quad r_{AC} = 0,1; \quad r_{BC} = 0,2.$$

Найти среднее квадратичное отклонение погрешности изготовления детали.

4.26. Даны законы распределения двух независимых случайных величин X и Y

X	–1	0	1	2
P	0,2	0,3	0,1	0,4

Найти характеристики $M(XY - 5X + 2Y - 7)$, $D(8X - 3Y + 4)$.

4.27. Независимые СВ X и Y заданы плотностью распределения:

$$f_1(x) = e^{-x} \ (0 \le x < \infty); \quad f_2(y) = \frac{1}{2} e^{-y/2} \ (0 \le y < \infty).$$

Найти плотность распределения СВ $Z = X + Y$.

Задачи, предлагаемые для решения на практических занятиях по первому разделу курса «Теория вероятностей»

Глава 5. «Предельные законы теории вероятностей».

Пример 5.1. Дана случайная величина X с математическим ожиданием m_x и дисперсией σ_x^2. Оценить сверху вероятность того, что величина X отклонится от своего математического ожидания не меньше, чем на $3\sigma_x$.

Решение. Полагая во втором неравенстве Чебышева $\varepsilon = 3\sigma_x$, получим

$$P\left[\left|X - m_x\right| \ge 3\sigma_x\right] \le \frac{D_x}{9\sigma_x^2} = \frac{1}{9},$$

т.е. вероятность того, что отклонение случайной величины от ее математического ожидания выйдет за пределы трех средних квадратических отклонений, не может быть больше 1/9 ни при каком законе распределения.

Замечание. На практике в большинстве случаев вероятность того, что величина X выйдет за пределы участка $m_x \pm 3\sigma_x$, значительно мень-

ше 1/9. Например, для нормального закона эта вероятность приблизительно равна 0.003.

Пример 5.2. Среднее значение скорости ветра в данной местности равно 16 км/час. Оценить вероятность того, что в данной местности скорость ветра (при одном наблюдении) не превышает 80 км/час.

Решение. По первой форме неравенства Чебышева находим

$$P(X \geq 80) \leq \frac{M[X]}{80} = 1/5 = 0.2, \quad \text{т.е.} \quad P(X \leq 80) \geq 4/5 = 0.8.$$

Пример 5.3. Математическое ожидание скорости ветра на данной высоте равно 25 км/час, а $\sigma_x = 4{,}5$ км/час. Какие скорости ветра можно ожидать на этой высоте с вероятностью не меньшей чем 0.9?

Решение. Пусть X - скорость ветра. Тогда по второму неравенству Чебышева имеем $P(|X - 25| \leq \varepsilon) \geq 1 - \left(\frac{4.5}{\varepsilon}\right)^2 > 0.9$, т.е. $\varepsilon > 14.2\ \frac{\text{км}}{\text{час}}$.

Следовательно, с вероятностью, большей 0.9, имеем $10.8 \leq X \leq 39.2\ \frac{\text{км}}{\text{час}}$.

Пример 5.4. 4 станка производят детали из стали марки A, 6 других – из стали марки B. Определить вероятность того, что из 500 взятых деталей количество деталей из стали марки A будет заключено в пределах от 180 до 220.

Решение. Воспользуемся неравенством Чебышева для оценки нижней границы искомой вероятности $P[|X - a| \leq \varepsilon] \geq 1 - \frac{D[X]}{\varepsilon^2}$.

Имеем $a = np$; $n = 500$; $p = 0.4$; $q = 1 - p = 0.6$; $a = 500 \cdot 0.4 = 200$; $\varepsilon = 20$.

$D[X] = npq = 500 \cdot 0.4 \cdot 0.6 = 120$.

$$P[|X - 200| \leq 20] \geq 1 - \frac{120}{20^2} = 1 - \frac{120}{400} = \frac{280}{400} = 0.7.$$

С другой стороны, эту вероятность можно более точно вычислить (оценить) по теореме Муавра–Лапласа:

$$P(m_1 \leq m \leq m_2) = \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right),$$

$$P(180 \leq m \leq 220) = \Phi\left(\frac{220 - 500 \cdot 0{,}4}{\sqrt{500 \cdot 0{,}4 \cdot 0{,}6}}\right) - \Phi\left(\frac{180 - 500 \cdot 0{,}4}{\sqrt{500 \cdot 0{,}4 \cdot 0{,}6}}\right) =$$

$$= 2\Phi(1{,}86) = 2 \cdot 0{,}4666 = 0.9332.$$

Пример 5.5. Станок–автомат требует подналадки в среднем один раз за 4 часа работы. Определить вероятность того, что за 10 суток непрерывной работы подналадка осуществлялась ровно 70 раз.

Решение. Имеем $n = 240$; $m = 70$; $p = 0.25$; $q = 0.75$.

По локальной теореме Муавра–Лапласа получаем

$$P_{70,240} = \frac{1}{\sqrt{240 \cdot 0{,}25 \cdot 0{,}75}} f\left(\frac{70 - 240 \cdot 0{,}25}{\sqrt{240 \cdot 0{,}25 \cdot 0{,}75}}\right) = \frac{1}{6{,}7} f(1{,}49) = \frac{1}{6{,}7} 0{,}132 = 0{,}020 .$$

По закону Пуассона эта вероятность $P_{70} = \frac{60^{70}}{70!} e^{-60} = 0{,}022$.

Пример 5.6. В страховой фирме застраховано 10 тысяч лиц одного возраста и одной социальной группы. Вероятность смерти в течение года для каждого лица равна 0,006. Каждый застрахованный в начале года вносит 12 долларов страховых, и, в случае смерти, его родственники получают от фирмы тысячу долларов. Найти вероятность того, что:

1) фирма потерпит убыток;

2) фирма получит прибыль, не меньшую чем x тысяч долларов $(x = 40)$.

Решение. Вероятность убытков для страховой фирмы есть вероятность смерти в течении года более чем 120 застрахованных.

Тогда по формуле Муавра–Лапласа имеем

$$P(m > 120) = P(120 < m < 10000) = \Phi\left(\frac{10000 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{120 - np}{\sqrt{npq}}\right).$$

У нас $n = 10000$; $p = 0.006$; $q = 1 - p = 0.994$.

Тогда с точностью до 10 знаков после запятой

$$P(m > 120) = \Phi\left(\frac{10000\text{-}60}{\sqrt{10000 \cdot 0.006 \cdot 0.994}}\right) - \Phi\left(\frac{120\text{-}60}{\sqrt{59.64}}\right) = 0.5 - \Phi(7.78) \approx 0 .$$

Получение прибыли в x тысяч долларов и более может быть, если в течении года из застрахованных умрёт не более чем $\frac{120000 - x \cdot 1000}{1000} = (120 - x)$ человек. В этом случае вероятность получения фирмой прибыли ($П$) не менее величины x равна:

$$P(П \geq x) = P(m \leq 120 - x) = P(0 \leq m \leq 120\text{-}x) = \Phi\left(\frac{120 - np - x}{\sqrt{npq}}\right) - \Phi\left(\frac{0 - np}{\sqrt{npq}}\right).$$

При $x = 40$ получим:

$$P(П \geq 40) = \Phi\left(\frac{80 - 60}{\sqrt{59.64}}\right) - \Phi\left(\frac{-60}{\sqrt{59.64}}\right) = \Phi(2.59) + \Phi(7.78) = 0.995 .$$

Пример 5.7. Три станка, производительности которых соотносятся как 5:3:2, производят детали на общий конвейер. Определить вероятность того, что из 240 деталей, взятых случайным образом с конвейера, деталей, произведенных вторым станком будет от 60 до 70.

Решение. Искомую вероятность определяем по интегральной формуле Муавра–Лапласа:

$$P(m_1 \le m \le m_2) = \Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right), \qquad \text{где } \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{e}}dt.$$

Имеем $p = 0{,}3$; $q = 1 - 0{,}3 = 0{,}7$; $m_1 = 60$; $m_2 = 70$; $n = 240$. Тогда

$$P(m_1 \le m \le m_2) = \Phi(-0{,}3) - \Phi(-1{,}7) = \Phi(1{,}7) - \Phi(0{,}3) = 0{,}4554 - 0{,}1179 = 0{,}3375.$$

Пример 5.8. Вероятность изделию быть бракованным равна 0,05. Сколько изделий надо взять, чтобы с вероятностью не меньшей 0,9 среди них оказалось не менее 50 бракованных?

Решение. Необходимо найти число n, удовлетворяющее интегральной формуле Муавра–Лапласа

$$P(50 \le m \le n) = \Phi\left(\frac{n - 0{,}05\cdot n}{\sqrt{n\cdot 0{,}05\cdot(1-0{,}05)}}\right) - \Phi\left(\frac{50 - 0{,}05\cdot n}{\sqrt{n\cdot 0{,}05\cdot(1-0{,}05)}}\right) \ge 0{,}9\ .$$

Оценим значение

$$\Phi\left(\frac{n - 0{,}05\cdot n}{\sqrt{n\cdot 0{,}05\cdot(1-0{,}05)}}\right) = \Phi\left(\frac{0{,}95\cdot n}{0{,}022\cdot\sqrt{n}}\right) \approx \Phi(4{,}3\sqrt{n}) \ge \Phi(30) \approx 0{,}5.$$

Тогда $\Phi\left(\frac{50 - 0{,}05\cdot n}{\sqrt{n\cdot 0{,}05\cdot(1-0{,}05)}}\right) \le -0{,}4.$

По таблице функции Лапласа находим, что $\Phi(x) = -0{,}4$ при $x = 1{,}28$.

Поэтому получаем соотношение $\frac{50 - 0{,}05\cdot n}{\sqrt{n\cdot 0{,}05\cdot(1-0{,}05)}} \le 1{,}28$

или $\quad 0{,}05\cdot n - 0{,}282\cdot\sqrt{n} - 50 \ge 0.$

Решая последнее неравенство, находим $n \ge 1196$, то есть следует взять не менее 1196 изделий.

Контрольные задачи к главе 5
«Предельные теоремы теории вероятностей»

5.1. Среднее потребление электроэнергии в мае в некотором населенном пункте составляет 360 000 кВт·ч. Оцените с помощью первого неравенства Чебышева вероятность того, что потребление электроэнергии в мае текущего года в этом населенном пункте превысит 10^6 кВт·ч.

5.2. Средний ежедневный расход воды в некотором населенном пункте составляет 50 000 *л*. Оценить с помощью первого Чебышева вероятность того, что в произвольно выбранный день расход воды в этом пункте превысит 150 000 *л*.

5.3. Среднее квадратическое отклонение погрешности изменения курса самолета равно 2^0. Считая математическое ожидание погрешности измерения

равным нулю, оцените с помощью второго неравенства Чебышева вероятность того, что погрешность одного измерения курса самолета превысит 5^0.

5.4. Вероятность появления события A в одном опыте равна 0,5. Можно ли с вероятностью, большей 0,97, утверждать, что число появлений события A в 1000 независимых опытов будет в пределах от 400 до 600?

5.5. Средний расход воды в населенном пункте составляет 50000 литров в день. Оценить вероятность того, что в этом населенном пункте в данный день расход воды не превысит 150000 литров.

5.6. Вероятность наступления события A в каждом испытании равна 2/3. Используя неравенство Чебышева, оценить вероятность того, что в 10000 испытаний отклонение относительной частоты события A от вероятности его появлений не превзойдет 0,01.

5.7. Вопрос#19: Среднее квадратическое изменение курса акции компании в течение одних биржевых торгов составляет 0,3%. Оценить вероятность того, что на ближайших торгах курс изменится более чем на 3%.

5.8. Отделение банка обслуживает в среднем 100 клиентов в день. Оценить вероятность того, что сегодня в отделении банка будет обслужено не более 200 клиентов.

5.9. Электростанция обслуживает сеть на 1600 электроламп, вероятность включения каждой из которых равна 0,9. Оценить с помощью неравенства Чебышева вероятность того, что число ламп, включенных в сеть вечером, отличается от своего математического ожидания не более чем на 100 (по абсолютной величине).

5.10. Вероятность того, что акции, переданные на депозит, будут востребованы, равна 0,08. Оценить с помощью неравенства Чебышева вероятность того, что среди 1000 клиентов от 70 до 90 востребуют свои акции.

5.11. Среднее значение длины детали 50 *см*, а дисперсия – 0,1. Оценить с помощью неравенства Чебышева вероятность того, что случайно взятая деталь окажется по длине не менее 49,5 *см* и не более 50,5 *см*. Уточнить вероятность того же события, если известно, что длина случайно взятой детали имеет нормальный закон распределения.

5.12. В течение времени t эксплуатируются 500 приборов. Каждый прибор имеет надежность 0,98 и выходит из строя независимо от других. Оценить с помощью неравенства Чебышева вероятность того, что доля надежных приборов отличается от 0,98 не более чем на 0,1 (по абсолютной величине).

5.13. Вероятность сдачи в срок всех экзаменов студентом факультета равна 0,7. Оценить с помощью неравенства Чебышева вероятность того, что доля сдавших в срок все экзамены из 2000 студентов заключена в границах от 0,66 до 0,74.

5.14. В среднем 10% работоспособного населения некоторого региона – безработная. Оценить с помощью неравенства Чебышева вероятность того, что уровень безработицы среди обследованных 10000 работоспособных жителей города будет в пределах от 9 до 11% включительно.

5.15. Сколько нужно провести измерений, чтобы с вероятностью, равной

0,9973, утверждать, что погрешность средней арифметической результатов этих измерений не превысит 0,01, если измерение характеризуется средним квадратическим отклонением, равным 0,03?

5.16. Случайная величина X является средним арифметическим из n независимых одинаково распределенных случайных величин, дисперсия каждой из которых равна 5. Оцените, какое число слагаемых n нужно взять для того, чтобы с вероятностью не менее 0,9973 случайная величина X отклонялась от своего среднего не более чем на 0,01.

