МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

КАЗАНСКИЙ ГОСУДАРСТВЕННЫЙ АРХИТЕКТУРНО-СТРОИТЕЛЬНЫЙ УНИВЕРСИТЕТ

В.Л. Крепкогорский

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Казань
2018

УДК 517.2
ББК 22.171 +22.172
К79
Крепкогорский В.Л.

К79 Теория вероятностей и математическая статистика: Учебное пособие / В.Л. Крепкогорский. – Казань: Изд-во Казанск. гос. архитект.-строит. ун-та, 2018.– 80 с.

ISBN 978-5-7829-0567-5

Печатается по решению Редакционно-издательского совета Казанского государственного архитектурно-строительного университета

Теоретический материал пособия предназначен для использования на занятиях и самостоятельного изучения студентами первого и второго курса, обучающимися по программам для бакалавров: 08.03.01 «Строительство»; 09.03.02 «Информационные системы и технологии»; 38.03.02 «Менеджмент».

В конце каждого раздела пособия приведены практические задания для самостоятельной проработки теоретического материала.

Рецензенты:
Доктор физико-математических наук,
профессор Казанского кооперативного института
А.В. Поташев

Кандидат физико-математических наук,
доцент кафедры прикладной математики Казанского государственного архитектурно-строительного университета
Ф.Г. Габбасов

УДК 517.2
ББК 22.171+22.172

ISBN 978-5-7829-0567-5

Глава I. Случайные события

1. Вероятность сложных событий

1.1. Случайные события, действия над событиями

Теорией вероятностей называется математическая наука, изучающая закономерности в случайных явлениях.

Основными понятиями теории вероятности являются испытание и случайное событие.

Испытание – это действие, которое может повторяться при неизменных условиях любое число раз.

Случайное событие – это любой случайный результат испытания, то есть событие, которое может произойти или не произойти в результате испытания.

Случайные события обозначаются заглавными буквами: $A, B, C, \ldots$

Достоверным называется событие, которое обязательно должно произойти в результате испытания. Обозначается – Ω.

Невозможным называется событие, которое не может произойти в результате испытания. Обозначается – $\emptyset$.

Два события называются **несовместными**, если они не могут произойти при одном испытании.

С испытанием может быть связано множество элементарных исходов – это множество попарно несовместных случайных событий, одно (и только одно) должно произойти при испытании. Говорят, что множество элементарных событий образует **полную группу событий**.

Примеры

1. Испытание – бросание игрального кубика имеет шесть элементарных исходов: E_1 – «выпала цифра 1», … , E_6 – «выпала цифра 6».

2. Испытание – произведен выстрел по цели. Оно имеет два элементарных исхода: «попали в цель», «не попали в цель».

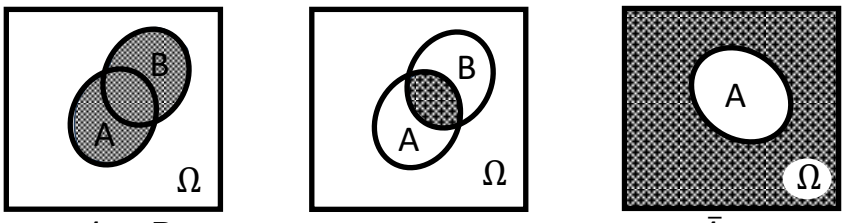
Элементарный исход A **благоприятствует** событию B, если из того, что произошло A следует, что произошло событие B.

Арифметические действия над событиями

Суммой событий A и B называется событие $A + B$ – «произошло хотя бы одно из событий – A или B».

Произведением событий A и B называется событие AB – «произошли оба события – A и B».

Событие $\bar{A}$ – «событие A не произошло» – называется **противоположным** к событию A.



$A + B$ $\quad$ AB $\quad$ $\bar{A}$

Рис. 1

Изображение операций над событиями на диаграммах Эйлера-Венна (рис.1) .

Действия над событиями обладают следующими свойствами:

1) $A + B = B + A, \quad AB = BA$;
2) $(A + B)C = AC + BC$;
3) $A + A = A, \ AA = A$;
4) $A + \bar{A} = \Omega, \ A\bar{A} = \emptyset$;
5) $\overline{A + B} = \bar{A}\bar{B}, \qquad \overline{AB} = \bar{A} + \bar{B}$;
6) $\Omega + \emptyset = \Omega, \ \Omega\emptyset = \emptyset$.

1.2. Определения вероятности события

а) **Статистическое определение вероятности**

Пусть n – число проведенных испытаний, n_A – число испытаний, в которых произошло событие A.

Относительной частотой (частостью) события A называют отношение:

$$w(A) = \frac{n_A}{n}.$$

Вероятностью события A называется число $P(A)$, равное пределу, к которому стремится относительная частота события при неограниченном увеличении числа испытаний.

б) **Классическое определение вероятности**

Пусть исходы некоторого испытания образуют полную группу событий и равновозможны. Такие исходы будем называть *случаями* или *шансами*.

Вероятность события A равна:

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

где m – число случаев, благоприятствующих событию A, а n – общее число случаев.

1.3. Элементы комбинаторики

Требуется выбрать m элементов из общего числа n. Полученные наборы элементов будем называть *комбинациями*. Если комбинации отличаются либо составом элементов, либо порядком расположения (либо и тем, и другим), то такие комбинации называют *размещениями* из n элементов по m. Число размещений равно:

$$A_n^m = n(n-1)(n-2)\dots(n-m+1) = n!/(n-m)!$$

Пример. Расписание одного дня состоит из 5 уроков. Определить число вариантов расписания при выборе из 11 дисциплин.

Решение. Число вариантов расписания равно числу размещений из 11 по 5: $A_{11}^5 = 11 \cdot 10 \cdot 9 \cdot 8 \cdot 7 = 55440$.

Если комбинации из n элементов по m отличаются только составом элементов, то их называют *сочетаниями* из n элементов по m. Число сочетаний равно:

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

Пример. В шахматном турнире участвуют 16 человек. Сколько партий должно быть сыграно в турнире, если между любыми двумя участниками должна быть сыграна одна партия?

Решение. Число партий равно числу сочетаний из 16 по 2:

$$C_{16}^2 = \frac{16 \cdot 15}{2} = 120.$$

1.4. Вероятность суммы событий

Теорема сложения. Вероятность суммы событий:

$$P(A+B) = P(A) + P(B) - P(AB).$$

Для несовместных событий:

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

<u>*Замечание*</u>. Последняя формула верна для любого числа слагаемых.

1.5. Вероятность произведения событий

Два события называются *независимыми*, если вероятность одного из них не зависит от того, произошло другое или нет.

Условной вероятностью $P(A/B)$ события B называется вероятность события B при условии, что событие A произошло.

Теорема умножения. Для любых случайных событий A и B справедлива формула:

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

Если события A и B независимы, то

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

Вероятность противоположного события

Так как $A + \bar{A} = \Omega$ и события A и $\bar{A}$ – несовместные, то $P(A) + P(\bar{A}) = 1$. Следовательно,

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

В случае, когда число событий больше двух, можно записать формулы:

$$a)\ P(A_1 + A_2 + \cdots A_n) = P(A_1) + P(A_2) + \cdots P(A_n)\ -$$

для несовместных событий;

$$b)\ P(A_1 A_2 \ldots A_n) = P(A_1)P(A_2) \ldots P(A_n) -$$

для независимых событий.

Для совместных, но независимых событий можно использовать равенство:

$$A_1 + A_2 + \cdots A_n = \overline{\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n} \Rightarrow P(A_1 + A_2 + \cdots A_n) =$$

$$1 - P(\bar{A}_1)P(\bar{A}_2) \ldots P(\bar{A}_n).$$

Примеры

1) Два стрелка сделали по выстрелу по мишеням. Вероятность попадания для первого стрелка равна 0,6, а для второго – 0,3. Найти вероятности следующих событий:

A – «оба попали»,

B – «хотя бы один стрелок попал»,

C – «только один попал»,

D – «оба не попали».

Решение. Рассмотрим элементарные события A_1 – «первый стрелок попал», A_2 – «второй стрелок попал». Тогда $P(A_1) = 0{,}6$ и $P(A_2) = 0{,}3$. Выразим все события через A_1 и A_2. При этом события A_1 и A_2 можно считать независимыми.

$$A = A_1 A_2 \Rightarrow P(A) = P(A_1)P(A_2) = 0{,}6 \cdot 0{,}3 = 0{,}18.$$

$$B = A_1 + A_2 \Rightarrow P(B) = P(A_1) + P(A_2) - P(A_1 A_2) =$$

$$= 0{,}6 + 0{,}3 - 0{,}18 = 0{,}72.$$

$$C = \bar{A}_1 A_2 + A_1 \bar{A}_2 \Rightarrow P(C) = P(\bar{A}_1 A_2) + P(A_1 \bar{A}_2) =$$
$$= (1 - 0{,}6) \cdot 0{,}3 + 0{,}6 \cdot (1 - 0{,}3) = 0{,}12 + 0{,}42 = 0{,}54.$$

$$D = \bar{A}_1 \bar{A}_2 \Rightarrow P(D) = P(\bar{A}_1) P(\bar{A}_2) = (1 - 0{,}6) \cdot (1 - 0{,}3) = 0{,}28.$$

2) Среди 25 студентов, из которых 15 девушек, разыгрывается четыре билета, причем каждый может выиграть только один билет. Какова вероятность того, что все 4 билета достанутся девушкам?

Решение. Количество вариантов выбора 4 человек из общего числа 25 равно C_{25}^4. Количество вариантов выбора 4 человек из 15 девушек равно C_{15}^4. Тогда вероятность того, что все билеты достанутся девушкам, равна:

$$\frac{C_{15}^4}{C_{25}^4} \approx 0{,}108.$$

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

1.1. ОТК обнаружил 5 бракованных деталей из 100 проверенных. Найти относительную частоту появления стандартных деталей.

1.2. В группе 16 девушек и 9 юношей. Из стопки тетрадей с контрольными работами, выполненными студентами этой группы, наудачу извлекается одна тетрадь. Какова вероятность, что эта тетрадь принадлежит юноше?

1.3. Для зачета билеты составляются из 100 базовых вопросов. Студент знает ответ на 40 вопросов. Чтобы сдать зачет, студент должен правильно ответить на 3 вопроса. Какова вероятность того, что студент сдаст зачет?

1.4. Порядок выступления 7 участников конкурса определяется жребием. Сколько различных вариантов жеребьевки при этом возможны?

1.5. В группе 30 человек. Необходимо выбрать старосту, его заместителя и профорга. Сколько существует способов это сделать?

1.6. В шахматном турнире участвуют 16 человек. Сколько партий должно быть сыграно в турнире, если между любыми двумя участниками должна быть сыграна одна партия?

1.7. В районах Крайнего Севера работы прекращаются при температуре –40° или при скорости ветра свыше 20 м/сек. Вероятность низкой температуры равна: $p_1 = 0{,}06$, вероятность ветра свыше 20 м/сек – $p_2 = 0{,}07$. Какова вероятность того, что работы будут приостановлены из-за погодных условий?