5.17. Среднее квадратическое отклонение погрешности изменения курса самолета равно 2^0. Считая математическое ожидание погрешности измерения равным нулю, оцените с помощью второго неравенства Чебышева вероятность того, что погрешность одного измерения курса самолета превысит 5^0.

5.18. Случайная величина X является средним арифметическим из 3200 независимых случайных величин, причем каждое слагаемое имеет математическое ожидание 3 и дисперсию 2. Оценим вероятность того, что X попадает в интервал (2,925, 3,075).

5.19. Производится выборочное обследование большой партии электрических лампочек. Среднее квадратическое отклонение времени горения лампочки равно $\sigma_x = 80\,ч$. Из всей партии наудачу выбирается 400 лампочек. Оценить вероятность того, что среднее время горения лампочки будет отличаться от наблюдаемого среднего времени горения выбранных 400 лампочек не более чем на 10 *ч*.

5.20. Три станка, производительности которых соотносятся как 3:2:1,производят детали на общий конвейер. Определить вероятность того, что из 200 деталей, взятых случайным образом с конвейера, деталей, произведенных вторым станком будет от 50 до 70.

5.21. 3 станка производят детали из стали марки *A* , 7 других – из стали марки *B*. Определить вероятность того, что из 400 взятых деталей количество деталей из стали марки *A* будет заключено в пределах от 100 до 130.

5.22. Отделение банка обслуживает в среднем 100 клиентов в день. Оценить вероятность того, что сегодня в отделении банка будет обслужено более 150 клиентов.

5.23. Три станка, производительности которых соотносятся как 5:3:2, производят детали на общий конвейер. Определить вероятность того, что из 200 деталей, взятых случайным образом с конвейера, деталей, произведенных вторым станком будет от 50 до 65.

5.24. Найти вероятность того, что при 720 бросаниях игральной кости «шестерка» выпадает от 100 до 130 раз.

5.25. Вероятность появления некоторого события в каждом из 800 независимых испытаний равна 1/4. Воспользовавшись вторым неравенством Чебышева, оцените вероятность того, что число X появлений этого события заключено в пределах от 150 до 250.

Глава 6 «Характеристические функции случайных величин»

Пример 6.1. Найти ХФ для распределений: биноминального, Пуассона равномерного, нормального.

Решение.

1) ХФ дискретной величины X с возможными значениями $x_1, x_2, ..., x_n$ и их вероятностями P_j $(j=1,...,n)$ определяются формулой:

$$q(t)=\sum_{j=1}^{n} e^{itx_j} P_j .$$

Тогда для биномиального распределения имеем:

$$q(t)=\sum_{m=0}^{n} C_n^m p^m q^{n-m} e^{itm} = \left(pe^{it}+q\right)^n .$$

2) Для распределения Пуассона

$$q(t)=\sum_{m=0}^{\infty}\frac{a^m}{m!}e^{-a}e^{itm}=e^{-a}\sum_{m=0}^{\infty}\frac{\left(ae^{it}\right)^m}{m!}=e^{-a}e^{ae^{it}}=e^{a(e^{it}-1)},$$ так как $e^x=\sum_{s=0}^{\infty}\frac{x^s}{s!}$.

3) Для равномерного распределения в интервале (a,b)

$$q(t)=\int_a^b \frac{e^{itx}}{b-a}dx=\frac{e^{itb}-e^{ita}}{it(b-a)}.$$

В частности, ХФ равномерно распределенной СВ на $[-a,\ a]$:

$$q(t)=\frac{1}{2a}\int_{-a}^{a} e^{itx}dx=\frac{\sin(at)}{at}.$$

4) Для нормальной СВ имеем:

$$q(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{itx-\frac{(x-m)^2}{2\sigma^2}}dx=\frac{e^{itm}}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{it(x-m)-\frac{(x-m)^2}{2\sigma^2}}dx=e^{itm-\frac{\sigma^2t^2}{2}},$$

так как $\int_{-\infty}^{\infty} e^{at-\frac{1}{2}bt^2}dt=\sqrt{\frac{2\pi}{b}}e^{\frac{a^2}{2b}}$.

Для n-мерного нормального распределения случайного вектора X

$$q(t)=\frac{1}{\sigma\sqrt{(2\pi)^n|R|}}\int_{-\infty}^{\infty} e^{it^Tx-\frac{1}{2}\left(x^T-m^T\right)R^{-1}(x-m)}dx .$$

Так как $\int\limits_{-\infty}^{\infty} e^{a^T t - \frac{1}{2} t^T bt} dt = \sqrt{\frac{(2\pi)^n}{|b|}} e^{-\frac{1}{2} a^T b^{-1} a}$, то получим $q(t) = e^{it^T m - \frac{1}{2} t^T Rt}$.

Пример 6.2. Получить выражение для всех центральных моментов нормальной СВ.

Решение. Для нормальной СВ ХФ имеет вид $q(t) = e^{itm - \frac{\sigma^2 t^2}{2}}$.

С другой стороны ХФ выражается через центральные моменты по соотношению

$$q(t) = e^{itm}\left[1 + \sum_{k=2}^{v} \frac{i^k \mu_k}{k!} t^k + R'_v\right].$$

Приравнивая эти выражения и учитывая, что $e^x = \sum_{s=0}^{\infty} \frac{x^s}{s!}$, $e^{-x} = \sum_{s=0}^{\infty} \frac{(-1)^s x^s}{s!}$, получим:

$$e^{-\frac{\sigma^2 t^2}{2}} = 1 + \sum_{k=2}^{v} \frac{i^k \mu_k}{k!} t^k + R'_v; \qquad \sum_{s=0}^{\infty} \frac{(-1)^s \sigma^{2s}}{2^s s!} t^{2s} = 1 + \sum_{k=2}^{\infty} \frac{i^k \mu_k}{k!} t^k;$$

$$\sum_{s=0}^{\infty} \frac{i^{2s} \sigma^{2s}}{2^s s!} t^{2s} = 1 + \sum_{k=2}^{\infty} \frac{i^k \mu_k}{k!} t^k; \qquad k = 2s;$$

$$\mu_{2s-1} = 0; \qquad \mu_{2s} = \frac{(2s)!}{2^s s!} \sigma^{2s}; \qquad \mu_0 = 1.$$

Пример 6.3. Найти ХФ суммы двух независимых распределений Пуассона с параметрами *a* и *b*.

Решение. ХФ суммы $Z = X + Y$ двух независимых СВ равна произведению соответствующих ХФ. Поэтому получаем:

$$q_Z(t) = e^{a(e^{it} - 1)} e^{b(e^{it} - 1)} = e^{(a+b)(e^{it} - 1)}.$$

Полученное выражение соответствует форме ХФ пуассоновского распределения с параметрами $(a + b)$, то есть сумма двух законов Пуассона дает снова закон Пуассона с соответствующими параметрами.

Пример 6.4. Найти распределение суммы всех независимых СВ X и Y, распределенных равномерно в интервалах $(-a, a)$ и $(-b, b)$ соответственно $(a < b)$.

Решение. $Z = X + Y$.

$$f(z)=\frac{1}{8\pi ab}\int_{-\infty}^{\infty}e^{-itz}dt\int_{-a}^{a}\int_{-b}^{b}e^{it(x+y)}dxdy=\frac{1}{2\pi ab}\int_{-\infty}^{\infty}e^{-itz}\frac{\sin(ta)\sin(tb)}{t^2}dt=$$

$$=\frac{1}{\pi ab}\int_{-\infty}^{\infty}\frac{\sin(ta)\sin(tb)\cos(tz)}{t^2}dt=$$

$$=\frac{1}{4\pi ab}\int_{-\infty}^{\infty}\left[\cos(a-b+z)t+\cos(a-b-z)t-\cos(a+b+z)t-\cos(a+b-z)t\right]\frac{dt}{\lambda^2}.$$

Отсюда на основании известной формулы $\int_0^{\infty}\frac{1-\cos(au)}{u^2}du=\frac{\pi}{2}|a|$ получаем:

$$f(z)=\frac{1}{8ab}\left(|a+b+z|+|a+b-z|-|a-b+z|-|a-b-z|\right).$$

Контрольные задачи к главе 6
«Характеристические функции случайных величин»

6.1. Ряд распределения случайной величины *X* представлен таблицей:

X	–2	0	2
P	¼	½	¼

Найти характеристическую функцию случайной величины *X*.

6.2. Найдите характеристическую функцию непрерывной случайной величины X, имеющей плотность распределения $f(x)=\frac{\sqrt{2}}{\pi(1+x^4)}$

6.3. Найдите характеристическую функцию случайной величины X, ряд распределения которой представлен в таблице:

X	0	1	2	3
P	$1/2$	$1/8$	$1/4$	$1/8$

6.4. Найти характеристическую функцию случайной величины *X*, имеющей равномерное на интервале (a,b) распределение:

6.5. Найдите плотность распределения случайной величины, имеющей характеристическую функцию

$$f(t)=\begin{cases}0, & |t|\geq 1,\\ 1-|t|, & |t|<1.\end{cases}$$

6.6. Найдите характеристическую функцию непрерывной случайной величины X, имеющей плотность распределения

6.7. Независимые случайные величины X_1 и X_2 распределены по экспоненциальному закону с параметрами λ_1 и λ_2. Найти характеристическую функцию случайной величины $Y=X_1+X_2$.

6.8. Случайная величина *X* имеет плотность распределения $f(x)=\frac{e^{-|x|}}{2}$.

Найти характеристическую функцию случайной величины X.

6.9. Найдите характеристическую функцию случайной величины X, ряд распределения которой представлен в таблице:

X	0	1	2	3
P	$1/2$	$1/8$	$1/4$	$1/8$

6.10. Найдите характеристическую функцию непрерывной случайной величины X, имеющей плотность распределения $f(x)=\begin{cases}0, & |x|\geq 1;\\ \dfrac{3(1-x^2)}{4}, & |x|<1.\end{cases}$

6.11. Найдите характеристическую функцию неотрицательной целочисленной случайной величины X, распределение которой задается вероятностями

$$p\{X=n\}=(1+n)p^2(1-p)^n,\ n=0,1,2,...;\ \ 0<p<1$$

6.12. Случайная величина X_1 распределена равномерно на интервале $(0,1)$, а случайная величина X_2 имеет стандартное нормальное распределение. Найдите характеристическую функцию случайной величины $Y=X_1+X_2$, если известно, что X_1 и X_2 являются независимыми.

6.13. Найти закон распределения случайной величины, характеристическая функция которой равна $f(t)=\dfrac{1+\cos t}{2}$.

6.14. Найдите характеристическую функцию случайной величины $Y=aX+b$, где X – случайная величина, имеющая плотность распределения

$$f(x)=\frac{\sqrt{2}}{\pi(1+x^4)}.$$

6.15. Найдите математическое ожидание и дисперсию случайной величины X, имеющей характеристическую функцию $f(t)=\begin{cases}\dfrac{\sin t}{t}, & t\neq 0;\\ 1, & t=0.\end{cases}$

6.16. Найти математическое ожидание и дисперсию случайной величины X, имеющей характеристическую функцию $f(t)=\dfrac{1}{1+2it}$.

6.17. Найти закон распределения случайной величины X, характеристическая функция которой равна $f(t)=e^{-|t|}$.

Задачи, предлагаемые для решения на практических занятиях по второму разделу курса «Случайные процессы»

Пример 2.1. Являются ли периодическими процессы

$$X(t) = a_1\sin(2t + \Theta_1) + a_2\sin(3t + \Theta_2) + a_3\sin(7t + \Theta_3);$$
$$Y(t) = a_1\sin(2t + \Theta_1) + a_2\sin(3t + \Theta_2) + a_3\sin(\sqrt{50}\,t + \Theta_3)\ ?$$

Решение. Сумма нескольких синусоидальных образует периодический процесс только в том случае, если отношения всех возможных пар частот представляет собой рациональные числа. Это означает, что существует некоторый основной период, удовлетворяющий формуле: $X(t) = X(t \pm nT_0)$, $n = 1,2,3\ldots$

Поэтому процесс $X(t)$ периодический, поскольку 2/3, 2/7, 3/7 – рациональные числа (с основным периодом равным 1).

Процесс $Y(t)$ не является периодическим, поскольку числа $2/\sqrt{50}$, $3/\sqrt{50}$ иррациональные (и основной период равен бесконечности). В этом случае процесс является почти периодическим, но соотношение, записанное для $X(t)$ не удовлетворяется при любых конечных значениях T_0.

Пример 2.2. Периодический процесс формируется в результате сложения трех синусоидальных волн с частотами 60, 75, 100 Гц. Определить основной период этого процесса. Как будет выглядеть ряд Фурье этого процесса?

Решение. Наибольший общий делитель указанных частот равен пяти, поэтому период результирующего периодического процесса составляет 0,2 секунд. Следовательно, при разложении в ряд Фурье значения c_n будут равны нулю при всех n, кроме $n = 12$, $n = 15$, $n = 20$.

Пример 2.3. Элементарная случайная функция имеет вид $Y(t) = aX + t$, где X – случайная величина с характеристиками m и σ; a – неслучайная величина. Требуется найти характеристики $Y(t)$.

Решение. Все характеристики выразим по их определению.

$$m_y(t) = aM[X] + t = am_x + t;$$

$$\overset{\circ}{Y}(t) = aX + t - am_x - t = a\overset{\circ}{X};$$

$$R_y(t_1, t_2) = M[\overset{\circ}{Y}(t_1)\overset{\circ}{Y}(t_2)] = M[a\overset{\circ}{X}\,a\overset{\circ}{X}] = a^2\sigma_x{}^2;$$

$$D_y(t) = R_y(t,t) = a^2\sigma^2.$$

Пример 2.4. Случайная функция имеет вид $Y(t) = Xe^{-t}$; $(t > 0)$, где X – случайная величина, распределенная по нормальному закону с параметрами m и σ. Найти основные характеристики функции $Y(t)$.

Решение. Математическое ожидание и дисперсию функции $Y(t)$ выразим по их определению:

$$m_y(t) = M[Xe^{-t}] = e^{-t}M[X] = me^{-t};$$

$$D_y(t) = D[Xe^{-2t}] = \sigma^2 e^{-2t}.$$

Найдем корреляционную функцию.

Центрируем функцию $Y(t)$: $\overset{\circ}{Y}(t) = Y(t) - m_y(t) = \overset{\circ}{X}e^{-t}$.

Тогда $R_y(t_1,t_2) = M[\overset{\circ}{Y}(t_1)\overset{\circ}{Y}(t_2)] = M[\overset{\circ}{X}e^{-t_1}\,\overset{\circ}{X}e^{-t_2}] = \sigma^2 e^{-(t_1+t_2)}$.

Пример 2.5. Мощность угольного пласта с математическим ожиданием (средним значением) 3,4 м является нормальной стационарной случайной функцией по направлению отработки с АКФ $R_x(r) = ae^{-b|r|}\cos(cr)$ (r – в метрах). Скорость продвижения забоя равна 1,5 м/ч. В текущий момент обработки пласта его мощность равна 4 м. Определить вероятность того, что через 2 часа работы мощность пласта будет больше 4,4 м, если a = 2 м^2; b = 0,1 м^2; c = 0,2 м$^{-1}$.

Решение. Обозначим $x_1 = X(t_1)$ = 4 м; $x_2 = X(t_1 + 2)$. Для условного закона распределения x_2 имеем $f(x_2 \mid x_1 = 4) = \dfrac{f(x_1, x_2)}{f(x_1)}$,

где $f(x_1,x_2)$ – нормальный закон распределения системы случайных величин с корреляционной матрицей $\begin{Vmatrix} R_x(0) & R_x(\tau) \\ R_x(\tau) & R_x(0) \end{Vmatrix}$,

τ = (2 часа)·(1,5 м/час) = 3 м.

Искомая вероятность определиться следующим образом:

$$P = \int_{4.4}^{\infty} f(x_2 \mid x_1 = 4) dx_2 .$$

Условный закон распределения определяется выражением:

$$f(y \mid x) = \frac{1}{\sigma_y \sqrt{2\pi}\sqrt{1-r^2}} e^{-\frac{\left[y-\bar{y}-r\frac{\sigma_y}{\sigma_x}(x-\bar{x})\right]^2}{2(1-r^2)\sigma_y{}^2}} .$$

Вычислим неизвестные параметры этой формулы.

Имеем $r = e^{-0,1|3|}\cos(0,2\cdot 3) = 0,74\cdot 0,83 = 0,61$. Тогда

$$f(x_2 \mid x_1 = 4) = \frac{1}{\sqrt{2}\sqrt{2\pi}\sqrt{1-0,61^2}} e^{-\frac{\left[x_2-3,4-0,61(4-3,4)^2\right]^2}{2(1-0,61^2)2}} = \frac{1}{1,6\sqrt{\pi}} e^{-\frac{[x_2-3,77]^2}{2\cdot 1,26}} .$$

Используя таблицу функции распределения нормального закона определяем искомую вероятность

$$P = \int_{4,4}^{\infty} f(x_2 \mid x_1 = 4) dx_2 = \frac{1}{2} - \Phi\left(\frac{4,4-3,77}{\sqrt{1,26}}\right) = \frac{1}{2} - \Phi(0,57) \approx 0,28.$$

Пример 2.6.

На вход дифференцирующего механизма поступает случайный сигнал $X(t)$ с математическим ожиданием

$$m_x(t) = \sin(t)$$

и корреляционной функцией

$$R_x(t_1, t_2) = \sigma_x^2 e^{-a(t_2-t_1)^2} .$$

Определить математическое ожидание и дисперсию сигнала на выходе системы.

Решение. Случайная функция $Y(t)$ на выходе системы связана с $X(t)$ оператором дифференцирования $Y(t) = \dfrac{dX(t)}{dt}$.

Применяя общие правила, имеем:

$$m_y(t) = \frac{dm_x(t)}{dt} = \cos(t);$$

$$R_y(t_1,t_2)=\frac{\partial^2 R_x(t_1,t_2)}{\partial t_1 \partial t_2}=\frac{\partial}{\partial t_1}[\sigma_x^2 e^{-a(t_2-t_1)^2}(-2at_2+2at_1)]=$$

$$=\frac{\partial}{\partial t_1}[-2a\sigma_x^2 t_2 e^{-a(t_2-t_1)^2}+2a\sigma_x^2 t_1 e^{-a(t_2-t_1)^2}]=$$

$$=2\sigma_x^2 a e^{-a(t_2-t_1)^2}[1-2a(t_2-t_1)^2].$$

Полагая $t_2=t_1$, имеем

$$\sigma_y^2(t)=2a\sigma_x^2,$$

т.е. дисперсия на выходе является постоянной величиной.

Пример 2.7. Может ли быть при каких–либо значениях аргументов:

1) Функция распределения процесса больше единицы?
2) Плотность распределения процесса больше единицы?
3) Функция распределения процесса отрицательной?
4) Плотность распределения процесса отрицательной?
5) Дисперсия процесса больше единицы?
6) Среднеквадратичное отклонение процесса меньше нуля?
7) Корреляционная функция процесса отрицательной?
8) Спектральная плотность процесса отрицательной?
9) Нормированная корреляционная функция процесса равна нулю?

Решение.

1) нет; 2) да; 3) нет; 4) нет; 5) да; 6) нет; 7) да; 8) нет; 9) да.

Пример 2.8. Как изменятся основные характеристики случайного процесса, если: 1) его значения умножить на постоянную величину a; 2) к процессу добавить постоянную величину a?

Решение. 1) Математическое ожидание умножится на a; дисперсия увеличится в a^2 раз; среднеквадратическое отклонение умножится на $|a|$; корреляционная функция увеличится в a^2 раз; спектральная плотность умножится на a^2; функция плотности в a раз увеличит масштаб по оси абсцисс и в a раз уменьшит масштаб по оси ординат.

2) К математическому ожиданию добавится величина a; график функции плотности сдвигается влево на a единиц, если $a < 0$, или на a единиц вправо, если $a > 0$; другие характеристики не изменятся.

Пример 2.9. Какова размерность: 1) функции распределения случайного процесса; 2) плотности распределения; 3) математического ожидания; 4) дисперсии; 5) среднеквадратического отклонения; 6) корреляционной функции; 7) спектральной плотности; 8) нормированной корреляционной функции; 9) нормированной спектральной плотности; 10) взаимной корреляционной функции?

Решение. 1) безразмерная; 2)обратная размерность случайного процесса; 3) размерность случайного процесса; 4) размерность квадрата случайного процесса; 5) размерность случайного процесса; 6) размерность квадрата случайного процесса; 7) размерность квадрата случайного процесса, деленная на размерность частоты; 8) безразмерная; 9) обратная размерность частоты; 10) размерность одного процесса, умноженная на размерность другого.

Пример 2.10. Корреляционная функция процесса определяется выражением $R_x(r) = \sigma_x^2 e^{-a|t|}$, где $a > 0$. Определить спектральную плотность процесса.

Решение. Воспользуемся следующим соотношением

$$S_x(2\pi f) = \int_{-\infty}^{+\infty} R_x(\tau) e^{-2\pi i f \tau} d\tau = 4\int_0^{\infty} R_x(\tau)\cos(2\pi f\tau) d\tau .$$

Имеем

$$S(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau} R(\tau) d\tau = \frac{\sigma_x^2}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau - a|\tau|} d\tau =$$

$$= \frac{\sigma_x^2}{2\pi}\left\{\int_{-\infty}^{0} e^{(a-i\omega)\tau} d\tau + \int_0^{\infty} e^{-(a+i\omega)r} d\tau\right\} =$$

$$= \frac{\sigma_x^2}{2\pi}\left(\frac{1}{a-i\omega} + \frac{1}{a+i\omega}\right) = \frac{a\sigma_x^2}{\pi(\omega^2 + a^2)}.$$

Пример 2.11. Нормированная АКФ процесса убывает по линейному закону от единицы до нуля. Определить нормированную спектральную плотность процесса.

Решение. Корреляционная функция выражается формулой

$$r_x(\tau)=\begin{cases}0\,, & \text{при} \quad \tau \ge \tau_0;\\ 1-\dfrac{\tau}{\tau_0}\,, & \text{при} \quad 0 \le \tau < \tau_0.\end{cases}$$

Нормированную спектральную плотность получим из соотношения

$$S_x^H(\omega)=\frac{2}{\pi}\int_0^\infty r_x(\tau)\cos\omega\tau d\tau=\frac{2}{\pi}\int_0^{\tau_0}(1-\frac{\tau}{\tau_0})\cos\omega\tau d\tau=\frac{2(1-\cos\omega\tau_0)}{\pi\tau_0\omega^2}\ .$$

Первый (абсолютный) максимум спектральной плотности достигается при ω = 0. Раскрытием неопределенности в этой точке убеждаемся, что он равен τ_0/π. Изменение τ_0 равносильно изменению масштаба кривой $S_x^H(\omega)$ по обеим осям при сохранении ее единичной площади.

Пример 2.12. Спектральная плотность изменения температуры воздуха в летний период (температура фиксировалась ежедневно в 12.00 часов) выражается зависимостью $S_x(\omega)=\dfrac{a\sigma_x^2}{\pi(\omega^2+a^2)}$.

Определить корреляционную функцию этого процесса.

Решение. Пользуясь преобразованием Фурье, имеем

$$R_x(\tau)=\frac{\sigma_x^2}{\pi}\int_{-\infty}^{\infty}\frac{ae^{i\omega\tau}}{\omega^2+a^2}d\omega\,.$$

Для вычисления этого интеграла применим теорию вычетов. При τ > 0 интеграл $\int_{-\infty}^{\infty}\frac{ae^{i\omega\tau}}{\omega^2+a^2}d\omega$ равняется интегралу, взятому по контуру, составленному из вещественной оси и замкнутой полуокружности бесконечного радиуса, расположенной в верхней полуплоскости.

Поэтому его значение равно вычету относительно единственного полюса $\omega = ia$, умноженному на $2\pi a$, т.е.

$$\int_{-\infty}^{\infty}\frac{ae^{i\omega\tau}}{\omega^2+a^2}d\omega=\frac{\pi}{a}e^{-a\tau}\quad(\tau>0).$$

Аналогично при τ < 0, замыкая вещественную ось через ниж-

нюю полуплоскость, получаем $\int\limits_{-\infty}^{\infty} \frac{ae^{i\omega\tau}}{\omega^2 + a^2} d\omega = \frac{\pi}{a} e^{-a\tau}$.

Следовательно, при любом знаке τ получим: $R_x(\tau) = \sigma_x^2 e^{-a|\tau|}$.

Пример 2.13. Найти корреляционную функцию стационарного случайного процесса $X(t)$, если ее спектральная плотность $S_x(\omega)$ постоянна на интервале (ω_1, ω_2) и равна c, а вне этого интервала равна нулю:

$$S_x(\omega) = \begin{cases} c, \ \omega \in (\omega_1, \omega_2), \\ 0, \ \omega \notin (\omega_1, \omega_2). \end{cases}$$

Решение. По определению корреляционной функции

$$R_x(\tau) = \int\limits_0^{\infty} S_x(\omega) \cos \omega\tau d\omega = \int\limits_{\omega_1}^{\omega_2} c \cos \omega\tau d\omega = \frac{c}{\omega} [\sin \omega_2 \tau - \sin \omega_1 \tau] =$$

$$= \frac{2c}{\tau} \cos\left(\frac{\omega_2 + \omega_1}{2} \tau \right) \sin\left(\frac{\omega_2 - \omega_1}{2} \right).$$

Дисперсия рассматриваемого случайного процесса $X(t)$ будет

$$D_x = \int\limits_0^{\infty} S_x(\omega) d\omega = \int\limits_{\omega_1}^{\omega_2} c d\omega = c(\omega_2 - \omega_1).$$

Откуда $c = \frac{D_x}{\omega_2 - \omega_1}$.

Следовательно, $r_x(\tau) = \frac{2D_x}{\omega_2 - \omega_1} \cos\left(\frac{\omega_2 + \omega_1}{2} \tau \right) \sin\left(\frac{\omega_2 - \omega_1}{2} \tau \right) \frac{1}{\tau}$.

Рассмотрим предел этого выражения при $\omega_2 \to \omega_1$:

$$\lim_{\omega_2 \to \omega_1} \left(\frac{D_x \cos\left(\frac{\omega_2 + \omega_1}{2} \tau \right) \sin\left(\frac{\omega_2 - \omega_1}{2} \tau \right)}{\left(\frac{\omega_2 - \omega_1}{2} \tau \right)} \right) =$$

$$= D_x \cos\omega_1\tau \lim_{\omega_2 \to \omega_1} \frac{\sin\left(\frac{\omega_2 - \omega_1}{2}\tau\right)}{\left(\frac{\omega_2 - \omega_1}{2}\tau\right)} = D_x \cos\omega_1\tau.$$

Таким образом, при $\omega_2 \to \omega_1$ мы получили случай, когда $X(t)$ является элементарным стационарным случайным процессом – случайные колебания на частоте ω_1.

Пример 2.14. Найти спектральную плотность процесса $X(t)$, представляющего собой случайную телеграфную волну с корреляционной функцией $R_x(\tau) = a^2 e^{-2\lambda|\tau|}$.