1.8. В двух ящиках по 100 деталей. В первом ящике 5 бракованных деталей, во втором – 8. Из каждого ящика наугад вынимают по одной детали. Найти вероятности событий:

A – «обе детали бракованные»,

B – «хотя бы одна деталь бракованная»,

C – «только одна деталь бракованная»,

D – «ни одна деталь не бракованная»,

E – «хотя бы одна деталь не бракованная».

1.9. Вероятность того, что на экзамене студент получит оценку «отлично», равна $p_1 = 0{,}3$, «хорошо» – $p_2 = 0{,}4$, «удовлетворительно» – $p_3 = 0{,}2$. Какова вероятность того, что студент сдаст экзамен?

1.10. Три стрелка стреляют по мишени. Вероятность попадания для первого стрелка равна 0,8, для второго – 0,6, для третьего – 0,5. Найти вероятность событий:

A – «все попали»,

B – «все не попали»,

C – «хотя бы один попал».

1.11. Машина состоит из трех блоков. Известна вероятность выхода из строя каждого блока: первого – 0,1; второго – 0,2; третьего – 0,4. Найти вероятности событий:

A – «вышли из строя три блока»,

B – «вышел из строя хотя бы один блок»,

C – «не вышел из строя ни один блок»,

D – «вышли из строя два блока».

1.12. Монета брошена четыре раза. Найти вероятность того, что орел выпадет хотя бы один раз.

1.13. Среди 25 студентов, среди которых 15 девушек, разыгрывается четыре билета, причем каждый раз может выиграть только один билет. Какова вероятность того, что среди обладателей билета окажутся а) 4 юноши; б) три юноши и одна девушка.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1.14. В витрине магазина в ряд выставлено 6 мобильных телефонов разных моделей. Определить число способов их размещения.

1.15. Клиент желает получить кредит в одном из не связанных между собой банков. Для данного клиента вероятность получить кредит в первом банке равна 0,4, во втором – 0,7. Какова вероятность, что кредит будет получен?

1.16. Вероятность того, что первый спортсмен выполнит норму мастера спорта, равна 0,9, второй – 0,7. Определить вероятность того, что: а) оба спортсмена не выполнят норму мастера спорта; б) только один спортсмен выполнит норму мастера спорта; в) хотя бы один спортсмен выполнит норму.

1.17. В первой группе 24 студента, из них 4 – отличники, во второй группе 30 студентов, из них 3 – отличники. Из каждой группы вызывается по одному курсанту. Найти вероятности событий:

A – «оба отличники»,

B – «хотя бы один – отличник»,

C – «только один – отличник»,

D – «ни один не является отличником».

1.18. Три лампочки соединены последовательно. Вероятности их перегорания за определенный срок $p_1 = 0{,}2,\ p_2 = 0{,}3,\ p_3 = 0{,}4.$ Найти вероятность того, что лампочки будут перегорать?

2. Формула полной вероятности и формула Байеса Повторные независимые испытания

2.1. Формула полной вероятности и формула Байеса

Рассмотрим некоторое испытание и случайное событие A, которое может произойти в результате испытания при появлении одного из событий $H_1, H_2, \dots, H_n$, удовлетворяющих условиям:

1) события $H_1, H_2, \dots, H_n$ несовместные;

2) события $H_1, H_2, \dots, H_n$ образуют полную группу, то есть:

$H_1 + H_2 + \cdots + H_n = \Omega \Rightarrow P(H_1) + P(H_2) + \cdots + P(H_n) = 1.$

События $H_1, H_2, \dots, H_n$ называют **гипотезами**.

Формула для вычисления вероятности события A имеет вид:

$$P(A) = \sum_{k=1}^{n} P(H_k)P(A/H_k)$$

и называется **формулой полной вероятности**.

Вероятность гипотезы H_k $(k = 1, 2, , \dots, n)$ при условии, что событие A произошло, находится по **формуле Байеса:**

$$P(A/H_k) = \frac{P(H_k)P(A/H_k)}{P(A)}.$$

Примеры

1) В начале учебного года в группе по списку было 30 студентов. Из них 20 получили стипендию. В середине семестра один студент взял академический отпуск. Определить вероятность того, что случайно отобранный после этого студент окажется стипендиатом.

Решение. Результат будет зависеть от того, получал ли стипендию ушедший в отпуск студент. Рассмотрим две гипотезы: H_1 – в отпуск ушел стипендиат; H_2 – ушедший в отпуск не получал стипендии. Тогда:

$$P(H_1) = \frac{20}{30} = \frac{2}{3},\ P(H_2) = \frac{10}{30} = \frac{1}{3}.$$

Если H_1 – верно, то осталось 29 студентов, из которых 19 получают стипендию. Поэтому $P(A/H_1) = 19/29$.

Если H_2 – верно, то из 29 студентов 20 получают стипендию. Поэтому $P(A/H_2) = 20/29$. Подставляя в формулу полной вероятности, получаем:

$$P(A) = \frac{2}{3} \cdot \frac{19}{29} + \frac{1}{3} \cdot \frac{20}{29} = \frac{38 + 20}{3 \cdot 29} = \frac{58}{87}.$$

2) В отряде 10 стрелков. Для пяти из них вероятность попадания в цель – 0,8, для трех других – 0,5 и для двух остальных – 0,25. Выстрел, произведенный кем-то из стрелков, дал попадание. Определить вероятность того, что этот выстрел был сделан стрелком из второй группы.

Решение. Пусть A – событие, состоящее в том, что случайно выбранный стрелок поразил цель, H_k – был выбран стрелок из $k-$ой по номеру группы. Тогда вероятность $P(A)$ может быть найдена по формуле полной вероятности.

$$P(A) = \frac{5}{10} \cdot 0{,}8 + \frac{3}{10} \cdot 0{,}5 + \frac{2}{10} \cdot 0{,}25 = 0{,}6.$$

Наша задача – найти $P(H_2/A)$. Найдем ее по формуле Байеса:

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{\frac{3}{10} \cdot 0{,}5}{0{,}6} = 0{,}25.$$

2.2. Повторные независимые испытания. Формула Бернулли

Проводится серия испытаний.

Теорема. Если вероятность p наступления события A в каждом испытании постоянна, то вероятность $P_n(m)$ того, что событие A наступит m раз в n независимых испытаниях, равна:

$$P_n(m) = C_n^m p^m q^{n-m},$$

где $q = 1 - p$.

Это равенство называется формулой Бернулли.

Считаем, что $0! = 1,\ C_n^0 = C_n^n = 1$.

Пример. Вероятность «малому предприятию быть банкротом» за время t равна 0,2. Найти вероятность того, что из пяти малых предприятий за время t обанкротится: а) два; б) более двух.

Решение: а) требуется найти:

$$P_5(2) = C_5^2 \cdot 0{,}2^2 0{,}8^3 = \frac{5 \cdot 4}{1 \cdot 2} \cdot 0{,}04 \cdot 0{,}512 = 0{,}2048 \approx 0{,}2;$$

б) требуется найти:

$$\sum_{k=3}^{5} P_5(k) = C_5^3 \cdot 0{,}2^3 \cdot 0{,}8^2 + C_5^4 \cdot 0{,}2^4 \cdot 0{,}8 + C_5^5 \cdot 0{,}2^5 =$$

$$= \frac{5 \cdot 4 \cdot 3}{1 \cdot 2 \cdot 3} \cdot 0{,}2^3 + \frac{5 \cdot 4 \cdot 3 \cdot 2}{1 \cdot 2 \cdot 3 \cdot 4} 0{,}2^4 + 0{,}2^5 = 10 \cdot 0{,}008 \cdot 0{,}64 + 5 \cdot 0{,}0016 \cdot 0{,}8 +$$

$$+0{,}00032 = 0{,}0512 + 0{,}0064 + 0{,}00032 = 0{,}05792.$$

При решении задач этого типа можно использовать функции программы EXCEL.

При больших значениях n и m решение подобных задач с помощью формулы Бернулли невозможно. В этом случае используются приближенные формулы.

Формула Пуассона. Рекомендуется использовать при большом количестве испытаний n и при условии: $\lambda = np \leq 10$.

$$P_n(m) \approx \frac{\lambda^m e^{-\lambda}}{m!} = P(m; \lambda).$$

Пример. На факультете насчитывается 1825 студентов. Какова вероятность того, что 1 сентября является днем рождения одновременно четырех студентов факультета?

Решение. Для простоты считаем, что в году 365 дней. Вероятность, что день рождения студента: $p = 1/365$. Так как число испытаний $n = 1825$ велико и $\lambda = np = 1825/365 = 5 < 10$, то можно применить формулу Пуассона $P_{1825}(4) = P(4; 5) = 0{,}1775$ (По табл. функции Пуассона. Приложение 3).

Формула Лапласа. Вероятность m событий при n испытаниях равна:

$$P_n(m) \approx \frac{f(x)}{\sqrt{npq}}, \text{где } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \text{ и } x = \frac{m - np}{\sqrt{npq}}.$$

Рекомендуется применять при больших n и выполнении условия: $np \geq 5$.

Вероятность, что число событий заключено между $a \le m \le b$, определяется интегральной формулой Лапласа.

$$P_n(a \le m \le b) = \frac{1}{2}\,[\Phi(x_2) - \Phi(x_1)],$$

где

$$\Phi(x) = \frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt, \quad x_1 = \frac{a - np}{\sqrt{npq}}, \quad x_2 = \frac{b - np}{\sqrt{npq}}.$$

Для симметричного относительно np интервала интегральная формула Лапласа имеет вид:

$$P_n(|m - np| \le \varepsilon) \approx \Phi\left(\frac{\varepsilon}{\sqrt{npq}}\right). \qquad (2.1)$$

Пример. Завод отправил на базу 10 000 стандартных изделий. Среднее количество изделий, повреждаемых при транспортировке, составляет 0,02%. Найти вероятность того, что из 10 000 изделий будет повреждено 3.

Решение. Вероятность того, что изделие будет повреждено при транспортировке, равна: $p = 0{,}0002$. Так как число p мало, $n = 10000$ – велико и $\lambda = np = 10000 \cdot 0{,}0002 = 2 \le 10$, следует применить формулу Пуассона: $P_{10000}(3) \approx \frac{2^3 e^{-2}}{3!} = 0{,}1804$.

Пример. По результатам проверок налоговыми инспекциями установлено, что в среднем каждое второе малое предприятие региона имеет нарушения финансовой дисциплины. Найти вероятность того, что из 1000 зарегистрированных в регионе малых предприятий имеют нарушения финансовой дисциплины: а) 480 предприятий; б) не менее 480.

Решение. По условию, $p = 0{,}5$. Так как $n = 1000$ достаточно велико (условие $npq = 1000 \cdot 0{,}5 \cdot (1 - 0{,}5) = 250 \ge 5$ выполнено), то можно применять формулы Лапласа.

а) Вычислим: $x = \frac{480 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = -1{,}265$, затем:

$$P_{1000}(480) \approx \frac{f(-1{,}265)}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = \frac{f(-1{,}265)}{\sqrt{250}} = \frac{0{,}1792}{\sqrt{250}} = 0{,}0113.$$

б) Необходимо найти $P_{1000}(m \ge 480) = P_{1000}(480 \le m \le 1000)$. Применим интегральную формулу Лапласа. Предварительно вычислим:

$$x_1 = \frac{480 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = -1{,}265, \quad x_2 = \frac{1000 - 1000 \cdot 0{,}5}{\sqrt{1000 \cdot 0{,}5 \cdot 0{,}5}} = 31{,}6.$$

Теперь

$$P_{1000}(480 \le m \le 1000) \approx \frac{1}{2}[\Phi(31{,}6) - \Phi(-1{,}265)] =$$
$$\frac{1}{2}[\Phi(31{,}6) + \Phi(+1{,}265)] \approx \frac{1}{2}(1 + 0{,}7941) = 0{,}897.$$

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

2.1. В группе спортсменов 20 лыжников, 6 велосипедистов, 4 бегуна. Вероятность выполнить квалификационную норму такова: для лыжника – 0,9; для велосипедиста – 0,8; для бегуна – 0,75. Найти: 1) вероятность того, что выбранный наугад спортсмен выполнит норму; 2) вероятность того, что спортсмен – лыжник, если известно, что он выполнил норму.

2.2. В магазин поступают однотипные изделия с трех заводов, причем первый завод поставляет 50% изделий, второй – 30%, третий завод 20%. Среди изделий первого завода 70% первосортных, второго завода – 80%, третьего завода – 90%. Куплено одно изделие.

1) Найти вероятность того, что оно первосортное.

2) Изделие оказалось первосортным. Найти вероятность того, что купленное изделие выпущено первым заводом.

2.3. В урне находится 10 шаров, из них 6 белых и 4 черных. Из урны выпал один шар неизвестного цвета. Затем извлекли еще один шар. Найти вероятность, что он оказался белый.

2.4. В некотором водоеме карпы составляют 80% всех рыб. Определить вероятность того, что из пяти выловленных в этом водоеме рыб окажется: 1) 4 карпа; 2) не менее 3 карпов; 3) не более 1 карпа.

2.5. В среднем по 15% договоров страховая компания выплачивает страховую сумму. Найти вероятность того, что из десяти договоров будут выплачены страховые суммы по: а) трем договорам; б) менее чем по двум договорам.

2.6. Два равносильных противника играют в шахматы. Что более вероятно: а) выиграть 2 партии из 4 или 3 партии из 6? б) не менее 2 партий из 4 или не менее 3 партий из 6? (Ничьи в расчет не принимаются).

2.7. Монету бросают 10 раз. Какова вероятность, что «герб» выпадет: а) 6 раз; б) 5 раз.

2.8. В банк отправлено 4000 пакетов денежных знаков. Вероятность, что пакет содержит недостаточное или избыточное число денежных знаков, равна 0,0001. Найдите вероятность того, что при проверке будет обнаружено: а) три ошибочно упакованных пакета; б) не более трех пакетов.

2.9. Аудиторную работу по теории вероятности с первого раза успешно выполняют 50% студентов. Найти вероятность того, что из 400 студентов работу успешно выполнят: а) 180 студентов, б) не менее 180 студентов.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

2.10. Страховая компания разделяет застрахованных по классам риска: I класс – малый риск, II класс – средний, III класс – большой риск. Среди этих клиентов 50% – первого класса риска, 30% – второго и 20% – третьего. Вероятность необходимости выплачивать страховое вознаграждение для первого класса риска равна 0,01, второго – 0,03, третьего – 0,08. Какова вероятность того, что: а) случайно выбранный застрахованный получит денежное вознаграждение за период страхования; б) получивший денежное вознаграждение застрахованный относится к группе малого риска?

2.11. Из 10 стрелков можно выделить 2 отличных, 5 хороших и 3 удовлетворительных. Вероятность попадания в мишень при одном выстреле для отличного стрелка равна 0,9; для хорошего – 0,7; для удовлетворительного – 0,3.

1) Найти вероятность того, что наугад выбранный стрелок попадет в мишень.

2) Стрелок попал в мишень. Какова вероятность того, что он из группы хороших стрелков.

2.12. Приобретено 4 лотерейных билета. Вероятность выигрыша по каждому – 0,2. Определить вероятности всевозможных исходов.

2.13. Вероятность того, что пассажир опоздает к отправлению поезда, равна 0,01. Найти вероятность того, что из 800 пассажиров 8 человек опоздают.

Указание: применить приближенную формулу Пуассона.

2.14. Монету бросают 100 раз. Сколько раз выпадет «герб» обозначим через m. а) Какова вероятность, что m =50; б) какова вероятность, что $40 \leq m \leq 60$.

Глава II. Случайные величины

3. Дискретная случайная величина

3.1. Дискретная случайная величина, законы распределения, числовые характеристики

Определения

Случайная величина (сл.в.) – это числовая функция, заданная на множестве элементарных событий испытания. **Обозначения**: случайные величины $X, Y, Z, \ldots$ их возможные значения – $x, y, z, \ldots$

Дискретная случайная величина – это величина, множество значений которой образует конечную или бесконечную последовательность.

Закон распределения случайной величины – это соотношение, устанавливающее связь между возможными значениями случайной величины и их вероятностями.

Закон распределения может быть задан с помощью ряда распределения, многоугольника распределения или функции распределения.

1) Ряд распределения – это таблица:

X	x_1	x_2	$\cdots$	x_n
$P\{X = x_k\}$	p_1	p_2	$\cdots$	p_n

Свойства p_k: а) $0 \le p_k \le 1$; б) $\sum_k^n p_k = 1$.

2) Многоугольник распределения (рис. 2).

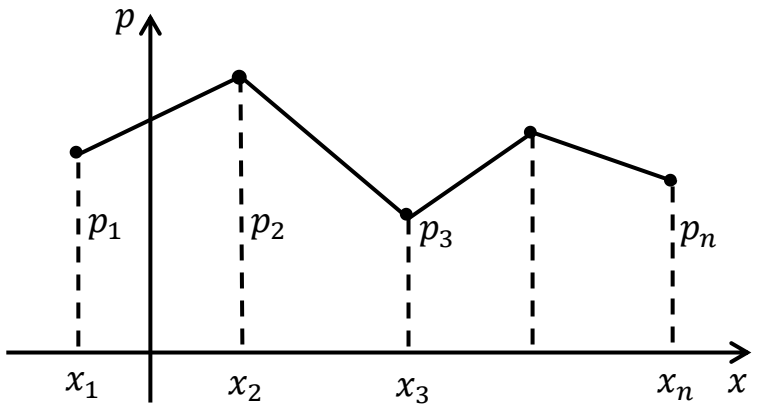


Рис.2

3) Функция распределения $F(x) = P\{X < x\}$.

Свойства функции распределения: а) $0 \le F(x) \le 1$; б) $F(-\infty) = 0$, $F(\infty) = 1$; в) $F(x)$ – неубывающая функция.

Числовые характеристики дискретной случайной величины X, заданной рядом распределения:

1) математическое ожидание $M[X] = m_x = \sum_{k=1}^{n} x_k p_k$;

2) дисперсия $D[X] = D_x = \sum_{k=1}^{n} (x_k - m_x)^2 p_k = \sum_{k=1}^{n} x_k^2 p_k - m_x^2$;

3) среднее квадратическое отклонение $\sigma[X] = \sigma_x = \sqrt{D_x}$.

Вычисление вероятности:

$$P\{a \le X \le b\} = \sum_{a \le x_k \le b} p_k.$$

Пример. Дискретная сл. в. X задана рядом распределения.

X	1	2	6	8
$P\{X = x_k\}$	0,1	0,25	0,4	0,25

Требуется:

1) построить многоугольник распределения;

2) найти функцию распределения $F(x)$ и построить ее график.

3) найти m_x, D_x, σ_x;

4) найти $P\{2 \le X \le 8\}$.

Решение

1) Построим многоугольник распределения (рис. 3).

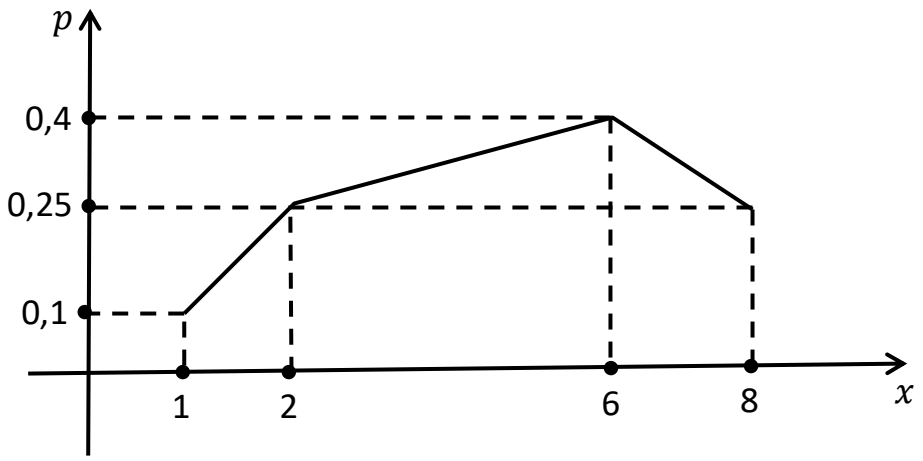


Рис. 3

2) Функцию $F(x)$ найдем в виде таблицы, и построим ее график (рис. 4).

x	$F(x) = P\{X < x\}$
$(-\infty, 1]$	0
$(1, 2]$	$0{,}1$
$(2, 6]$	$0{,}1 + 0{,}25 = 0{,}35$
$(6, 8]$	$0{,}35 + 0{,}4 = 0{,}75$
$(8, \infty)$	$0{,}75 + 0{,}25 = 1$

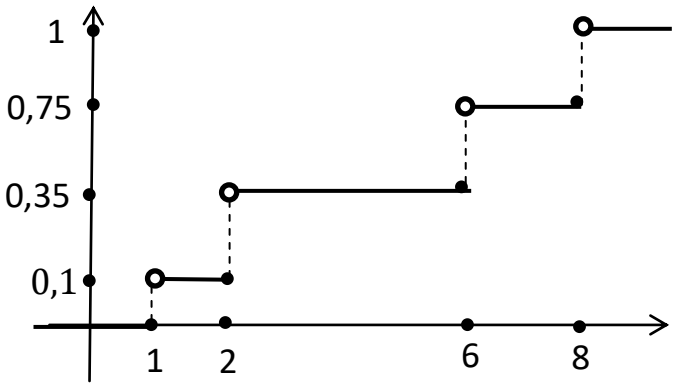


Рис.4

3) $m_x = \sum_{k=1}^{n} x_k p_k = 0{,}1 \cdot 1 + 0{,}25 \cdot +0{,}4 \cdot 6 + 0{,}25 \cdot 8 = 5;$

$$D_x = \sum_{k=1}^{n} x_k^2\, p_k - m_x^2 =$$

$$= 0{,}1 \cdot 1 + 0{,}25 \cdot 4 + 0{,}4 \cdot 36 + 0{,}25 \cdot 64 - 25 = 6{,}5;$$

$$P\{2 \le X \le 8\} = 0{,}25 + 0{,}4 + 0{,}25 = 0{,}9.$$

3.2. Биномиальный закон распределения

Производится серия из n испытаний, в каждом из которых событие A может произойти с вероятностью p. Дискретная случайная величина X «число испытаний, в которых произошло событие A». Ряд распределения в этом случае определяется формулой Бернулли.

$$P\{X = k\} = P_n(k) = C_n^k p^k q^{n-k}, \quad k = 0, 1, 2, \dots, n.$$

Числовые характеристики:

а) математическое ожидание $m_x = np$,

б) дисперсия $D_x = npq$,

в) среднее квадратическое отклонение: $\sigma_x = \sqrt{D_x} = \sqrt{npq}$.