Решение.

$$S_x^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} R_x(\tau)e^{i\omega\tau}d\tau = \frac{a^2}{2\pi}\left[\int_{-\infty}^{0} e^{-2\lambda|\tau|}e^{i\omega\tau}d\tau + \int_{0}^{\infty} e^{-2\lambda|\tau|}e^{i\omega\tau}\right] =$$

$$= \frac{a^2}{2\pi}\left[\int_{-\infty}^{0} e^{2\lambda\tau + i\omega\tau}d\tau + \int_{0}^{\infty} e^{-2\lambda\tau}d\tau\right] = \frac{a^2}{2\pi}\left[\frac{1}{2\lambda + i\omega} - \frac{1}{-2\lambda + i\omega}\right] =$$

$$= \frac{a^2}{2\pi}\left[\frac{1}{2\lambda + i\omega} + \frac{1}{2\lambda - i\omega}\right] = \frac{a^2}{2\pi}\cdot\frac{4\lambda}{(2\lambda)^2 + \omega^2} = \frac{a^2}{2\pi}\cdot\frac{2\lambda}{(2\lambda)^2 + \omega^2}.$$

Пример 2.15. Показать, что не существует никакой стационарной случайной функции $X(t)$, корреляционная функция которой $R_x(\tau)$ постоянна в каком–то интервале $(-\tau, \tau)$ и равна нулю вне его.

Решение. Предположим противное, т.е. что существует случайная функция $X(t)$, для которой корреляционная функция равна значению $b \neq 0$ при $|\tau| < \tau_1$ и равна 0 при $|\tau| > \tau_1$.

Попробуем найти спектральную плотность случайной функции $X(t)$:

$$S_x(\omega) = \frac{1}{\pi}\int_{0}^{\infty} R_x(\tau)\cos\omega\tau d\tau = \frac{1}{\pi}\int_{0}^{\tau_1} b\cos\omega\tau d\tau = \frac{b}{\pi}\cdot\frac{\sin\omega\tau_1}{\omega}.$$

Из этого выражения видно, что функция $S_x(\omega)$ для некоторых

значений ω отрицательна, что противоречит свойствам спектральной плотности, и следовательно, корреляционная функция указанного выше вида существовать не может.

Пример 2.16. Показать, что стационарный «белый шум» $X(t)$ имеет постоянную спектральную плотность.

Решение. У стационарного белого шума корреляционная функция может быть записана в виде $R_x(\tau) = c\delta(\tau)$.

Отсюда

$$S_x^*(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} c\delta(\tau)e^{-i\omega\tau}d\tau = \frac{c}{2\pi}.$$

Величина c называется интенсивностью белого шума.

Таким образом, стационарный белый шум представляет собой случайные колебания на всех частотах.

При этом дисперсия этих колебаний, приходящихся на элементарный участок $\Delta\omega$, остается постоянной и не зависит от частоты колебаний ω.

Эта дисперсия не зависит от частоты ω и будет приближенно равна величине

$$\Delta D_x \approx S_x^*(\omega)\Delta\omega = \frac{c}{2\pi}\Delta\omega \ .$$

Пример 2.17.

Система описывается диф. уравнением:

$$2\frac{dy(t)}{dt} - 3y(t) = 4\frac{dx(t)}{dt} + x(t).$$

Найти частотные характеристики системы.

Решение. Найдем передаточную функцию системы:

$$W(P) = \frac{b_1P + b_0}{a_1P + a_0} = \frac{4P+1}{2P-3}.$$

Амплитудно–фазовая функция системы: $W(i\omega) = \dfrac{4i\omega+1}{2i\omega-3}$.

Выражение для амплитудной частотной характеристики найдем как отношение модулей:

$$H(\omega) = \frac{|b_1i\omega + b_0|}{|a_1i\omega + a_0|} = \frac{|4i\omega+1|}{|2i\omega-3|} = \frac{\sqrt{16\omega^2+1}}{\sqrt{4\omega^2+9}}.$$

а для фазовой частотной характеристики – как разность аргументов числителя и знаменателя:

$$\Phi(\omega) = \arg(b_1 i\omega + b_0) - \arg(a_1 i\omega + a_0) =$$

$$= arctg\,\frac{b_1\omega}{b_0} - arctg\,\frac{a_1\omega}{a_0} = arctg\,4\omega + \arg tg\,\frac{2}{3}\omega\,.$$

Пример 2.18. Найти переходную функцию $h(t)$ элемента, описываемого уравнением $(a_1 P + a_0)\cdot y(t) = b_0 \cdot x(t)$.

Решение. Переходная функция имеет две составляющие:

$$h(t) = h_{\text{в}}(t) + h_{\text{с}}(t).$$

Вынужденная составляющая в данном случае равна:

$$h_{\text{в}}(t) = \frac{b_0}{a_0} = \text{const}\,.$$

Свободную составляющую будем искать в виде:

$$h_{\text{с}}(t) = c \cdot e^{-\frac{a_0}{a_1}\cdot t}\,.$$

Учитывая начальное условие $y(0) = 0$, получим: $c = -\dfrac{b_0}{a_0}$.

Тогда $h(t) = \dfrac{b_0}{a_0}\cdot\left(1 - e^{-\frac{a_0}{a_1}\cdot t}\right)$.

Пример 2.19.

Определить реакцию элемента, описываемого уравнением $(a_1 P + a_0)\cdot y(t) = b_0 x(t)$, на воздействие $x(t) = at1(t)$.

Решение.

Импульсная переходная функция элемента:

$$g(t) = \frac{dh(t)}{dt} = \frac{b_0}{a_1}\cdot e^{-\frac{a_1}{a_0}\cdot t}\,.$$

Функцию $y(t)$, описывающую изменение выходной величины после подачи линейного воздействия, получим, подставляя последнее выражение в интеграл Дюамеля:

$$y(t) = \int_0^t x(\tau)\cdot g(t-\tau)d\tau = \int_0^t a\tau\frac{b_0}{a_1}\cdot e^{-\frac{a_1}{a_0}\cdot(t-\tau)}d\tau = a\frac{b_0}{a_0}\left[t - \frac{a_1}{a_0}\cdot\left(1 - e^{-\frac{a_1}{a_0}\cdot t}\right)\right]1(t).$$

Задачи, предлагаемые для решения на практических занятиях по третьему разделу курса «Математическая статистика»

Пример 3.1. Исследовалась взаимосвязь между тремя показателями: производительностью труда (X_1), возрастом (X_2), и производственным стажем (X_3). По выборке из 100 рабочих одной и той же специальности вычислены парные коэффициенты корреляции: $r_{12} = 0{,}2$, $r_{13} = 0{,}41$, $r_{23} = 0{,}82$. Вычислить множественные и частные (парциальные) коэффициенты корреляции. Оценить значимость этих коэффициентов.

Решение. Вычислим множественные коэффициенты корреляции, характеризующие связь между одним зависимым и двумя независимыми показателями.

$$R_{y(x,v)} = \sqrt{\frac{r_{yx}^2 + r_{yv}^2 - 2r_{yx}r_{yv}r_{xv}}{1 - r_{xv}^2}};$$

$$R_{1(2,3)} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}} = \sqrt{\frac{0,2^2 + 0,41^2 - 2\cdot 0,2\cdot 0,41\cdot 0,82}{1 - 0,82^2}} = 0,47;$$

$$R_{2(1,3)} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21}r_{23}r_{13}}{1 - r_{13}^2}} = \sqrt{\frac{0,2^2 + 0,82^2 - 2\cdot 0,2\cdot 0,82\cdot 0,41}{1 - 0,41^2}} = 0,83;$$

$$R_{3(1,2)} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31}r_{32}r_{12}}{1 - r_{12}^2}} = \sqrt{\frac{0,41^2 + 0,82^2 - 2\cdot 0,41\cdot 0,82\cdot 0,2}{1 - 0,2^2}} = 0,86.$$

Оценим значимость вычисленных множественных коэффициентов корреляции при уровне значимости $\alpha = 0{,}05$.

Проверка гипотезы о значимости множественного коэффициента корреляции (H_0: $R_{y(x,v)} = 0$) осуществляется с использованием F – критерия. Вычислим наблюдаемое значение критерия:

$$F_H = \frac{R_{y(x,v)}^2 (n - m)}{(1 - R_{y(x,v)}^2)(m - 1)},$$

где $n = 100$ – объем выборки, $m = 3$ – количество показателей.

$$F_H = \frac{R_{1(2,3)}^2 (n - m)}{(1 - R_{1(2,3)}^2)(m - 1)} = \frac{0,47^2 (100 - 3)}{(1 - 0,47^2)(3 - 1)} = 13,75;$$

$$F_H = \frac{R^2_{2(1,3)}\left(n-m\right)}{(1-R^2_{2(1,3)})(m-1)} = \frac{0,83^2\left(100-3\right)}{(1-0,83^2)(3-1)} = 107,4;$$

$$F_H = \frac{R^2_{3(1,2)}\left(n-m\right)}{(1-R^2_{3(1,2)})(m-1)} = \frac{0,86^2\left(100-3\right)}{(1-0,86^2)(3-1)} = 137,75.$$

По таблице квантилей F – распределения найдем критическую точку.

$$F_{кр} = F\left(1-\alpha, m-1, n-m\right) = F\left(1-0,05; 3-1; 100-3\right) = F\left(0,95;\ 2; 97\right) = 3,1.$$

Так как $F_H > F_{кр}$, то гипотезу H_0 отвергаем, т.е. множественные коэффициенты корреляции значимы.

Вычислим частные (парциальные) коэффициенты корреляции, характеризующие связь между двумя показателями при исключенном влиянии третьего.

$$r_{xy|v} = \frac{r_{xy} - r_{yv}r_{xv}}{\sqrt{(1-r^2_{yv})(1-r^2_{xv})}};$$

$$r_{12|3} = \frac{r_{12} - r_{23}r_{13}}{\sqrt{(1-r^2_{23})(1-r^2_{13})}} = \frac{0,2-0,82\cdot 0,41}{\sqrt{(1-0,82^2)(1-0,41^2)}} = -0,26;$$

$$r_{23|1} = \frac{r_{23} - r_{31}r_{21}}{\sqrt{(1-r^2_{31})(1-r^2_{21})}} = \frac{0,82-0,41\cdot 0,2}{\sqrt{(1-0,41^2)(1-0,2^2)}} = 0,83;$$

$$r_{13|2} = \frac{r_{13} - r_{32}r_{12}}{\sqrt{(1-r^2_{32})(1-r^2_{12})}} = \frac{0,41-0,82\cdot 0,2}{\sqrt{(1-0,82^2)(1-0,2^2)}} = 0,44.$$

Оценим значимость вычисленных частных (парциальных) коэффициентов корреляции при уровне значимости $\alpha = 0,05$.

Проверка гипотезы о значимости частного коэффициента корреляции ($H_0 : r_{xy|v} = 0$) осуществляется с использованием T − распределения Стьюдента.

Вычислим наблюдаемое значение критерия:

$$T_H = r_{xy|v}\sqrt{\frac{n-2}{1-r^2_{xy|v}}};$$

$$T_H = r_{12|3}\sqrt{\frac{n-2}{1-r^2_{12|3}}} = -0,26\sqrt{\frac{100-2}{1-\left(-0,26\right)^2}} = -2,67;$$

$$T_H = r_{23|1}\sqrt{\frac{n-2}{1-r^2_{23|1}}} = 0,83\sqrt{\frac{100-2}{1-0,83^2}} = 14,7;$$

$$T_H = r_{13|2}\sqrt{\frac{n-2}{1-r_{13|2}^2}} = 0{,}44\sqrt{\frac{100-2}{1-0{,}44^2}} = 4{,}85.$$

По таблице квантилей T − распределения Стьюдента найдем критическую точку.

$$T_{кр} = \mp T\left(1-\frac{\alpha}{2}; n-2\right) = T\left(1-\frac{0{,}05}{2}; 100-2\right) = T(0{,}975; 98) = 1{,}985.$$

Так как $|T_H| > T_{кр}$, то гипотезу H_0 отвергаем, т.е. частные (парциальные) коэффициенты корреляции значимы.

Пример 3.2. На конкурсе инвестиционных проектов 11 участников получили следующие оценки (по стобальной системе) за экологичность (экологическую безопасность) и экономическую обоснованность расчетов:

№ проекта	1	2	3	4	5	6	7	8	9	10	11
Экологичность	70	66	89	60	75	78	53	52	55	59	90
Экономическая обосноснованность	73	53	85	55	78	89	63	64	65	51	75

Связаны ли между собой экологичность и экономическая обоснованность расчетов?

Решение. Тесноту связи между экологичностью и экономической обоснованностью расчетов определим с помощью ранговых коэффициентов корреляции. Ранжируем участников конкурса:

№ участника	11	3	6	5	1	2	4	10	9	7	8
Ранг за экологичность, x_i	1	2	3	4	5	6	7	8	9	10	11
Ранг за экон. обосн., y_i	4	2	1	3	5	10	9	11	6	8	7

Вычислим ранговый коэффициент Спирмена:

$$r_c = 1-\frac{6}{n^3-n}\sum_{i=1}^{n}(x_i-y_i)^2 = 1-\frac{6}{11^3-11}\Big((1-4)^2+(2-2)^2+(3-1)^2+(4-3)^2+$$

$$+(5-5)^2+(6-10)^2+(7-9)^2+(8-11)^2+(9-6)^2+(10-8)^2+(11-7)^2\Big) = 0{,}673.$$

Справа от y_1 имеется $k_1 = 7$ рангов больших, чем y_1, справа от y_2 имеется $k_2 = 8$ рангов больших, чем y_2. Аналогично находим: $k_3 = 8,\ k_4 = 7,\ k_5 = 6,\ k_6 = 1,\ k_7 = 1,\ k_8 = 0,\ k_9 = 2,\ k_{10} = 0.$

Вычислим ранговый коэффициент Кендалла:

$$r_k = \frac{4}{n(n-1)}\sum_{i=1}^{n-1} k_i - 1 = \frac{4}{11\cdot 10}(7+8+8+7+6+1+1+0+2+0) - 1 = 0{,}46.$$

Оценим значимость вычисленных коэффициентов Спирмена и Кендалла при уровнях значимости $\alpha = 0{,}01$, $\alpha = 0{,}05$, $\alpha = 0{,}1$.

Проверка гипотезы о значимости коэффициента ранговой корреляции Спирмена ($H_0 : r_c = 0$) осуществляется с использованием T – критерия Стьюдента.

Вычислим наблюдаемое значение критерия:

$$T_H = \tilde{r}_c \sqrt{\frac{n-2}{1-\tilde{r}_c^{\,2}}} = 0{,}673\sqrt{\frac{11-2}{1-0{,}673^2}} = 2{,}73.$$

Для различных уровней значимости по таблице квантилей T – критерия Стьюдента найдем критические точки.

$$T_{кр} = \mp T\left(1-\frac{\alpha}{2}; n-2\right);$$

При $\alpha = 0{,}01$ $T_{кр} = T\left(1-\frac{0{,}01}{2}; 9\right) = T(0{,}995;9) = 3{,}25$;

При $\alpha = 0{,}05$ $T_{кр} = T\left(1-\frac{0{,}05}{2}; 9\right) = T(0{,}975;9) = 2{,}262$;

При $\alpha = 0{,}10$ $T_{кр} = T\left(1-\frac{0{,}1}{2}; 9\right) = T(0{,}95;9) = 1{,}833$.

При уровнях значимости $\alpha = 0{,}05$ и $\alpha = 0{,}1$ данные противоречат гипотезе о незначимости коэффициента Спирмена, т.е. между техникой и артистизмом есть связь, а при уровне значимости $\alpha = 0{,}01$ нет оснований отвергать гипотезу и полученный коэффициент незначим.

Проверим значимость коэффициента ранговой корреляции Кендалла ($H_0 : r_k = 0$). Вычислим наблюдаемое значение критерия:

$$Z_H = \tilde{r}_k \sqrt{\frac{9n(n+1)}{2(2n+5)}} = 0{,}46\sqrt{\frac{9\cdot 11\cdot 12}{2\cdot 27}} = 2{,}16.$$

Для различных уровней значимости с помощью табличной функции Лапласа найдем критические точки.

$$Z_{кр} = \mp \arg\Phi\left(\frac{1-\alpha}{2}\right);$$

При $\alpha = 0{,}01$ $Z_{кр} = \arg\Phi\left(\frac{1-0{,}01}{2}\right) = \arg\Phi(0{,}495) = 2{,}6$;

При $\alpha = 0{,}05$ $Z_{кр} = \arg\Phi\left(\frac{1-0{,}05}{2}\right) = \arg\Phi(0{,}475) = 1{,}96$;

При $\alpha = 0{,}10$ $Z_{кр} = \arg\Phi\left(\frac{1-0{,}1}{2}\right) = \arg\Phi(0{,}45) = 1{,}65$.

При уровнях значимости $\alpha = 0{,}05$ и $\alpha = 0{,}1$ данные противоречат гипотезе о не значимости коэффициента Кендалла, т.е. между экологичностью и экономической обоснованностью расчетов есть связь (это характеризует качество всего проекта), а при уровне значимости $\alpha = 0{,}01$ нет оснований отвергать гипотезу и полученный коэффициент не значим.

Пример 3.3. Группа из 5 экспертов оценивает качество однотипной продукции, выпускаемой на 7 предприятиях. Предпочтения экспертов (их ранги) представлены в таблице:

Эксперт	Предприятия						
	1	2	3	4	5	6	7
1	1	3	4	2	6	7	5
2	1	2	5	3	6	4	7
3	2	1	7	5	6	4	3
4	1	2	4	6	3	5	7
5	3	1	5	4	2	6	7

Взаимосвязаны (согласуются) ли мнения экспертов? Рассчитать коэффициент конкордации и оценить его значимость.

Решение. Вычислим коэффициент конкордации по формуле:

$$R_k = \frac{12}{m^2\left(n^3 - n\right)} \sum_{i=1}^{n} \left(\sum_{j=1}^{m} k_{ij} - \frac{m(n+1)}{2} \right)^2,$$

где k_{ij} – ранги i–го наблюдения для j–го показателя,

$n = 7$ – объем выборки,

$m = 5$ – количество ранговых показателей (экспертов).

$$\frac{m(n+1)}{2} = \frac{5(7+1)}{2} = 20.$$

Промежуточные расчеты представлены в таблице:

Расчетные величины	Предприятия							Итого
	1	2	3	4	5	6	7	
$\sum_{j=1}^{m} k_{ij}$	8	9	25	20	23	26	29	140

$\sum_{j=1}^{m} k_{ij} - \frac{m(n+1)}{2}$	–12	–11	5	0	3	6	9	0
$\left(\sum_{j=1}^{m} k_{ij} - \frac{m(n+1)}{2}\right)^2$	144	121	25	0	9	36	81	416

$$R_k = \frac{12}{5^2\left(7^3 - 7\right)} 416 = 0{,}594.$$

Значимость коэффициента конкордации $(H_0 : R_k = 0)$ проверяется по χ^2 – распределению.

Вычислим наблюдаемое значение критерия:

$$\chi_H^2 = m(n-1)R_k = 5 \cdot (7-1) \cdot 0{,}594 = 17{,}82.$$

Критическое значение $\chi_{кр}^2$ найдем по таблице квантилей χ^2– распределения по заданному уровню значимости α и числу степеней свободы $\nu = n - 1$.

$$\chi_{кр}^2 = \chi^2(1-\alpha; n-1) = \chi^2(1-0{,}05; 7-1) = \chi^2(0{,}95; 6) = 12{,}59.$$

Так как $\chi_H^2 > \chi_{кр}^2$, то гипотезу H_0 отвергаем, т.е. коэффициент конкордации R_k значим. Значит, мнение экспертов по оценки качества выпускаемой продукции на предприятиях согласуются.

Пример 3.4. Исследовалось влияние индивидуальных качеств (особенностей, заслуг) продавцов на объемы выручки магазина. Получены следующие данные:

Номера продавцов	Число отработанных дней	Дневная выручка магазина (тыс. долл.)
1	8	16.5; 15.0; 16.2; 18.9; 20.1; 19.3; 10.1; 12.8
2	13	16.7; 16.3; 14.0; 15.0; 16.7; 12.4; 7.9; 9.8; 14.4; 10.8; 11.1; 13.0; 10.7
3	9	10.7; 9.0; 13.9; 11.3; 9.4; 11.9; 10.5; 9.7; 7.4

Определить, влияют ли индивидуальные качества продавцов на объемы выручки магазина?

Решение. По имеющимся данным y_{ij} определим средние значения дневной выручки (переменной *Y*) на каждом (из трех) *i*–м уровне:

$$\bar{y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} .$$

Получим

$$\bar{y}_1 = 16{,}11; \qquad \bar{y}_2 = 12{,}98; \qquad \bar{y}_3 = 10{,}44.$$

Среднее значение переменной Y по всем значениям

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} y_{ij} = 13{,}05.$$

Результаты других расчетов сведены в табл.3.1.

Таблица 3.1

Результаты расчета примера однофакторного дисперсионного анализа

Источник изменчивости	Сумма квадратов отклонений	Число степеней свободы	Дисперсия	F_H
Факторная (систематическая)	131,67	2	65,83	8,46
Случайная составляющая	210,05	27	7,78	
Общая изменчивость	341,72	29		

Критическое значение находим по таблицам квантилей F – распределения при уровне значимости $\alpha = 0{,}05$, что соответствует квантилю порядка 0,95

$$F_{кр} = F(0{,}95;\ 2;\ 27) = 3{,}35.$$

Так как $F_H > F_{кр}$, то гипотеза H_0 о равенстве дневной выручки у каждого продавца отвергается, и принимается конкурирующая гипотеза H_1 – средние выручки продавцов (на различных уровнях) различны.

Это означает, что индивидуальные качества (особенности) продавцов влияют на объемы выручки магазина.

Пример 3.5. Катализатор для химической реакции получался четырьмя различными способами. В экспериментах проверялась активность катализатора, причем для каждого уровня (способа получения) было сделано по пять наблюдений.

Получены следующие данные:

Способы получения катализатора	Активность катализатора
A1	56; 55; 62; 59; 60
A2	64; 61; 50; 55; 56
A3	45; 46; 45; 39; 43
A4	42; 39; 45; 43; 41

Необходимо проверить независимость качества (активности) катализатора от способа получения

Решение. По имеющимся данным y_{ij} определим средние значения активности катализатора при каждом (из четырех) i–м способе его получения. Получим $\bar{y}_1 = 58,4;\quad \bar{y}_2 = 57,2;\quad \bar{y}_3 = 43,6;\quad \bar{y}_4 = 42,0.$

Среднее значение переменной Y по всем наблюдениям $\bar{Y} = 50.3$.

Результаты других расчетов сведены в табл. 3.2.

Таблица 3.2

Результаты расчета примера однофакторного дисперсионного анализа

Источник изменчивости	Сумма квадратов отклонений	Число степеней свободы	Дисперсия	F_H
Факторная (систематическая)	1135	3	378,3	29,8
Случайная составляющая	203,2	16	12,7	
Общая изменчивость	1338,2	19		

Критическое значение критерия находим по таблицам квантилей F – распределения при уровне значимости $\alpha = 0.05$, что соответствует квантилю порядка 0,95: $F_{кр} = F(0,95;\ 2;\ 27) = 3,35.$

Так как $F_H > F_{кр}$, то гипотеза H_0 – о равенстве активности катализатора при различных способах его получения (на каждом уровне) отвергается, и принимается конкурирующая гипотеза H_1 – активность катализатора при различных способах его получения различна.

Пример 3.6. Четыре различных типа покрышек испытывались на автомобилях четырех различных марок. Измерялся износ шин в мм после пробега 40 тысяч км. Получены следующие результаты:

Тип покрышек	Марка автомобиля (фактор B на 4 уровнях)			
	ВОЛГА	ЖИГУЛИ	ШКОДА	МОСКВИЧ
1	14	11	9	10
2	11	11	9	8
3	10	10	7	8
4	10	5	6	6

Влияет ли тип покрышек и марка автомобиля на износ шин?

Решение. Имеем задачу двухфакторного дисперсионного анализа с однократным наблюдением на каждой комбинации уровней ($k = 4$, $m = 4$), следовательно можно воспользоваться схемой расчета из табл. 4.2.

По имеющимся исходным данным получаем:

$$\bar{y}_{1\cdot} = 11; \qquad \bar{y}_{2\cdot} = 9,75; \qquad \bar{y}_{3\cdot} = 8.75; \qquad \bar{y}_{4\cdot} = 6,75;$$

$$\bar{y}_{\cdot 1} = 11,25; \quad \bar{y}_{\cdot 2} = 9,25; \qquad \bar{y}_{\cdot 3} = 7.75; \qquad \bar{y}_{\cdot 4} = 8,00;$$

$$\bar{Y} = 9,06 \ .$$

Другие результаты вычислений представлены в табл. 3.3

Таблица 3.3

Результаты расчета примера двухфакторного дисперсионного анализа

Источник изменчивости	Сумма квадратов отклонений	Число степеней свободы	Дисперсия	F_H
Фактор *A*	30,6	3	10,2	7,8
Фактор *B*	38,6	3	12,9	9,9
Случайная составляющая	11,7	9	1,3	
Общая изменчивость	80,9	15		

Критические значения критерия находим по таблицам квантилей *F* – распределения:

$$F_{кр} = F(0,99; \ \ 3; \ \ 9) \ = \ 7,0; \qquad F_{кр} = F(0,95; 3; 9) \ = \ 3,86.$$

Так как $F_H > F_{кр}$ для фактора *A* и для фактора *B*, то гипотеза H_0 – о равенстве износов на каждом уровне отвергается, и принимается конкурирующая гипотеза H_1 – средние износы на различных уровнях различны.

Следовательно различные типы покрышек и марки автомобилей влияют на износ покрышек (шин).

Поэтому не безразлично, какие покрышки устанавливаются на автомобили той или иной марки.

Пример 3.7. За прошедший год собраны данные о затратах на рекламу (*X*, долл./нед.) и объеме реализации продукции (*Y*, долл./нед.).

Получили следующие данные:

x_i \ y_j	1200 – 1300	1300 – 1400	1400 – 1500	1500 – 1600	1600 – 1700
40 – 45	1				
45 – 50	3	2	3		
50 – 55	2	4	4	3	
55 – 60		4	5	6	2
60 – 65			3	4	2
65 – 70					2

Есть ли взаимосвязь между вложениями в рекламу и объемом продаж? Если взаимосвязь есть, то найти уравнение связи.

Спрогнозировать объем продаж, если вложения в рекламу составят 80 долл./нед.

Сколько средств надо вкладывать в рекламу, чтобы получить объем реализации продукции в 1000 долл.?

Решение. Пусть k и m – количество интервалов группировки исходных показателей x_i и y_j.

Имеем: $k = 6$, $m = 5$.