Наивероятнейшее значение Mo (мода распределения) числа наступлений события A при проведении n испытаний вычисляется по формуле:

$$np - q \le Mo \le np + p.$$

Пример. Стрелок производит 12 выстрелов. Вероятность попадания в мишень при каждом выстреле равна 0,7. Случайная величина X – число попаданий. Определить:

а) Mo – наиболее вероятное число попаданий и вероятность этого числа попаданий;

б) математическое ожидание, дисперсию и среднее квадратическое отклонение сл. вел. X.

Решение. В нашем примере $n = 12,\ p = 0{,}7,\ q = 0{,}3.$ Тогда

$$m_x = np = 12 \cdot 0{,}7 = 8{,}4;\ \ D_x = npq = 12 \cdot 0{,}7 \cdot 0{,}3 = 2{,}52;$$

$$\sigma_x = \sqrt{2{,}52} = 1{,}59.$$

Чтобы найти Mo, составим неравенства: $12 \cdot 0{,}7 - 0{,}3 \leq Mo \leq 12 \cdot 0{,}7 + 0{,}7 \ \Leftrightarrow 8{,}1 \leq Mo \leq 9{,}1$. Единственное целое число в этом интервале равно 9. Значит, наиболее вероятное число попаданий равно 9.

Вероятность этого числа попаданий: $P_{12}(9) = C_{12}^{9} \cdot 0{,}7^9 \cdot 0{,}3^3 \approx 0{,}24.$

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

3.1. Дискретная сл. в. X задана рядом распределения:

X	2	4	5	10
$P\{X = x_k\}$	0,3	0,1	0,4	0,2

Требуется:

1) построить многоугольник распределения;
2) найти функцию распределения $F(x)$ и построить ее график;
3) найти m_x, D_x, σ_x;
4) найти $P\{2 \leq X \leq 6\}$.

3.2. Дискретная сл.в. X задана рядом распределения:

X	-1	0	4	10
$P\{X = x_k\}$	0,1	0,5	0,2	0,2

Требуется:

1) построить многоугольник распределения;
2) найти функцию распределения $F(x)$ и построить ее график;
3) найти m_x, D_x, σ_x;
4) найти $P\{0 \leq X \leq 5\}$.

3.3. Два стрелка независимо друг от друга делают по одному выстрелу в мишень. Вероятность попадания в мишень первого стрелка равна 0,8 , а второго – 0,6. Случайная величина X – суммарное число попаданий в мишень. Построить ряд распределения случайной величины X, найти m_x, D_x, σ_x.

3.4. Дискретная сл. в. X задана рядом распределения:

X	3	5
$P\{X = x_k\}$	p_1	p_2

Известно, что $m_x = 3{,}6$. Найти p_1и p_2.

3.5. Два шахматиста играют три партии (без учета ничейных игр). Известна вероятность выигрыша первого игрока в каждой партии, равная: $p = 0{,}6$. Случайная величина X – «число партий, выигранных первым игроком».

Требуется:

1) Построить ряд распределения случайной величины X;

2) найти m_x, D_x, σ_x; 3) найти $P\{2 \leq X\}$; 4) найти моду распределения.

3.6. Игральная кость бросается 100 раз. Определить наивероятнейшее число выпадений шести очков на верхней грани кости.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

3.7. Дискретная сл. в. X задана рядом распределения:

X	-2	0	1	3
$P\{X = x_k\}$	0,3	0,2	0,4	0,1

Требуется:

1) построить многоугольник распределения;

2) найти функцию распределения $F(x)$ и построить ее график;

3) найти m_x, D_x, σ_x;

4) найти $P\{0 \leq X \leq 2\}$.

3.8. Дискретная сл. в. X задана рядом распределения:

X	0	x_2	9
$P\{X = x_k\}$	0,1	0,5	0,4

Математическое ожидание $m_x = 5{,}6$. Найти x_2.

3.9. Вероятность выигрыша по лотерейному билету равна 0,4. Составить закон распределения случайной величины X – числа выигрышных билетов из имеющихся четырех билетов.

3.10. Необходимо исследовать 1000 проб руды. Вероятность промышленного содержания металла в каждой пробе равна 0,3. Определить математическое ожидание и дисперсию числа проб с промышленным содержанием металла.

4. Непрерывная случайная величина. Равномерный закон распределения случайной величины

4.1. Непрерывная случайная величина

Случайная величина X называется **непрерывной**, если ее функция распределения непрерывна в любой точке и дифференцируема всюду, кроме может быть, отдельных точек.

Вероятность любого отдельно взятого значения непрерывной случайной величины равна нулю.

Плотностью вероятности (**плотностью распределения** или просто **плотностью**) непрерывной случайной величины X называется производная ее функции распределения:

$$f(x) = F'(x).$$

Свойства функции плотности распределения $f(x)$:

1. Плотность распределения неотрицательна: $f(x) \geq 0$;
2. $P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x)dx$;
3. $F(x) = \int_{-\infty}^{x} f(x)dx$;
4. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Числовые характеристики непрерывной случайной величины X:

Математическое ожидание:

$$a = M(X) = m_x = \int_{-\infty}^{\infty} x\, f(x)dx.$$

Дисперсия:

$$D(X) = D_x = \int_{-\infty}^{\infty} (x-a)^2\, f(x)dx = \int_{-\infty}^{\infty} x^2\, f(x)dx - a^2.$$

Модой $Mo(X)$ случайной величины X называют ее наиболее вероятное значение (для которого вероятность p_i или плотность вероятности $f(x)$ достигает максимума).

Медианой $Me(X)$ непрерывной случайной величины X называется такое ее значение, для которого:

$P(X < Me(X)) = P(X > Me(X)) = \frac{1}{2}$ (рис. 5).

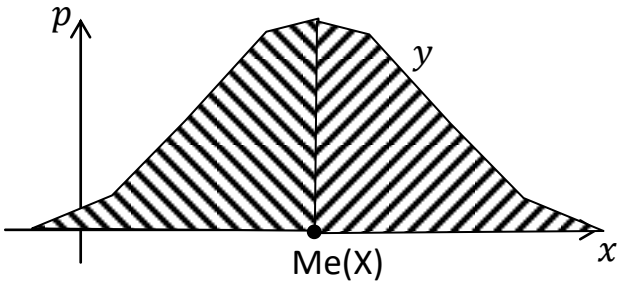


Рис. 5

Пример. Для случайной величины X плотность вероятности определяется равенством $f(x) = 3x^2, x\epsilon[0,1]$. Найти функцию распределения, математическое ожидание, моду и медиану сл. вел. X.

Решение. Вычислим функцию распределения $F(x) = \int_{-\infty}^{x} f(x)dx$ при $x\epsilon[0,1]$:

$$F(x) = \int_{-\infty}^{0} 0\,dx + \int_{0}^{x} 3x^2 dx = x^3\Big|_0^x = x^3.$$

При $x\epsilon[1,\infty)$

$$F(x) = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} 3x^2 dx + \int_{1}^{x} 0\,dx = x^3\Big|_0^1 = 1.$$

Следовательно,

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

Чтобы найти медиану, решим уравнение:

$$F(x) = 1/2 \Rightarrow x^3 = 1/2 \Rightarrow Me(X) = 1/\sqrt[3]{2} \approx 0{,}79.$$

Математическое ожидание:

$$M(X) = \int_{-\infty}^{\infty} x\,f(x)dx = \int_{-\infty}^{0} 0\,dx + \int_{0}^{1} x \cdot 3x^2 dx + \int_{1}^{\infty} 0\,dx = \int_{0}^{1} 3x^3 dx =$$

$$= \frac{3}{4}x^4\Big|_0^1 = 0{,}75.$$

Моду определим из рисунка 6, как точку, в которой функция плотности достигает наибольшего значения.

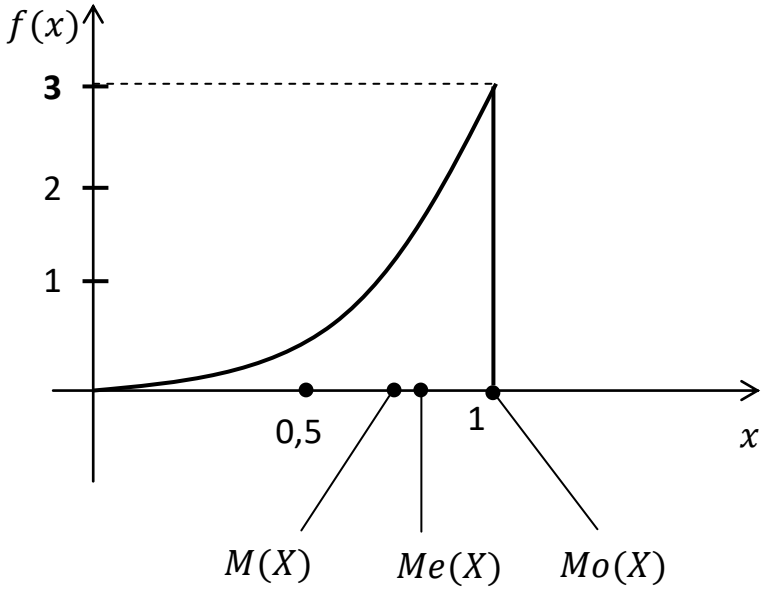


Рис. 6

4.2. Равномерный закон распределения

Непрерывная случайная величина X имеет **равномерный закон** распределения на отрезке $[a, b]$, если ее плотность вероятности $f(x)$ постоянна на этом отрезке и равна нулю вне его, т.е.:

$$f(x) = \begin{cases} \dfrac{1}{b-a} \text{ при } a \le x \le b, \\ 0 \text{ при } x < a, \ \ x > b. \end{cases}$$

Математическое ожидание:

$$M(X) = \frac{a+b}{2}.$$

Дисперсия и среднее квадратическое отклонение:

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

Вероятность попадания равномерно распределенной сл. вел. X в интервал $[\alpha, \ \beta] \subset [a, b]$ равна:

$$P(X \in [\alpha, \beta]) = \frac{\beta - \alpha}{b-a}.$$

Пример. Поезда метро идут регулярно с интервалом 2 мин. Пассажир выходит на платформу в случайный момент времени. Какова вероятность, что

ждать пассажиру придется не более полминуты. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение случайной величины X – времени ожидания поезда.

Решение. Время ожидания поезда равномерно распределено от $a = 0$, до $b = 2$. Если время ожидания не превосходит 0,5 минуты, то $\alpha = 0$ и $\beta = 0{,}5$.

$$P(X \in [\alpha, \beta]) = \frac{0{,}5 - 0}{2 - 0} = \frac{1}{4}.$$

Числовые характеристики:

$$M(X) = \frac{0 + 2}{2} = 1 \text{ мин}, \qquad D(X) = \frac{(2 - 0)^2}{12} = \frac{1}{3}.$$

$$\sigma(X) = \frac{1}{\sqrt{3}} \approx 0{,}58 \text{ мин}.$$

Равномерный закон распределения используется при анализе ошибок округления при проведении числовых расчетов в ряде задач массового обслуживания, при статистическом моделировании наблюдений, подчиненных заданному распределению.

4.3. Нормальный закон распределения случайной величины

Определение. Непрерывная случайная величина X распределена по **нормальному закону с параметрами $\boldsymbol{m}$** и $\boldsymbol{\sigma}$, если ее плотность имеет вид:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}. \qquad (4.1)$$

График $f(x)$ имеет вид (рис. 7):

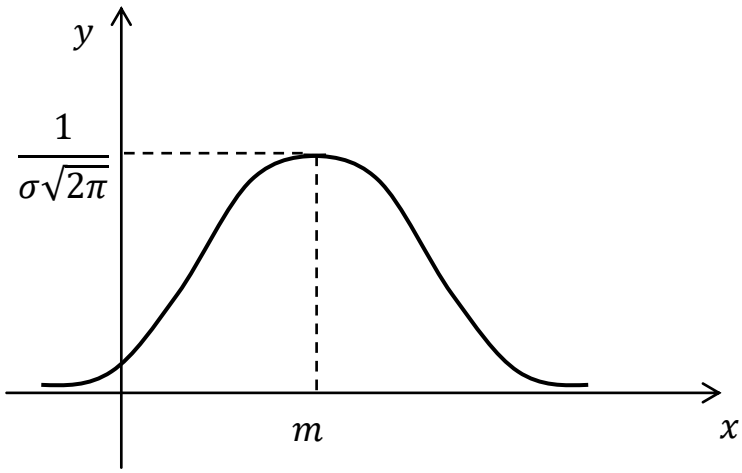


Рис.7

Числовые характеристики: $m_x = m,\ D_x = \sigma^2,\ \ \sigma_x = \sigma$. Вычисление вероятности попадания в интервал:

$$P\{\alpha \le X \le \beta\} = \frac{1}{2}\left[\Phi\left(\frac{\beta - m}{\sigma}\right) - \Phi\left(\frac{\alpha - m}{\sigma}\right)\right], \qquad (4.2)$$

$$P\{|X - m| \le l\} = \Phi\left(\frac{l}{\sigma}\right), \qquad (4.3)$$

где функция $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}}\, dt$ – функция Лапласа, значения которой находятся из таблицы (приложение 1). При этом используется равенство: $\Phi(-x) = -\Phi(x)$.

Функция распределения случайной величины X, распределенной по нормальному закону, выражается через функцию Лапласа $\Phi(x)$ по формуле:

$$F(x) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{x - m}{\sigma}\right). \qquad (4.4)$$

Вычислим по формуле (4.3) вероятности $P\{|X - m| \le \Delta\}$ при различных значениях Δ (используем приложение 1). Получим при:

$$\Delta = \sigma \quad P\{|X - m| \le \sigma\} = \Phi(1) = 0{,}6827;$$

$$\Delta = 2\sigma \quad P\{|X - m| \le 2\sigma\} = \Phi(2) = 0{,}9545;$$

$$\Delta = 3\sigma \quad P\{|X - m| \le 3\sigma\} = \Phi(3) = 0{,}9973.$$

Отсюда следует «правило трех сигм». Отклонение нормально распределенной случайной величины от среднего значения больше, чем на 3 σ (по абсолютной величине), является событием практически невозможным, так как его вероятность весьма мала:

$$P\{|X - m| > 3\sigma\} = 1 - P(|X - m| > 3\sigma) = 1 - 0{,}9973 = 0{,}0027.$$

Пример. Полагая, что рост мужчины определенной возрастной группы есть нормально распределенная случайная величина X с параметрами: $m = 173$ и $\sigma^2 = 36$, найти: а) выражение плотности вероятности и функции распределения случайной величины X; б) доли костюмов 4-го роста (176 –182 см) и 3-го роста (170 – 176 см), которое нужно предусмотреть в общем объеме производства для данной возрастной группы.

Решение. а) По формулам (4.1) и (4.4) запишем:

$$f(x) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{(x-173)^2}{2\cdot 36}}.$$

$$F(x) = \frac{1}{2} + \frac{1}{2}\Phi\left(\frac{x-173}{6}\right).$$

б) Доля костюмов 4-го роста (176–182 см) в общем объеме производства определяется по формуле (4.2) как вероятность:

$$P\{176 \le X \le 182\} = \frac{1}{2}\left[\Phi\left(\frac{182-173}{6}\right) - \Phi\left(\frac{176-173}{6}\right)\right] =$$

$$= \frac{1}{2}[\Phi(1{,}5) - \Phi(0{,}5)] = \frac{1}{2}(0{,}8664 - 0{,}3829) = 0{,}2418.$$

Долю костюмов 3-го роста (170-176 см) можно найти по более простой формуле (4.3), так как этот интервал симметричен относительно $m = 173$. Поэтому неравенство $170 \le X \le 176$ равносильно неравенству: $|X - 173| \le 3$:

$$P(170 \le X \le 176) = P(|X - 173| \le 3) = \Phi(3/6) = \Phi(0{,}5) = 0{,}3829.$$

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

4.1. X – непрерывная случайная величина с плотностью распределения:

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

Требуется: 1) построить кривую распределения; 2) найти функцию распределения; 3) найти m_x, D_x, σ_x; 4) найти $P(1 < X < 2)$.

4.2. Автобусы некоторого маршрута идут строго по расписанию. Интервал движения – 5 мин. Сл. вел. X – время ожидания автобуса. Записать плотность вероятности сл. в. X, найти вероятность того, что пассажир, подошедший к остановке, будет ждать не более трех минут. Найти m_x, D_x, σ_x.

4.3. Цена деления шкалы измерительного прибора равна 0,2. Показания прибора округляют до ближайшего целого деления. Найти вероятность того, что при отсчете будет сделана ошибка: а) меньшая 0,04; б) большая 0,05.

4.4. Математическое ожидание нормально распределенной случайной величины X равно 10, среднее квадратическое отклонение равно 4. Определить вероятность того, что X примет значения, принадлежащие интервалу $(8; 20)$.

4.5. Размер детали подчиняется закону распределения со средней арифметической 15 мм и дисперсией 0,25. Определить ожидаемый процент брака, если допустимые размеры находятся в пределах от 14 мм до 17 мм. Найти выражения плотности вероятности и функции распределения.

4.6. Математическое ожидание нормально распределенной случайной величины равно 15, среднее квадратическое отклонение равно 5. Определить вероятность того, что абсолютная величина отклонения $X - M(X)$ будет меньше 10.

4.7. Случайная величина распределена по нормальному закону с математическим ожиданием: $m_x = 100$ м и средним квадратическим отклонением $\sigma_x = 5$м. Найти интервал, симметричный относительно 100 м, попадание в который имеет вероятность 0,9973.

4.8. Ошибка измерения – нормально распределенная случайная величина с дисперсией, равной 100. Систематическая ошибка отсутствует. Найти вероятность того, что ошибка измерения окажется в интервале (3; 6).

4.9.Текущая цена акции может быть смоделирована с помощью нормального закона распределения с математическим ожиданием 15 ден. ед. и средним квадратическим отклонением 0,2 ден. ед.

1. Найти вероятность того, что цена акции: а) не выше 15,3 ден. ед.; б) не ниже 15,4 ден. ед. ; в) от 14,9 до 15,3 ден. ед.

2. С помощью правила трех сигм найти границы, в которых будет находиться текущая цена акции.

4.10. Коробки с конфетами упаковываются автоматически. Их средняя масса равна 540 г. Известно, что 5% коробок имеет массу, меньшую 500 г. Каков процент коробок, масса которых: а) менее 470 г; б) от 500 до 550 г; в) более 550 г; г) отличается от средней не более, чем на 30 г (по абсолютной величине).

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

4.11. Цена деления шкалы измерительного прибора равна 0,5. Показания прибора округляются до ближайшего деления. Сл. вел. X – ошибка округления. Записать плотность распределения сл. вел. X и найти:

1) вероятность того, что ошибка округления не будет превышать 0,1;

2) вероятность того, что из пяти измерений в двух случаях ошибка округления не будет превышать 0,1.

4.12. X – непрерывная случайная величина с плотностью:

$$f(x) = \begin{cases} x + 1/2, & \text{если } x \in [0, 1]; \\ 0, & \text{если } x < 0, \text{или } x > 1. \end{cases}$$

Требуется: 1) построить кривую распределения; 2) найти $F(x)$, построить график; 3) найти m_x, D_x, σ_x; 4) найти $P(1/2 < X < 1)$.

4.13. Вес вылавливаемых в пруду рыб подчиняется закону нормального распределения с математическим ожиданием, равным 375 г, и средним квадратическим отклонением 25 г. Определить вероятность того, что вес одной пойманной рыбы будет заключен в пределах от 300 до 400 г. Найти выражение функции плотности распределения и функции распределения.

4.14. Ошибка взвешивания – случайная величина, имеющая нормальное распределение. Среднее квадратическое отклонение равно 8. Систематическая ошибка отсутствует. Найти вероятность того, что ошибка измерения попадет в интервал (2; 5).

4.15. Диаметр изготовленных цехом деталей распределяется по нормальному закону с математическим ожиданием, равным 4,5 см, и средним квадратическим отклонением 0,5 мм. Определить вероятность того, что размер диаметра наудачу взятой детали отклонится от математического ожидания не более чем на 1 мм.

5. Показательный закон распределения и закон Пуассона Простейший поток событий. Системы нескольких случайных величин

5.1. Показательный закон распределения

Непрерывная случайная величина X распределена по показательному закону с параметром $\lambda > 0$, если ее плотность вероятности имеет вид:

$$f(x) = \begin{cases} 0, & \text{если } x < 0, \\ \lambda e^{-\lambda x}, & \text{если } \geq 0. \end{cases}$$

Показательное распределение часто встречается в теории массового обслуживания.

Числовые характеристики:

а) математическое ожидание:

$$m_x = \int_0^\infty x\,\lambda e^{-\lambda x}\,dx = \frac{1}{\lambda};$$

б) дисперсия:

$$D_x = \int_0^\infty x^2 \lambda e^{-\lambda x}\,dx - \frac{1}{\lambda^2} = \frac{1}{\lambda^2};$$

в) среднее квадратическое отклонение:

$$\sigma_x = \sqrt{D_x} = \frac{1}{\lambda}.$$

Вероятность попадания случайной величины X в интервал $[a, b]$:

$$P(a \le X \le b) = e^{-\lambda a} - e^{-\lambda b}.$$

Функция распределения:

$$F(x) = \begin{cases} 0, & \text{если } x < 0, \\ 1 - e^{-\lambda x} & \text{если } x \ge 0. \end{cases}$$

Пример. Установлено, что время ремонта телевизора есть случайная величина, распределенная по показательному закону. Определить вероятность того, что на ремонт телевизора потребуется не менее 20 дней, если среднее время ремонта телевизора составляет 15 дней. Найти плотность вероятности, функцию распределения и среднее квадратическое отклонение случайной величины X.