Тогда общее количество наблюдений

$$n = \sum_{i=1}^{k}\sum_{j=1}^{m} n_{ij} = 1+3+2+3+2+4+4+3+4+5+6+2+3+4+2+2 = 50.$$

Вычислим основные характеристики, взяв за x_i и y_j соответствующие середины интервалов:

$$m_x = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m} x_i n_{ij}\ ; \quad m_y = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m} y_j n_{ij}\ ;$$

$$m_x = \frac{1}{50}\Big(42{,}5\cdot 1 + 47{,}5\cdot(3+2+3) + 52{,}5\cdot(2+4+4+3) + 57{,}5\cdot(4+5+6+2) +$$

$$+62{,}5\cdot(3+4+2) + 67{,}5\cdot 2\Big) = \frac{1}{50}\cdot 2780 = 55{,}6 \quad \text{долл./нед.};$$

$$m_y = \frac{1}{50}\Big(1250\cdot(1+3+2) + 1350\cdot(2+4+4) + 1450\cdot(3+4+5+3) +$$

$$+1550\cdot(3+6+4) + 1650\cdot(2+2+2)\Big) = \frac{1}{50}\cdot 72800 = 1456 \quad \text{долл./нед.};$$

$$\sigma_x^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m}(x_i - m_x)^2 n_{ij}\ ; \quad \sigma_y^2 = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m}(y_j - m_y)^2 n_{ij}\ ;$$

$$\sigma_x^2 = \frac{1}{50}\Big((42,5-55,6)^2 \cdot 1 + (47,5-55,6)^2(3+2+3) +$$

$$+(52,5-55,6)^2(2+4+4+3) + (57,5-55,6)^2(4+5+6+2) +$$

$$+(62,5-55,6)^2(3+4+2) + (67,5-55,6)^2 \cdot 2\Big) = \frac{1}{50} \cdot 1594,5 =$$

$$= 31,89 \quad (\text{долл./нед.})^2;$$

$$\sigma_y^2 = \frac{1}{50}\Big((1250-1456)^2(1+3+2) + (1350-1456)^2(2+4+4) +$$

$$+(1450-1456)^2(3+4+5+3) + (1550-1456)^2(3+6+4) +$$

$$+(1650-1456)^2(2+2+2)\Big) = \frac{1}{50} \cdot 708200 = 14164 \quad (\text{долл./нед.})^2;$$

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{31,89} = 5,65 \quad \text{долл./нед.};$$

$$\sigma_y = \sqrt{\sigma_y^2} = \sqrt{14164} = 119 \quad \text{долл./нед.};$$

$$R_{xy} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{m}(x_i - m_x)(y_j - m_y)n_{ij} = \frac{1}{50}\Big((42,5-55,6)(1250-1456) \cdot 1 +$$

$$+(47,5-55,6)\big((1250-1456) \cdot 3 + (1350-1456) \cdot 2 + (1450-1456) \cdot 3\big) +$$

$$+(52,5-55,6)\big((1250-1456) \cdot 2 + (1350-1456) \cdot 4 + (1450-1456) \cdot 4 +$$

$$+(1550-1456) \cdot 3\big) + (57,5-55,6)\big((1350-1456) \cdot 4 + (1450-1456) \cdot 5 +$$

$$+(1550-1456) \cdot 6 + (1650-1456) \cdot 2\big) + (62,5-55,6)\big((1450-1456) \cdot 5 +$$

$$+(1550-1456) \cdot 6 + (1650-1456) \cdot 2\big) + (67,5-55,6)(1650-1456) \cdot 2\Big) =$$

$$= \frac{1}{50} \cdot 23284,4 = 441,4 \quad (\text{долл./нед.})^2;$$

$$r_{xy} = \frac{R_{xy}}{\sigma_x \sigma_y} = \frac{441,4}{5,65 \cdot 119} = 0,6565.$$

Определим, есть ли взаимосвязь между вложениями в рекламу и объемом продаж с помощью линейного коэффициента корреляции.

Проверка гипотезы о значимости коэффициента корреляции ($H_0 : r_{xy} = 0$) осуществляется с использованием T − распределения Стьюдента.

Вычислим наблюдаемое значение критерия:

$$T_H = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}} = 0{,}6565\sqrt{\frac{50-2}{1-(0{,}6565)^2}} = 6{,}03.$$

По таблице квантилей T − распределения Стьюдента найдем критическую точку для двухсторонней критической области.

$$T_{кр} = \mp T\left(1-\frac{\alpha}{2};\ n-2\right) = T\left(1-\frac{0{,}05}{2};\ 50-2\right) = T\left(0{,}975;\ 48\right) = 2{,}03.$$

Так как $T_H > T_{кр}$, то гипотезу H_0 отвергаем, т.е. коэффициент корреляции при уровне значимости $\alpha = 0{,}05$ значим.

Получим уравнение регрессии (рис.5.2):

$$\bar{y}_x = m_y + r_{xy}\frac{\sigma_x}{\sigma_y}\left(x - m_x\right) = 1456 + 0{,}6565\frac{119}{5{,}65}\left(x - 55{,}6\right) = 13{,}83x + 687{,}2.$$

$$\bar{x}_y = m_x + r_{xy}\frac{\sigma_x}{\sigma_y}\left(y - m_y\right) = 55{,}6 + 0{,}6565\frac{5{,}65}{119}\left(y - 1456\right) = 0{,}0312y + 10{,}22.$$

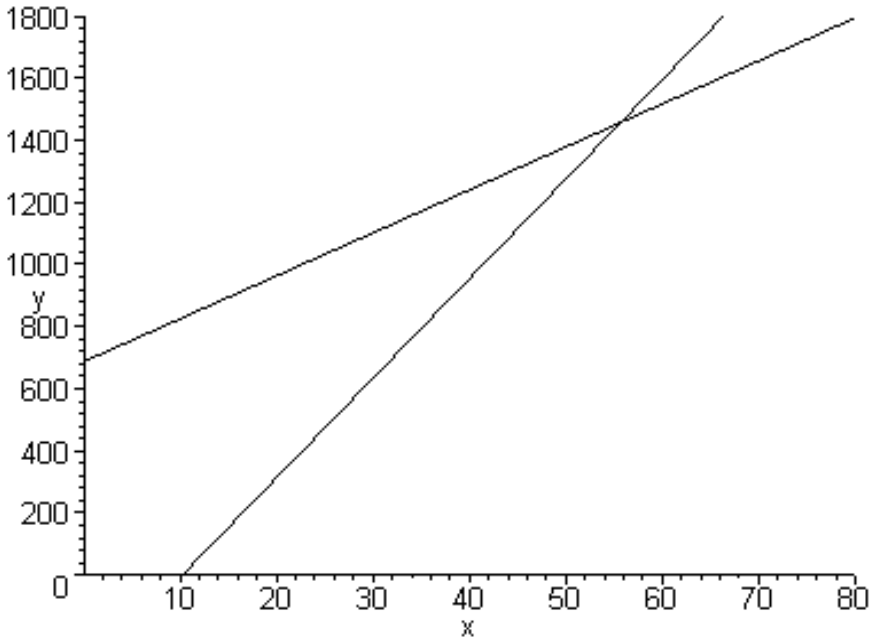


Рис. 3.1. Графическое представление уравнений регрессии

Для проверки гипотезы о значимости полученных моделей регрессии (H_0 : уравнение не значимо) вычислим значение F –критерия:

$$F_H = \frac{R^2(n-m-1)}{(1-R^2)m} = \frac{(0,6565)^2(50-1-1)}{\left(1-(0,6565)^2\right)\cdot 1} = 36,36,$$

где $n = 50$ – количество наблюдений, $m = 1$ – количество показателей.

По таблице квантилей F – распределения найдем критическую точку.

$$F_{кр} = F(1-\alpha;\ m;\ n-m-1) = F(0,95;\ 1;\ 48) = 4,08.$$

Так как $F_H > F_{кр}$, то гипотезу H_0 отвергаем, т.е. уравнения регрессии значимы при уровне значимости $\alpha = 0,05$.

Прогнозную величину еженедельных продаж, если расходы на рекламу составят 80 долл./нед., определим по уравнению регрессии:

$$\bar{y}_{x=80} = \bar{y}(80) = 13,83\cdot 80 + 687,2 = 1793,6 \quad \text{долл./нед.}$$

По уравнению регрессии для $\bar{x}_y$ вычислим сколько надо вкладывать в рекламу, чтобы объем продаж составил 1000 долл./нед.

$$\bar{x}_{y=1000} = \bar{x}(1000) = 0,0312\cdot 1000 + 10,22 = 41,42 \quad \text{долл./нед.}$$

Пример 3.8. Месячные объемы продаж ткани (X, тыс.м.) и величины премиального фонда (Y, тыс. руб.) в девяти филиалах торговой фирмы характеризовались следующими данными:

i	1	2	3	4	5	6	7	8	9
x_i	2,3	2,8	1,9	3,4	2,6	3,3	4,2	3,0	1,7
y_i	6,1	8,2	7,1	14,9	9,1	9,0	15,8	8,2	7,7

Исследовать зависимость премиального фонда от объема продаж. Получить прогноз объема премиального фонда при продажах 3,4 тыс. м ткани. Оценить качество прогноза. Продавцы полагают, что если объем продаж составит 5 тыс. метров, то их премия будет свыше 18 тыс. руб. в месяц. Какова вероятность этого?

Решение.

1. Найдем основные характеристики исследуемых величин:

$$m_x = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{9}(2,3+2,8+1,9+3,4+2,6+3,3+4,2+3,0+1,7) = 2,8;$$

$$m_y = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{9}(6,1+8,2+7,1+14,9+9,1+9+15,8+8,2+7,7) = 9,57;$$

$$\sum_{i=1}^{n}(x_i - m_x)^2 = ((2,3-2,8)^2 + ... + (1,7-2,8)^2) = 4,928;$$

$$\sigma_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)^2} = \sqrt{\frac{1}{9}(4,928)} = 0,74;$$

$$\sigma_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - m_y)^2} = \sqrt{\frac{1}{9}((6,1-9,57)^2 + ... + (7,7-9,57)^2)} = 3,21.$$

2. Исследуем тесноту линейной зависимости показателей. Вычислим корреляционный момент показателей X и Y.

$$R_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - m_x)(y_i - m_y) = \frac{1}{9}((2,3-2,8)\cdot(6,1-9,57) + (2,8-2,8)\times$$
$$\times(8,2-9,57) + (1,9-2,8)\cdot(7,1-9,57) + (3,4-2,8)\cdot(15,9-9,57) + (2,6-$$
$$-2,8)\cdot(9,1-9,57) + (3,3-2,8)\cdot(9,0-9,57) + (4,2-2,8)\cdot(15,8-9,57) +$$
$$+(3,0-2,8)\cdot(8,2-9,57) + (1,7-2,8)\cdot(7,7-9,57)) = 1,94;$$

Линейный коэффициент корреляции:

$$r_{xy} = \frac{R_{xy}}{\sigma_x \sigma_y} = \frac{1,94}{0,74\cdot 3,21} = 0,82.$$

Исследуем найденный коэффициент на значимость.

Воспользуемся T – критерием Стьюдента.

Гипотеза $H_0 : r_{xy} = 0$, конкурирующая гипотеза $H_1 : r_{xy} \neq 0$.

Возьмем уровень значимости $\alpha = 0,01$.

$$T_H = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}} = 0,82\sqrt{\frac{9-2}{1-0,67}} = 3,78;$$

$$T_{кр} = T(1-\alpha/2;\ n-2) = T(0,995;\ 7) = 3,499.$$

Так как $T_H > T_{кр}$, то принимаем гипотезу H_1, т.е. линейный коэффициент корреляции значим и между X и Y существует линейная зависимость.

3. Найдем линейную зависимость в виде: $\overline{y}_x = a + bx,$

где a и b находятся по формулам:

$$b = r_{xy}\frac{\sigma_y}{\sigma_x} = 0,82 \cdot \frac{3,21}{0,74} = 3,56;$$

$$a = m_y - bm_x = 9,57 - 3,56 \cdot 2,8 = -0,398.$$

Уравнение регрессии имеет вид: $\bar{y}_x = -0,398 + 3,56x.$

Вычислим значения $\bar{y}_{x_i}$ по полученному уравнению регрессии:

x_i	2,300	2,800	1,900	3,400	2,600	3,300	4,200	3,000	1,700
$\bar{y}_{x_i}$	7,790	9,570	6,366	11,706	8,858	11,350	14,554	10,282	5,654

Проверим полученное уравнение регрессии на значимость при уровне значимости $\alpha = 0,05$. Для этого воспользуемся F −критерием. Гипотеза H_0: уравнение не значимо, конкурирующая гипотеза H_1: уравнение значимо.

Вычислим значение критерия по формуле: $F_H = \dfrac{r_{xy}^2(n-m-1)}{(1-r_{xy}^2)m}$,

где n − количество наблюдений, m − количество независимых переменных $(m=1)$.

$$F_H = \frac{0,67 \cdot (9-1-1)}{1-0,67} = 14,21.$$

$$F_{кр} = F(1-\alpha;\ m;\ n-m-1) = F(1-0,05;\ 1;\ 7) = 5,59.$$

Так как $T_H > T_{кр}$, то принимаем гипотезу H_1, т.е. уравнение регрессии значимо.

Проверим на значимость коэффициенты уравнения регрессии.

Для этого вычислим стандартную ошибку по соотношению

$$\sigma_\varepsilon = \sqrt{\frac{S_{ост}^2}{n-m-1}} = \sqrt{\sigma_{ост}^2}.$$

Вычислим соответствующую сумму квадратов отклонений.

$$S_{ост}^2 = \sum_{i=1}^{n}(y_i - \bar{y}_{x_i})^2 = (7,1-7,79)^2 + (8,2-9,57)^2 + (7,1-6,366)^2 + (14,9 -$$

$$-11,706)^2 + (9,1-8,858)^2 + (9,0-11,35)^2 + (15,8-14,554)^2 + (8,2-10,282)^2 +$$

$$+(7,7-5,654)^2 = 31,13.$$

Тогда:

$$\sigma_{ост}^2 = \sigma_\varepsilon^2 = \frac{S_{ост}^2}{n-m-1} = \frac{31{,}13}{9-1-1} = 4{,}45; \qquad \sigma_\varepsilon = \sqrt{4{,}45} = 2{,}1.$$

Проверим на значимость коэффициент b. Гипотеза H_0: коэффициент b не значим ($b = 0$), конкурирующая гипотеза H_1: коэффициент b значим ($b \neq 0$). Возьмем уровень значимости $\alpha = 0{,}05$.