Решение. По условию, математическое ожидание $M(X) = \frac{1}{\lambda} = 15$, откуда параметр $\lambda = 1/15$, следовательно, плотность вероятности и функция распределения имеют вид:

$$f(x) = \frac{1}{15} e^{-\frac{1}{15}x}; \quad F(X) = 1 - e^{-\frac{1}{15}x} \ (x > 0).$$

Найдем искомую вероятность, используя функцию распределения:

$$P(X \ge 20) = 1 - P(X < 20) = 1 - F(20) = 1 - \left(1 - e^{-\frac{20}{15}}\right) = e^{-\frac{20}{15}} =$$
$$= 0{,}264.$$

Осталось найти среднее квадратическое отклонение $\sigma_x = M(X) =$ 15 дней.

Время безотказной работы отдельных элементов есть случайная величина, имеющая показательное распределение. При этом параметр λ является ин-

тенсивностью отказов, т.е. средним числом отказов за единицу времени. Величина $\frac{1}{\lambda}$, обратная интенсивности отказов, называется **наработкой на отказ**. Пусть X – время безотказной работы элемента. Функция надежности $R(x) = e^{-\lambda x}$ определяет вероятность того, что за время x элемент не выйдет из строя.

Пример. Среднее время безотказной работы устройства (наработка на отказ) равна 300 часов. Найти вероятность того, что устройство проработает без отказа 400 часов.

Решение. Параметр $\lambda = \frac{1}{300}$. Надо вычислить функцию надежности

при $x = 400$. Вероятность того, что устройство проработает 400 часов, равна: $R(400) = e^{-\frac{400}{300}} = e^{-1,33}$=0,264.

5.2. Закон распределения Пуассона

Дискретная случайная величина X распределена по закону Пуассона с параметром $a > 0$, если ее возможные значения равны: 0, 1, 2, … , а соответствующие вероятности находятся по формулам:

$$P(X = k) = \frac{a^k}{k!} e^{-a}, \qquad \text{где } a = np.$$

Распределение Пуассона может быть получено из биномиального распределения путем предельного перехода при $n \to \infty$, $p \to 0$ при условии, что $np = a = const$ и в этом случае называется законом «редких явлений». Если n достаточно велико, а p мало, то формулу Пуассона используют в качестве приближения вместо биномиальных формул.

Числовые характеристики:

$$m_x = a, \qquad D_x = a, \qquad \sigma_x = \sqrt{a}.$$

5.3. Простейший поток событий

Рассмотрим последовательность событий, которые появляются в случайные моменты времени. Например, вызовы на станцию скорой помощи, заявки на телефонную станцию, отказы в электронной аппаратуре и т.д. Последовательность таких случайных событий называют **потоком событий**.

Поток событий называют **простейшим**, если он обладает свойствами:

а) **стационарность** – вероятность появления k событий из потока в интервале времени длины τ зависит лишь от длины интервала, а не от его расположения на оси времени;

б) **ординарность** – вероятность появления более одного события за малый промежуток времени пренебрежимо мала по сравнению с вероятностью появления только одного события;

в) **отсутствие последствий** – вероятность появления k событий в интервале времени τ не зависит от числа событий, появившихся в интервалах времени, предшествующих данному.

Интенсивностью потока λ называется среднее число событий из потока в единицу времени. С простейшим потоком событий связаны две случайные величины:

1) X – число событий из потока, попадающих в интервал времени длины τ. Случайная величина X распределена по закону Пуассона с параметром $a = \lambda\tau$.

Следовательно,

$$P(X = k) = \frac{a^k}{k!}e^{-a}.$$

2) T – время между двумя событиями в потоке. Случайная величина T распределена по показательному закону с параметром λ. Следовательно,

$$P(\alpha \leq T \leq \beta) = e^{-\lambda\alpha} - e^{-\lambda\beta}.$$

Пример. Дан простейший поток событий с интенсивностью 20 событий в час. Найти вероятности событий:

A – «за 5 минут произойдут два события»,

B– «за 5 минут произойдет менее двух событий».

C– «за 5 минут произойдет хотя бы одно событие».

D – «время между событиями колеблется от 3 до 5 минут».

Решение. Рассмотрим две случайные величины: X – число событий за $\tau = 5$ минут, T – время между соседними событиями.

Интенсивность потока $\lambda = \frac{20}{60} = \frac{1}{3} = 0{,}33$ событий в минуту.

Случайная величина X распределена по закону Пуассона с параметром $a = \lambda\tau = \frac{5}{3} \approx 1{,}67$.

Отсюда:

$$P(A) = P(X = 2) = \frac{25}{2! \cdot 9} e^{-1,67} = 0,26;$$

$$P(B) = P(X = 0) + P(X = 1) = \frac{(1,67)^0}{0!} e^{-1,67} + \frac{(1,67)^1}{1!} e^{-1,67} = 0,51;$$

$$P(C) = P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{(1,67)^0}{0!} e^{-1,67} = 0,81.$$

Случайная величина T распределена по показательному закону с параметром $\lambda = 0,33$, следовательно,

$$P(D) = P(3 \leq T \leq 5) = e^{-3 \cdot 0,33} - e^{-5 \cdot 0,33} = 0,15.$$

5.4. Системы нескольких случайных величин. Ковариация

В одном испытании могут измеряться две или несколько случайных величин. Например, это может быть длина и ширина детали.

Определение. Совокупность случайных величин $X, Y, Z, \ldots$, связанных одним испытаний, называется **системой случайных величин**.

Рассмотрим случай, когда в системе случайных величин (X, Y) обе случайные величины – дискретные. В этом случае закон распределения системы случайных величин может быть задан с помощью таблицы.

Пример. Пусть случайная величина X может принимать два значения, а случайная величина Y – три значения. Тогда закон распределения системы задается с помощью таблицы. Здесь $p(x_i, y_j)$ – вероятность того, что одновременно выполняются равенства: $\{X = x_i,\ Y = y_j\}$.

X	Y		
	y_1	y_2	y_3
x_1	$p(x_1, y_1)$	$p(x_1, y_2)$	$p(x_1, y_3)$
x_2	$p(x_2, y_1)$	$p(x_2, y_2)$	$p(x_2, y_3)$

Сумма вероятностей $p(x_i, y_j)$ должна равняться единице. Сложив вероятности по столбцам, вычислим: $p(y_j)$, т.е. получим закон распределения случайной величины Y. Сложив по строкам – получим $p(x_i)$, т.е. найдем распределение X.

Условным распределением составляющей X при $Y = y_j$ называется совокупность условных вероятностей $p(x_i / y_j)$, вычисленных в предположении,

что событие $Y = y_j$ уже наступило. Условные вероятности можно вычислить по формулам:

$$p(x_i/y_j) = p(x_i, y_j)/p(y_j) \text{ и } p(y_j/x_i) = p(x_i, y_j)/p(x_i) . \qquad (5.1)$$

Условным математическим ожиданием дискретной величины Y при фиксированном значении $X = x$ называется величина:

$$M(Y/X = x) = \sum_{j=1}^{m} y_j p(y_j/x)$$

и

$$M(X/Y = y) = \sum_{i=1}^{n} x_i p(x_i/y).$$

Условное математическое ожидание $M(Y/x)$ есть функция от x:

$$M(Y/x) = f(x),$$

которую называют **функцией регрессии** Y на X. Аналогично, $M(X/y)$ – функция регрессии X на Y.

Пример. Распределение системы случайных величин задано таблицей:

X	Y	
	2	8
1	0,3	0,2
4	0,2	0,3

Найти: а) законы распределения одномерных случайных величин X и Y; б) условные вероятности $p(x_i/Y = 2)$; в) найти корреляционный момент и коэффициент корреляции.

Решение. а) Складывая вероятности в таблице по строкам, найдем закон распределения X:

$$p(x_1) = P(x = 1) = 0{,}3 + 0{,}2 = 0{,}5;$$
$$p(x_2) = P(x = 4) = 0{,}2 + 0{,}3 = 0{,}5.$$

Складывая вероятности в таблице по столбцам, найдем закон распределения Y:

$$p(y_1) = P(y = 2) = 0{,}3 + 0{,}2 = 0{,}5;$$
$$p(y_2) = P(y = 8) = 0{,}2 + 0{,}3 = 0{,}5.$$

б) Вычислим условные вероятности по формулам (5.1):

$$p(x_1/y_1) = p(x_1, y_1)/p(y_1) \ = \frac{0{,}3}{0{,}5} = 0{,}6 \ ,$$

$$p(x_2/y_1) = p(x_2, y_1)/p(y_1) \ = \frac{0{,}2}{0{,}5} = 0{,}4 \ .$$

г) Используя закон распределения X, найдем числовые характеристики.

X:

x_i	1	4
p_i	0, 5	0, 5

$$M(X) = 1 \cdot 0{,}5 + 4 \cdot 0{,}5 = 2{,}5,$$

$$D(X) = 1^2 \cdot 0{,}5 + 4^2 \cdot 0{,}5 - 2{,}5^2 = 2{,}25,$$

$$\sigma(X) = 1{,}5.$$

Аналогично для Y:

Y:

y_j	2	8
p_j	0, 5	0,5

$$M(Y) = 2 \cdot 0{,}5 + 8 \cdot 0{,}5 = 5,$$

$$D(X) = 2^2 \cdot 0{,}5 + 8^2 \cdot 0{,}5 - 5^2 = 9,$$

$$\sigma(X) = 3.$$

Найдем корреляционный момент по формуле:

$$K_{xy} = M(XY) - M(X) \cdot M(Y).$$

$$M(XY) = 1 \cdot 2 \cdot 0{,}3 + 4 \cdot 2 \cdot 0{,}2 + 1 \cdot 8 \cdot 0{,}2 + 4 \cdot 8 \cdot 0{,}3 = 13{,}4.$$

$$K_{xy} = 13{,}4 - 2{,}5 \cdot 5 = 0{,}9.$$

Тогда коэффициент корреляции:

$$\rho_{xy} = \frac{0{,}9}{1{,}5 \cdot 3} = 0{,}2 \neq 0.$$

Мы получили число близкое, но не равное нулю. Это показывает, что величины X и Y слабо коррелированы, но не являются независимыми.

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

5.1. На станцию скорой помощи поступает поток вызовов, в среднем 20 вызовов в час. Найти вероятность событий:

A – «за пять минут поступят два вызова»,

B – «за пять минут поступит менее двух вызовов»,

C – «время между вызовами колеблется от 2 до 5 минут».

5.2. Поток грузовых железнодорожных составов прибывает на сортировочную горку с интенсивностью 4 состава в час. Найти вероятность событий:

A – «за 30 минут прибудет один состав»,

B – «за 30 минут прибудет хотя бы один состав»,

C – «время между прибытиями составов колеблется от 10 до 20 минут».

5.3. При работе ЭВМ возникают сбои. Среднее число сбоев равно 1,5 сбоев в сутки. Найти вероятность событий:

A – «за 2 суток не будет ни одного сбоя»,

B – «за 2 суток будет не менее двух сбоев»,

C – «время между сбоями колеблется от 12 до 24 часов».

5.4. Дан простейший поток событий с интенсивностью 30 событий в час. Найти вероятности событий.

A – «за 5 минут произойдут три события»,

B – «за 5 минут произойдет хотя бы одно событие»,

C – «время между событиями колеблется от 2 до 5 минут».

5.5. В страховую компанию поступают требования на выплату страховых сумм с интенсивностью $\lambda = 3$ требования в неделю. Найти вероятность того, что на текущей неделе поступят 4 требования.

5.6. Случайная величина X имеет показательное распределение с параметром λ, равным 2. Найти вероятность попадания X в интервал (0,3; 1).

5.7. Среднее время безотказной работы ламп – 50 часов. Найти вероятность того, что лампа не перегорит за 100 часов.

5.8. Закон распределения дискретной двумерной случайной величины (X, Y) задан таблицей:

x_i \ y_j	-1	0	1	2
1	0,10	0,25	0,30	0,15
2	0,10	0,05	0,00	0,05

Найти: а) законы распределения одномерных случайных величин X и Y; б) условные законы распределения случайной величины X при условии: $Y = 2$

и случайной величины Y при условии: $X = 1$; в) вычислить $P(Y < X)$; г) определить ковариацию и коэффициент корреляции случайных величин X и Y.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

5.9. В библиотеку за 10 часов работы обращаются в среднем 120 читателей. Найти вероятность событий:

A – «за 5 минут произойдут три события»,

B – «за 5 минут произойдет хотя бы одно событие»,

C – «время между событиями колеблется от 2 до 5 минут».

5.10. Интенсивность движения трамвая – 12 трамваев в час. Найти вероятность того, что в течение 5 минут не будет трамвая.

5.11. Закон распределения двумерной дискретной случайной величины (X, Y) задан в таблице:

x_i \ y_j	0	1	2	3
-1	0,02	0,03	0,09	0,01
0	0,04	0,16	0,16	0,10
1	0,12	0,10	0,15	0,02

Найти: а) законы распределения одномерных случайных величин X и Y; б) определить ковариацию и коэффициент корреляции этих случайных величин.

6. Неравенства Маркова и Чебышева. Закон больших чисел

6.1.Теорема Маркова и неравенство Чебышева

Теорема Маркова. Если случайная величина $X \geq 0$ имеет математическое ожидание, то для любого положительного числа A верны неравенства:

$$P(x > A) \leq \frac{M(X)}{A} \quad \text{или } P(x \leq A) \geq 1 - \frac{M(X)}{A}.$$

Неравенство Чебышева. Для любой случайной величины X, справедливо неравенство:

$$P(|X - a| > \varepsilon) \leq \frac{D(X)}{\varepsilon^2},$$

где $a = M(X), \varepsilon > 0$.

Для случайной величины $X = m$, имеющей биномиальный закон распределения с математическим ожиданием $a = M(X) = np$ и дисперсией: $D(X) = npq$:

$$P(|m - np| \leq \varepsilon) \geq 1 - \frac{npq}{\varepsilon^2}. \qquad (6.1)$$

Пример 1. Средний расход воды на животноводческой ферме составляет 1000 л в день, а среднее квадратичное отклонение этой случайной величины не превышает 200 л. Оценить вероятность того, что расход воды на ферме в любой выбранный день не превзойдет 2000 л,

используя: а) неравенство Маркова; б) неравенство Чебышева.

Решение. а) Пусть X – расход воды на ферме (л). По условию, $M(X) = 1000$. Используя неравенство Маркова, получим $P(X \leq 2000) \geq 1 - \frac{1000}{2000} = 0{,}5$, т.е. *не менее* чем 0,5.

б) Дисперсия $D(X) = \sigma^2 \leq 200^2$. Так как границы интервала $0 \leq X \leq 2000$ симметричны относительно математического ожидания $M(X) = 1000$, то для оценки вероятности искомого события можно применить неравенство Чебышева:

$$P(X \leq 2000) = P(0 \leq X \leq 2000) =$$
$$= P(|X - 1000| \leq 1000) \geq 1 - \frac{200^2}{1000^2} = 0{,}96,$$

т.е. не менее чем 0,96. В данной задаче оценку вероятности события, найденную с помощью неравенства Маркова ($P \geq 0{,}5$), удалось уточнить с помощью неравенства Чебышева ($P \geq 0{,}96$).

Пример 2. Вероятность выхода с автомата стандартной детали равна 0,96. Оценить с помощью неравенства Чебышева вероятность того, что число бракованных среди 2000 деталей находится в границах от 60 до 100. Уточнить вероятность того же события с помощью интегральной формулы Лапласа (2.1).

Решение. По условию, вероятность того, что деталь бракованная, равна: $p = 1 - 0{,}96 = 0{,}04$. Число бракованных деталей $X = m$ имеет биномиальный закон распределения, а его границы 60 и 100 симметричны относительно математического ожидания: $a = M(X) = np = 2000 \cdot 0{,}04 = 80$.

Следовательно, оценку вероятности искомого события:

$$P(60 \leq m \leq 100) = P(-20 \leq m - 80 \leq 20) = P(|m - 80| \leq 20)$$

можно найти по формуле (6.1):

$$P(|m-80| \le 20) \ge 1 - \frac{2000 \cdot 0{,}04 \cdot 0{,}96}{20^2} = 1 - \frac{76{,}8}{400} = 0{,}808,$$

т.е. *не менее* чем 0,808.

Применяя интегральную формулу Лапласа (2.1), получим:

$$P(|m-80| \le 20) = \Phi\left(\frac{20}{\sqrt{76{,}8}}\right) = \Phi(2{,}28) \approx 0{,}979,$$

т.е. вероятность искомого события приближенно равна 0,979.

Полученный результат $P \approx 0{,}979$ не противоречит оценке, найденной с помощью неравенства Чебышева: $P \ge 0{,}808$.

6.2. Теорема Чебышева. Закон больших чисел

Теорема Чебышева. Если дисперсии попарно независимых случайных величин X_i $i = 1,2,\dots,n$ ограничены сверху постоянным числом C, то:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n} M(X_i)\right| \le \varepsilon\right) > 1 - \frac{C}{n\varepsilon^2}.$$

Закон больших чисел. При неограниченном увеличении числа независимых испытаний над случайной величиной X для любого сколь угодно малого $\varepsilon > 0$ справедлива оценка:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} x_i - M(X)\right| \le \varepsilon\right) > 1 - \frac{C}{n\varepsilon^2}.$$

Пример 3. Сколько надо провести измерений данной величины, чтобы с вероятностью не менее 0,95 гарантировать отклонение средней арифметической этих измерений от истинного значения величины не более, чем на 1 (по абсолютной величине), если среднее квадратическое отклонение каждого из измерений не превосходит 5?

Решение. Пусть X_i – результат i-го измерения $(i = 1,2,\dots,n)$; a – величины, т.е. $M(X_i) = a$ при любом i.

Надо найти n, при котором:

$$P\left(\left|\frac{X_1 + X_2 + \dots + X_n}{n} - a\right| \le 1\right) \ge 0{,}95.$$

В соответствии с законом больших чисел данное неравенство будет выполняться, если :

$$1-\frac{C}{n\varepsilon^2}=1-\frac{5^2}{n\cdot 1^2}\geq 0{,}95\Rightarrow\frac{25}{n}\leq 0{,}05\ \Rightarrow n\geq\frac{25}{0{,}05}=500,$$

т.е. потребуется <u>не менее</u> 500 измерений.

6.3. Теорема Бернулли

Теорема. При неограниченном увеличении числа независимых испытаний с одинаковой вероятностью p появления события A в каждом испытании относительная частота $\frac{m}{n}$ события A сходится по вероятности к p. При этом:

$$P\left(\left|\frac{m}{n}-p\right|\leq\varepsilon\right)>1-\frac{pq}{n\varepsilon^2}.$$

Пример 4. Всхожесть семян характеризуется вероятностью 0,7. Определить, сколько нужно посеять семян, чтобы с вероятностью 0,9 можно было утверждать, что частость проросших семян будет отличаться от вероятности их всхожести не более, чем на 0,03 по абсолютной величине.

Решение. Воспользуемся теоремой Бернулли. По условию, $p=0{,}7$; $\varepsilon=0{,}03$. Вероятность $P\left(\left|\frac{m}{n}-p\right|\leq\varepsilon\right)\geq 0{,}9$. Это будет выполнено, если $1-\frac{pq}{n\varepsilon^2}\geq 0{,}9$ или $1-\frac{0{,}7\cdot 0{,}3}{n\cdot 0{,}03^2}\geq 0{,}9\Leftrightarrow 0{,}1\geq\frac{0{,}21}{n\cdot 9\cdot 10^{-4}}\Leftrightarrow n\geq\frac{2{,}1}{9}10^4=2333{,}3.$

Так как n – целое, то получаем условие: $n\geq 2334$.

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

6.1. Дан закон распределения случайной величины X. С помощью неравенства Маркова оценить вероятность того, что случайная величина X примет значение, не большее 6.

X	2	4	6	8	10
P	0,2	0,4	0,2	0,1	0,1

6.2. Средний расход воды в рабочем поселке составляет 250 м^3. Оценить вероятность того, что в поселке в ближайшие сутки расход воды не превысит 350 м^3.

6.3. Оценить вероятность того, что отклонение любой случайной величины от ее математического ожидания будет не более трех средних квадратичных отклонений (по абсолютной величине) – правило трех сигм для произвольного распределения.

6.4. Средняя длина детали составляет 10 см, а дисперсия равна 0,03. С помощью неравенства Чебышева оценить вероятность того, что изготовленная деталь окажется по длине не менее 9,75 см и не более 10,25 см.

6.5. Считая равновероятным рождение девочки и мальчика, оценить с помощью неравенства Чебышева вероятность того, из 4000 родившихся детей мальчиков будет от 1950 до 2050 (включительно).

6.6. 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.

6.7. Используя теорему Бернулли, определить число подбрасываний монеты, достаточное, чтобы с вероятностью, большей чем 0,95, можно было утверждать, что относительная частота выпадения герба будет заключена в интервале от 0,49 до 0,51?

6.8. Для определения среднего размера было проверено 2000 деталей. Известно, что дисперсия не превышает 9. С помощью закона больших чисел оценить вероятность того, что средний размер детали в проверенной партии отклонится от его математического ожидания не более, чем на 0,3.

6.9. Для нахождения среднего процента успеваемости проверена успеваемость 900 студентов. Дисперсия по успеваемости не превышает 6. Оценить вероятность того, что средний процент успеваемости по группе проверенных студентов будет отличаться от успеваемости всех студентов не более, чем на 0,5 процента.

6.10. При выпуске воротников процент возврата составляет 4,6, т.е. вероятность возврата: $p = 0{,}046$. Оценить вероятность того, что при проверке 10000 воротников процент возврата будет отличаться от 4,6 по абсолютной величине не более, чем на 0,5 процента.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

6.11. Среднее изменение курса акции компании в течение одних биржевых торгов составляет 0,3%. С помощью неравенства Маркова оценить вероятность того, что на ближайших торгах курс изменится более, чем на 3%?

6.12. Отделение банка обслуживает в среднем 100 клиентов в день. Оценить вероятность того, что сегодня в отделении банка будет обслужено: а) не более 200 клиентов; б) более 150 клиентов.

6.13. Электростанция обслуживает сеть в 20 000 ламп. Вероятность включения каждой лампы в зимнее время равна 0,8. Оценить с помощью теоремы Бернулли вероятность того, что число ламп, включенных в сеть в зимнее время, будет отличаться от своего математического ожидания по абсолютной величине не более чем на 150.