Вычислим наблюдаемое значение критерия по формуле:

$$T_{H(b)} = \frac{b}{\sigma_b};$$

где $\sigma_b = \sqrt{\dfrac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(x_i - m_x)^2}} = \sqrt{\dfrac{4{,}45}{4{,}928}} = 0{,}95;$

Тогда $T_{H(b)} = \dfrac{3{,}56}{0{,}95} = 3{,}75;$

$$T_{кр} = T(1-\frac{\alpha}{2};\ n-m-1) = T(1-\frac{0{,}05}{2};\ 7) = T(0{,}975;\ 7) = 2{,}365.$$

Так как $T_H > T_{кр}$, то принимаем гипотезу H_1: коэффициент b значим.

Для коэффициента b можно получить доверительный интервал:

$$I_b = (b \pm T_{кр}\sigma_b) = 3{,}56 \pm 2{,}365 \cdot 0{,}95 = 3{,}56 \pm 2{,}25 = (1{,}31;\ 5{,}81) \text{ для } p_д = 0{,}95.$$

Проверим на значимость коэффициент a.

Гипотеза H_0: ($a = 0$), конкурирующая гипотеза H_1: ($a \neq 0$).

Вычислим значение критерия по формуле: $T_{H(a)} = \dfrac{a}{\sigma_a}$,

где $\sigma_a = \sqrt{\dfrac{\sigma_\varepsilon^2 \sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}} = \sqrt{\dfrac{4{,}45 \cdot 75{,}48}{9 \cdot 4{,}928}} = 2{,}75.$

Тогда $T_{H(a)} = \dfrac{-0{,}398}{2{,}75} = -0{,}145;$

$$T_{кр} = T(1-\frac{\alpha}{2};\ n-m-1) = T(1-\frac{0{,}05}{2};\ 7) = 2{,}365.$$

Так как $|T_H| < T_{кр}$, то нет оснований отвергать гипотезу H_0, т.е. коэффициент a не значим.

4. Получим прогноз объема премиального фонда при $x_{пр} = 3,4$.

По уравнению регрессии найдем:

$$y_{пр} = \bar{y}_{x=3,4} = -0,398 + 3,56 \cdot 3,4 = 11,706.$$

Получим доверительный интеграл для уравнения регрессии:

$$\sigma_{\bar{y}_{x=3,4}} = \sigma_\varepsilon \sqrt{\frac{1}{n} + \frac{(x_{пр} - m_x)^2}{\sum_{i=1}^{n}(x_i - m_x)^2}} = 2,1\sqrt{\frac{1}{9} + \frac{(3,4-2,8)^2}{4,928}} = 0,9;$$

$$p_\partial = 0,95; \quad \alpha = 0,05; \quad T_{кр} = T(0,975;\ 7) = 2,365;$$

$$I_{\bar{y}_{x=3,4}} = (\bar{y}_{x=3,4} \pm T_{кр} \cdot \sigma_{\bar{y}_{x=3,4}}) = (11,706 \pm 2,365 \cdot 0,9) = (9,58;\ \ 13,83).$$

5. Получим прогноз объема премиального фонда при $x_{пр} = 5,0$.

По уравнению регрессии найдем:

$$\bar{y}_{x=5} = -0,398 + 3,56 \cdot 5 = 17,402.$$

Доверительный интервал для данного индивидуального прогноза:

$$\sigma_{\bar{y}_{x=5}} = \sigma_\varepsilon \sqrt{1 + \frac{1}{n} + \frac{(x_{пр} - m_x)^2}{\sum_{i=1}^{n}(x_i - m_x)^2}} = 2,1\sqrt{1 + \frac{1}{9} + \frac{(5-2,8)^2}{4,928}} = 3,75;$$

$$p_\partial = 0,95; \quad \alpha = 0,05; \quad T_{кр} = T(0,975;\ 7) = 2,365;$$

$$I_{\bar{y}_{x=5}} = (\bar{y}_{x=5} \pm T_{кр} \cdot \sigma_{\bar{y}_{x=5}}) = (17,402 \ \pm \ 2,365 \cdot 3,75) = (8,533;\ 26,271).$$

Искомую вероятность определим с использованием нормального закона распределения с параметрами

$$m_{y|x=5} = 17,402; \qquad \sigma_{y|x=5} = 3,75.$$

$$P = 0,5 - \Phi\left(\frac{18 - 17,402}{3,75}\right) = 0,5 - \Phi(0,16) = 0,5 - 0,06 = 0,44.$$

Пример 3.9. Исследовать зависимость урожайности капусты Y (ц/га) от количества использованной воды при искусственном поливе X (м^3/га) в период роста культур.

Опытные данные по 9 полям представлены ниже в таблице:

№ поля	1	2	3	4	5	6	7	8	9
Кол–во воды при поливе x_i (м^3/га)	18,2	16,3	17,0	19,4	20,4	22,1	23,2	24,3	25,1
Урожайность y_i(ц/га)	25,3	23,1	24,2	30,5	35,6	33,7	30,8	28,2	22,5

Выполнить регрессионный анализ исследуемых переменных.

Решение. Можно предположить, что увеличение объема полива приводит к урожайности до некоторого предела, после чего урожайность будет снижаться.

С учетом расположения точек корреляционного поля (таблица) можно предположить, что наиболее подходящим уравнением регрессии будет уравнение параболы:

$$\bar{y}_x = a_0 + a_1 x + a_2 x^2.$$

Параметры модели (a_0, a_1, a_2) находим, используя МНК:

$$Z = \sum_{i=1}^{n} (y_i - \bar{y}_{x_i})^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i - a_2 x_i^2)^2 \to \min.$$

Приравняв частные производные критерия по неизвестным параметрам к нулю $\frac{\partial z}{\partial a_j} = 0;\ (j = 0, 1, 2)$, получим (после соответствующих преобразований) систему нормальных уравнений:

$$\begin{cases} a_0 n + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i; \\ a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + a_2 \sum_{i=1}^{n} x_i^3 = \sum_{i=1}^{n} y_i x_i; \\ a_0 \sum_{i=1}^{n} x_i^2 + a_1 \sum_{i=1}^{n} x_i^3 + a_2 \sum_{i=1}^{n} x_i^4 = \sum_{i=1}^{n} y_i x_i^2. \end{cases}$$

Для расчета необходимых сумм составим вспомогательную таблицу:

i	x_i	y_i	x_i^2	x_i^3	x_i^4
1	18,2	25,3	331,2	6028,6	109719,9
2	16,3	23,1	265,7	4330,7	70591,2
3	17,0	24,2	289,0	4913,0	83521,0
4	19,4	30,5	376,4	7301,4	141646,9
5	20,4	35,6	416,2	8489,7	173189,1
6	22,1	33,7	488,4	10793,9	238544,4
7	23,2	30,8	538,2	12487,2	289702,3
8	24,3	28,2	590,5	14348,9	348678,4
9	25,1	22,5	630,0	15813,3	396912,6
Итого	186,0	253,9	3925,6	84506,6	1852505,8
i	$y_i x_i$	$y_i x_i^2$	y_i^2	$\overline{y}_{x_i}$	$(y_i - \overline{y}_{x_i})^2$
1	460,5	8380,4	640,1	29,0	13,7
2	376,5	6137,4	533,6	20,9	4,8
3	411,4	6993,8	585,6	24,4	0,1
4	591,7	11479,0	930,3	32,0	2,3
5	726,2	14815,3	1267,4	33,3	5,3
6	744,8	16459,4	1135,7	32,8	0,8
7	714,6	16577,8	948,6	30,8	0,0
8	685,3	16651,8	795,2	27,4	0,6
9	564,8	14175,2	506,3	24,0	2,3
Итого	5275,7	111670,1	7342,8	254,6	29,8

Теперь система примет вид:

$$\begin{cases} 9 \cdot a_0 + 186 \cdot a_1 + 3925,6 \cdot a_2 = 253,9; \\ 186 \cdot a_0 + 3925,6 \cdot a_1 + 84506,6 \cdot a_2 = 5275,7; \\ 3925,6 \cdot a_0 + 84506,6 \cdot a_1 + 1852505,8 \cdot a_2 = 111670,1. \end{cases}$$

Решая эту систему (например, методом Гаусса) получим:

$$a_0 = -216,351;\ a_1 = 23,783;\ a_2 = -0,566$$

Уравнение регрессии будет иметь вид:

$$\overline{y}_x = -216,351 + 23,783 \cdot x - 0,566 \cdot x^2$$

Оценим значимость (адекватность) полученной модели.

Вычислим необходимые суммы квадратов отклонений.

$$S_{ост}^2 = \sum_{i=1}^{n} (\bar{y}_{x_i} - y_i)^2 = 29,8;$$

$$S_{общ}^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{(\sum_{i=1}^{n} y_i)^2}{n} = 7342,8 - \frac{(253,9)^2}{9} = 180,0;$$

$$S_{факт}^2 = S_{общ}^2 - S_{ост}^2 = 180,0 - 29,8 = 150,2.$$

Тогда значение критерия адекватности модели будет равно:

$$F_H = \frac{S_{факт}^2 (n - m - 1)}{S_{ост}^2 m} = \frac{150,2 \cdot (9 - 2 - 1)}{29,8 \cdot 2} = 15,12;$$

$$F_{кр} = F(0,95;\ 2;\ 6) = 5,14.$$

Здесь m – количество параметров при независимой переменной (m = 2).

Уравнение регрессии значимо.

Для оценки тесноты связи между переменными X и Y вычислим индекс корреляции: $\eta_{yx} = \sqrt{1 - \frac{S_{ост}^2}{S_{общ}^2}} = \sqrt{1 - \frac{29,8}{180,0}} = \sqrt{0,8344} = 0,913.$

Полученная зависимость весьма веская и значимая.

Коэффициент детерминации $R^2 = 0,8344$ показывает, что вариация урожайности зерновых культур на 83,44 обусловлена регрессией, т.е. изменчивостью количества воды при поливе.

Точность модели оценим стандартной ошибкой уравнения регрессии

$$\sigma_\varepsilon = \sqrt{\frac{S_{ост}^2}{n-2}} = \sqrt{\frac{29,8}{7}} = \sqrt{4,257} = 2,063.$$

Пример 3.10. По семи предприятиям легкой промышленности региона получена информация, характеризующая зависимость объема выпуска продукции (Y, млн. руб.) от объема капиталовложений (X, млн. руб.):

Предприятие i	1	2	3	4	5	6	7
y_i	152	148	146	134	137	136	134
x_i	86	94	100	96	93	104	122

Для характеристики связи объема выпуска продукции от объема капиталовложений построить следующие модели: линейную, степенную, показательную, гиперболическую. Оценить качество каждой модели,

определив индекс корреляции, среднюю относительную ошибку, коэффициент детерминации, *F*–критерий Фишера.

Решение. Имеем $n = 7$ – количество наблюдений, $m = 1$ – количество независимых показателей.

1. Уравнение линейной регрессии имеет вид: $\overline{y}_x = a + bx$.

Рассчитаем основные средние величины показателей.

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{7}\left(86+94+100+96+93+104+122\right) = 99,3;$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{7}\left(152+148+146+134+137+136+134\right) = 141;$$

$$\overline{yx} = \frac{1}{n}\sum_{i=1}^{n} y_i x_i = \frac{1}{7}\left(152\cdot 86+148\cdot 94+146\cdot 100+134\cdot 96+137\cdot 93+\right.$$

$$\left.+136\cdot 104+134\cdot 122\right) = 13954,4;$$

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{7}\left(86^2+94^2+100^2+96^2+93^2+104^2+122^2\right) = 9971.$$

Значение параметров a и b линейной модели определим, используя вычисленные средние значения по данным исходной таблицы.

$$b = \frac{\overline{yx} - \overline{y}\cdot\overline{x}}{\overline{x^2} - \overline{x}^2} = \frac{13954,4 - 141\cdot 99,3}{9971 - 99,3^2} = -0,424;$$

$$a = \overline{y} - b\cdot\overline{x} = 141 - \left(-0,424\right)\cdot 99,3 = 183,14.$$

Получаем уравнение линейной регрессии: $\overline{y}_x = 183,14 - 0,424\cdot x.$

Это означает, что предприятие работает не эффективно. При увеличении капиталовложений на 1 млн. руб. объем выпускаемой продукции уменьшится на 424 тыс. руб. Возможно, это связано с реконструкцией предприятия, когда основное внимание уделяется техническому и технологическому переоборудованию производства, а не максимальному выпуску продукции в рассматриваемый период времени.

Определим линейный коэффициент корреляции по формуле:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}};$$

$$r_{xy}=\frac{(86-99,3)(152-141)+\cdots+(122-99,3)(134-141)}{\sqrt{\left((86-99,3)^2+\cdots+(122-99,3)^2\right)\left((152-141)^2+\cdots+(134-141)^2\right)}}=$$

$$=\frac{-314}{\sqrt{793,43\cdot 334}}=-0,61.$$

Можно сказать, что связь между объемом капиталовложений и объемом выпуска продукции не очень сильная и обратно–пропорциональная.

Рассчитаем коэффициент детерминации: $R^2=r_{xy}^2=0{,}372$.

Вариация объема выпуска продукции на 37,2 % объясняется вариацией фактора объема капиталовложений.

Оценку значимости уравнения регрессии (H_0 : уравнение не значимо) проведем с помощью *F*–критерия Фишера.

Вычислим наблюдаемое значение критерия:

$$F_H=\frac{R^2\left(n-m-1\right)}{\left(1-R^2\right)m}=\frac{0,372\cdot(7-1-1)}{1-0,372}=2,962.$$

По таблице F – распределения найдем критическую точку.

$$F_{кр}=F\left(1-\alpha;\ m;\ n-m-1\right)=F\left(0,95;\ 1;\ 5\right)=6,61.$$

Так как $F_H<F_{кр}$, то нет оснований отвергать гипотезу H_0, т.е. уравнение регрессии не значимо при уровне значимости $\alpha=0{,}05$.

Для построения значимого линейного уравнения необходимы дополнительные данные или поиск нелинейной зависимости межу исследуемыми показателями.

Определим среднюю относительную ошибку:

$$\bar{\varepsilon}_{отн}=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{\varepsilon_i}{y_i}\right|\cdot 100\%=\frac{1}{7}\cdot 0,257\cdot 100\,\%=3,67\,\%.$$

Промежуточные расчеты с используемой анализируемой модели регрессии $\bar{y}_i=y(x_i)=183,14-0,424x_i$ сведены в таблицу:

Предприятие i	x_i	y_i	$y(x_i)$	$\left\lvert\frac{y_i-y(x_i)}{y_i}\right\rvert$
1	86	152	146,676	0,035
2	94	148	143,284	0,032
3	100	146	140,740	0,036
4	96	134	142,436	0,063
5	93	137	143,708	0,049
6	104	136	139,044	0,022
7	122	134	131,412	0,019
Итого:				0,257

В среднем расчетные значения объема выпуска продукции для линейной модели отличаются от фактических значений на 3,67 %.

2. Уравнение степенной модели имеет вид: $y = ax^b$.

Для построения этой модели необходимо произвести линеаризацию переменных.

Для этого произведем логарифмирование обеих частей уравнения: $\lg y = \lg a + b \lg x$.

Обозначим

$$Y = \lg y,\ X = \lg x,\ A = \lg a.$$

Тогда уравнение примет линейный вид:

$$Y = A + bX.$$

Промежуточные расчеты для получения линеаризованной модели регрессии сведены в таблицу:

i	y_i	x_i	Y_i	X_i	$Y_i \cdot X_i$	X_i^2
1	152	86	2,182	1,935	4,221	3,742
2	148	94	2,170	1,973	4,282	3,893
3	146	100	2,164	2,000	4,329	4,000
4	134	96	2,127	1,982	4,217	3,929
5	137	93	2,137	1,969	4,206	3,875
6	136	104	2,134	2,017	4,303	4,068
7	134	122	2,127	2,086	4,438	4,353
Итого	987	695	15,041	13,962	29,996	27,861
Ср.знач.	141	99,3	2,149	1,995	4,285	3,980

Рассчитаем параметры уравнения, используя данные таблицы:

$$b = \frac{\overline{Y \cdot X} - \overline{Y} \cdot \overline{X}}{\overline{X^2} - \overline{X}^2} = \frac{4,285 - 2,149 \cdot 1,995}{3,98 - 1,995^2} = -0,297;$$

$$A = \overline{Y} - b\overline{X} = 2,149 - (-0,297) \cdot 1,995 = 2,74.$$

Уравнение линейной регрессии будет иметь вид:

$$Y = 2,74 - 0,297X.$$

Перейдем к исходным переменным x и y, выполнив потенцирование данного уравнения:

$$\overline{y}_x = 10^{2,74} \cdot x^{-0,297}.$$

Получим уравнение модели регрессии в виде:

$$\overline{y}_x = 549,54 \cdot x^{-0,297}.$$

Для оценки качества полученной модели промежуточные расчеты сведем в таблицу:

Пред-приятие	x_i	$y(x_i)$	$\left(y_i - y(x_i)\right)^2$	$\left(y_i - \bar{y}\right)^2$	$\left\lvert \frac{y_i - y(x_i)}{y_i} \right\rvert$
1	86	146,370	31,692	121	0,037
2	94	142,554	29,656	49	0,037
3	100	139,959	36,500	25	0,041
4	96	141,666	58,763	49	0,057
5	93	143,008	36,094	16	0,044
6	104	138,338	5,465	25	0,017
7	122	131,932	4,276	49	0,015
Итого			202,446	334	0,249

Определим индекс корреляции: $\eta_{yx} = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - y(x_i))^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$,

где $y(x_i) = 549{,}54 \cdot x_i^{-0{,}297}$.

Тогда, используя данные таблицы, получим:

$$\eta_{yx} = \sqrt{1 - \frac{202{,}446}{334}} = 0{,}628.$$

Связь между показателями не очень сильная. Коэффициент детерминации $R^2 = \left(\eta_{yx}\right)^2 = 0{,}628^2 = 0{,}39.$

Вариация результата *Y* (объема выпуска продукции) по степенной модели на 39 % объясняется вариацией фактора *X* (объема капиталовложений).

Оценку значимости уравнения регрессии (H_0 : уравнение не значимо) проведем с помощью *F*–критерия Фишера.

Вычислим наблюдаемое значение критерия:

$$F_H = \frac{R^2\left(n - m - 1\right)}{\left(1 - R^2\right)m} = \frac{0{,}39 \cdot (7 - 1 - 1)}{1 - 0{,}39} = 3{,}2. \quad \text{При этом } F_{кр} = 6{,}61.$$

Так как $F_H < F_{кр}$, то нет оснований отвергать гипотезу H_0, т.е. уравнение регрессии не значимо при уровне значимости $\alpha = 0{,}05$.

Определим среднюю относительную ошибку:

$$\bar{\varepsilon}_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y(x_i)}{y_i}\right| \cdot 100\,\% = \frac{1}{7} \cdot 0{,}249 \cdot 100\,\% = 3{,}56\,\%,$$

где $y(x_i) = 549{,}54 \cdot x_i^{-0{,}297}$.

В среднем расчетные значения результирующего показателя для степенной модели отличаются от фактических значений на 3,56 %.

3. Уравнение показательной модели(зависимости): $y = ab^x$.

Для построения этой модели необходимо произвести линеаризацию переменных.

Для этого осуществим логарифмирование обеих частей уравнения: $\lg y = \lg a + x \lg b$.

Обозначим: $Y = \lg y,\ B = \lg b,\ A = \lg a.$

Получим линейное уравнение регрессии: $Y = A + Bx.$

Промежуточные расчеты для получения линеаризованной модели регрессии сведены в таблицу:

Предприятие i	y_i	x_i	Y_i	$Y_i \cdot x_i$	x_i^2
1	152	86,0	2,182	187,639	7396
2	148	94,0	2,170	204,005	8836
3	146	100,0	2,164	216,435	10000
4	134	96,0	2,127	204,202	9216
5	137	93,0	2,137	198,715	8649
6	136	104,0	2,134	221,888	10816
7	134	122,0	2,127	259,507	14884
Итого	987	695,0	15,041	1492,390	69797
Средн. знач.	141	99,3	2,149	213,200	9971

Рассчитаем параметры уравнения, используя данные таблицы:

$$B = \frac{\overline{Y \cdot x} - \overline{Y} \cdot \overline{x}}{\overline{x^2} - \overline{x}^2} = \frac{213,2 - 2,149 \cdot 99,3}{9971 - 99,3^2} = -0,00121;$$

$$A = \overline{Y} - B \cdot \overline{x} = 2,148 - (-0,00121) \cdot 99,3 = 2,269.$$

Линейное уравнение будет иметь вид: $Y = 2,269 - 0,00121 \cdot x$.

Перейдем к исходным переменным, выполнив потенцирование данного уравнения:

$$\overline{y}_x = 10^{2,269} \cdot (10^{-0,00121})^x = 185,8 \cdot (0,997)^x.$$

Определим индекс корреляции:

$$\eta_{yx} = \sqrt{1 - \frac{\sum_{i=1}^{n} (y_i - y(x_i))^2}{\sum_{i=1}^{n} (y_i - \overline{y})^2}},$$

где $y(x_i) = 185,8 \cdot (0,997)^x$.

Для оценки качества модели расчеты сведем в таблицу:

Предприятие i	y_i	x_i	$y(x_i)$	$\left(y_i - y(x_i)\right)^2$	$\left(y_i - \bar{y}\right)^2$	$\left\lvert \frac{y_i - y(x_i)}{y_i} \right\rvert$
1	152	86	143,184	77,728	121	0,056
2	148	94	139,783	67,518	49	0,054
3	146	100	137,286	75,938	25	0,058
4	134	96	138,946	24,459	49	0,039
5	137	93	140,204	10,264	16	0,026
6	136	104	135,646	0,126	25	0,001
7	134	122	128,505	30,198	49	0,039
Итого				286,231	334	0,271

Тогда, используя данные таблицы, получим:

$$\eta_{yx} = \sqrt{1 - \frac{286,23}{334}} = 0,45.$$

Связь между показателями слабая.

Коэффициент детерминации $R^2 = \left(\eta_{yx}\right)^2 = 0,45^2 = 0,2025.$

Вариация результата объема выпуска продукции на 20,25 % объясняется вариацией объема капиталовложений.

Вычислим наблюдаемое значение F–критерия Фишера для модели:

$$F_H = \frac{0,2025 \cdot (7-1-1)}{1-0,2025} = 1,27. \quad \text{При этом} \quad F_{кр} = 6,61.$$

Так как $F_H < F_{кр}$, то нет оснований отвергать гипотезу H_0, т.е. уравнение регрессии не значимо при уровне значимости $\alpha = 0,05$.

Определим среднюю относительную ошибку:

$$\bar{\varepsilon}_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y(x_i)}{y_i}\right| \cdot 100\% = \frac{1}{7} \cdot 0,271 \ \cdot 100\% = 3,88\%,$$

где $y(x_i) = 184,5 \cdot \left(0,997\right)^{x_i}$.

В среднем расчетные значения результирующего показателя для степенной модели отличаются от фактических значений на 3,88%.

4. Уравнение гиперболической функции: $\bar{y}_x = a + \frac{b}{x}$.

Произведем линеаризацию модели путем замены $X = \frac{1}{x}$.

В результате получим линейное уравнение регрессии: $y = a + bX$.

Промежуточные расчеты для получения линеаризованной модели регрессии сведены в таблицу:

Предприятие i	y_i	x_i	X_i	$y_i \cdot X_i$	X_i^2
1	152	86,0	0,012	1,767	0,0001
2	148	94,0	0,011	1,575	0,0001
3	146	100,0	0,010	1,460	0,0001
4	134	96,0	0,010	1,396	0,0001
5	137	93,0	0,011	1,473	0,0001
6	136	104,0	0,010	1,308	0,0001
7	134	122,0	0,008	1,098	0,0001
Итого	987	695,0	0,071	10,077	0,0007
Средн. знач.	141	99,3	0,010	1,440	0,0001

Рассчитаем параметры модели по данным таблицы.

$$b = \frac{\overline{y \cdot X} - \overline{y} \cdot \overline{X}}{\overline{X^2} - \overline{X}^2} = \frac{1,44 - 141 \cdot 0,01}{0,0001 - (0,01)^2} = 3480,75.$$

$$a = \overline{y} - b \cdot \overline{X} = 141 - 3480,75 \cdot 0,01 = 105,57.$$

Получим следующее уравнение гиперболической модели:

$$\overline{y}_x = 105,57 + \frac{3480,75}{x}.$$

Для оценки качества полученной модели промежуточные расчеты сведем в таблицу:

Предприятие	x_i	y_i	$y(x_i)$	$(y_i - y(x_i))^2$	$(y_i - \overline{y})^2$
1	86	152	142,489	35,476	121
2	94	148	139,105	29,168	49
3	100	146	136,619	31,613	25
4	96	134	138,271	61,275	49
5	93	137	139,523	35,969	16
6	104	136	134,987	9,234	25
7	122	134	127,881	0,010	49
Итого				202,744	334

Определим индекс корреляции: $\eta_{yx} = \sqrt{1 - \frac{\sum_{i=1}^{n}(y_i - y(x_i))^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}}$,

где $y(x_i) = 105,57 + \frac{3480,75}{x_i}$.

Тогда индекс корреляции будет равен величине

$$\eta_{yx} = \sqrt{1 - \frac{202,74}{334}} = 0,63.$$

Связь между показателями и в этой модели слабая.

Коэффициент детерминации незначителен:

$$R^2 = \left(\eta_{yx}\right)^2 = 0,63^2 = 0,39.$$

Вычислим наблюдаемое значение F–критерия Фишера:

$$F_H = \frac{R^2\left(n - m - 1\right)}{(1 - R^2)m} = \frac{0,39 \cdot (7 - 1 - 1)}{1 - 0,39} = 3,2. \quad F_{кр} = 6,61.$$

И эта модель регрессии не значима при уровне значимости $\alpha = 0{,}05$.

Определим среднюю относительную ошибку:

$$\bar{\varepsilon}_{отн} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\varepsilon_i}{y_i}\right| \cdot 100\% = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y(x_i)}{y_i}\right| \cdot 100\%,$$

где $y(x_i) = 105,57 + \dfrac{3480,75}{x_i}$.

Тогда получим:

$$\bar{\varepsilon}_{отн} = \frac{1}{7}\left(\left|\frac{152 - 142,49}{152}\right| + \left|\frac{148 - 139,105}{148}\right| + \left|\frac{146 - 136,62}{146}\right| + \left|\frac{134 - 138,3}{134}\right| + \right.$$

$$\left. + \left|\frac{137 - 139,52}{137}\right| + \left|\frac{136 - 134,99}{136}\right| + \left|\frac{134 - 127,88}{134}\right|\right) \cdot 100\% = 3,42\%.$$

В среднем расчетные значения результирующего показателя для степенной модели отличаются от фактических значений на 3,42 %.