6.14. Вероятность сдачи в срок всех экзаменов студентом университета равна 0,7. С помощью неравенства Чебышева оценить вероятность того, что доля сдавших все экзамены из 2000 студентов заключена в границах от 0,66 до 0,74.

Глава III. Математическая статистика

7. Статистические методы обработки экспериментальных данных

7.1. Случайная выборка из генеральной совокупности

Пусть требуется изучить совокупность однородных объектов относительно некоторого признака, характеризующего эти объекты. Например, рассматриваются размеры детали и величина отклонения размеров от заданного значения. Иногда проводят сплошное обследование, т.е. исследуется каждый объект. Однако обычно исследуется только небольшая часть всех объектов.

Выборочной совокупностью, или просто **выборкой**, называется совокупность случайно отобранных объектов.

Генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Значение интересующего нас признака – случайная величина X. Случайная величина X распределена по некоторому закону с неизвестными параметрами, который называется **распределением генеральной совокупности**.

Проведем n испытаний при одних и тех же условиях. Случайная величина X принимает значения: $x_1, x_2, \dots, x_n$.

Множество значений $x_1, x_2, \dots, x_n$ называется **случайной выборкой объема n.**

Пусть выборка объема n содержит m различных чисел. Изменив нумерацию, запишем их в виде: $x_1, x_2, \dots, x_m$, причем $x_1 < x_2 < \dots < x_m$. Число m называется размахом выборки.

Пусть значение x_i встречается в выборке n_i раз, $i = 1,2, \dots, m$.

Число n_i называется **абсолютной частотой**, а число $w_i = \frac{n_i}{n}$ – относительной частотой или **частостью** элемента x_i.

При большом объеме выборки ее элементы объединяются в группы. Интервал, содержащий все элементы выборки, разбивается на k непересекающихся интервалов $(a_0, a_1), (a_1, a_2), \dots, (a_{k-1}, a_k)$, не обязательно равных по длине.

Пусть n_i – число элементов выборки, попавших в интервал (a_{i-1}, a_i), а $w_i = \frac{n_i}{n}$ – частость (относительная частота).

Вариационным рядом называется ранжированный (т.е. построенный в порядке убывания или убывания) ряд вариантов с соответствующими частотами или частостями.

Пример 1. Сгруппированный вариационный ряд:

	Баллы по ЕГЭ	Частота	Частость	Накопленная частота	Накопленная частость
1	27-40	4	0,182	4	0,182
2	40-60	8	0,364	12	0,545
3	60-100	10	0,454	22	1

Здесь изображены баллы по ЕГЭ в одной из групп первокурсников. Предполагается, что $n = 22$.

Пример 2. Дискретный вариационный ряд. Дано распределение 50 рабочих цеха по разрядам:

Тарифный разряд	1	2	3	4	5	6	Σ
Частота n_i	2	3	6	8	22	9	50

Для графического изображения дискретного вариационного ряда построим полигон частот (рис. 8) и кумуляту (рис. 9)

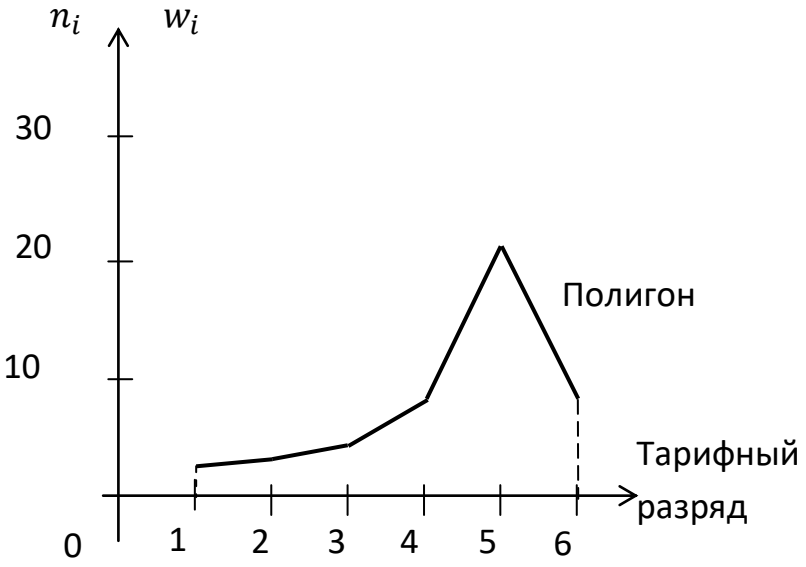


Рис.8

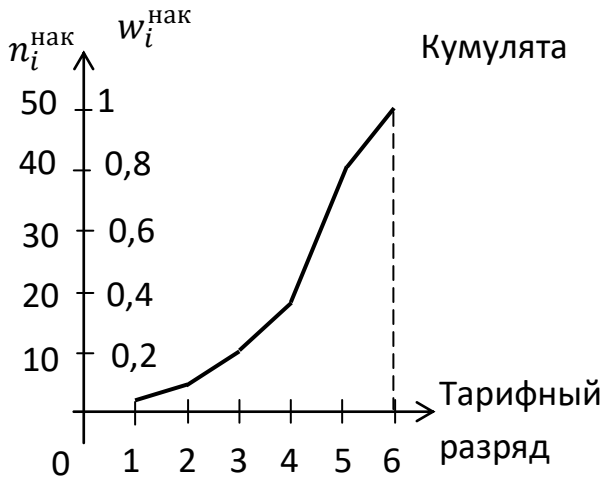


Рис.9

Для изображения интервального вариационного ряда используются гистограммы. Гистограмма относительных частот – это ступенчатая фигура, состоящая из прямоугольников с основаниями (a_{i-1}, a_i) и высотами $f_i = \frac{w_i}{a_i - a_{i-1}}, i = 1,2, \dots, k$.

Пример 3. Построить гистограмму, соответствующую данным примера 1.

Решение. Построим таблицу:

X	(27;40)	(40;60)	(60;100)
$a_i - a_{i-1}$	13	20	40
w_i	0,182	0,364	0,454
f_i	0,014	0,0182	0,01135

Теперь построим гистограмму (рис.10):

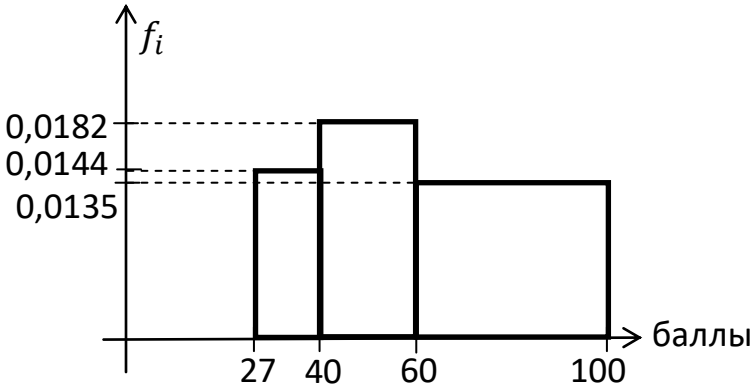


Рис.10

7.2. Точечные оценки параметров распределения

Пусть $x_1, x_2, \dots, x_n$ – выборка наблюдений случайной величины X.

Рассматривают выборочную среднюю:

$$m_B = \bar{x} = \frac{\sum_{i=1}^{m} x_i n_i}{n} = \sum_{i=1}^{n} x_i w_i$$

и выборочную дисперсию:

$$D_B = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}, \quad \sigma_x = \sqrt{D_B}.$$

Выборочную дисперсию можно вычислить по формуле:

$$D_B = \bar{x}^2 - [\bar{x}]^2.$$

В случае, когда значения $x_1, x_2, \dots, x_k$ имеют соответственные частоты $n_1, n_2, \dots, n_k$, то:

$$D_B = \left(\sum_{i=1}^{k} n_i (x_i - \bar{x})^2 \right) \Big/ n.$$

В случае группированного вариационного ряда формулы для нахождения точечных оценок принимают вид:

$$m_B = \bar{x} = \frac{\sum_{i=1}^{m} Z_i n_i}{n}, \qquad D_B = \left(\sum_{i=1}^{k} n_i (Z_i - \bar{x})^2 \right) \Big/ n,$$

где Z_i – середина каждого интервала (a_{n-1}, a_n), $i = 1, 2, \dots, k$.

Пример 4. Найти дисперсию по заданному распределению:

x_i	1	2	3	4
n_i	20	15	10	5

Решение. Найдем общую среднюю:

$$\bar{x} = \frac{20 \cdot 1 + 15 \cdot 2 + 10 \cdot 3 + 5 \cdot 4}{20 + 15 + 10 + 5} = \frac{100}{50} = 2.$$

Найдем среднюю квадратов значений признака:

$$\bar{x}^2 = \frac{20 \cdot 1^2 + 15 \cdot 2^2 + 10 \cdot 3^2 + 5 \cdot 4^2}{50} = 5.$$

Тогда дисперсия:

$$D_B = \bar{x}^2 - [\bar{x}]^2 = 5 - 2^2 = 1.$$

Используя выборочную среднюю, можно оценить математическое ожидание для генеральной совокупности. Для этого положим: $M(X) \approx \bar{x}$.

Если в качестве оценки генеральной дисперсии использовать выборочную дисперсию, то эта оценка будет приводить к систематическим ошибкам, давая заниженное значение генеральной дисперсии. Поэтому для оценки генеральной дисперсии используют исправленную выборочную дисперсию, которую обозначают:

$$s^2 = \frac{n}{n-1} D_B = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$$

Для оценки среднего квадратического отклонения генеральной совокупности используют величину:

$$s = \sqrt{\frac{n}{n-1} D_B}.$$

Пример 5. Найти исправленные выборочную дисперсию и среднее квадратическое отклонение, используя данные примера 4.

Решение. Исправленная дисперсия: $s^2 = \frac{n}{n-1} D_B = \frac{50}{49} \cdot 1 \approx 1{,}02$; исправленное среднее квадратическое отклонение: $s = \sqrt{1{,}020408} \approx 1{,}01$.

Свойства выборочной дисперсии аналогичны свойствам дисперсии случайной величины. Однако есть и отличия. Отметим такое свойство.

Если вариационный ряд состоит из нескольких групп наблюдений, то общая дисперсия равна сумме средней арифметической групповых дисперсий и межгрупповой дисперсии:

$$D_B = \bar{D}_\iota + \delta^2,$$

где D_B – общая дисперсия (дисперсия всего ряда);

$$\bar{D}_\iota = \frac{\sum_{i=1}^{\ell} D_i n_i}{n} \tag{7.1}$$

– средняя арифметическая групповых дисперсий;

$$D_i = \frac{\sum_{j=1}^{m}\left(x_j - \bar{x}_i\right)^2 n_j}{n_i};$$

$$\delta^2 = \frac{\sum_{i=1}^{\ell}(\bar{x}_i - \bar{x})^2 n_i}{n} \tag{7.2}$$

– межгрупповая дисперсия.

Пример 6. Имеются следующие данные о средних и дисперсиях заработной платы двух групп рабочих.

Таблица 7.1

Группа рабочих	Число рабочих	Средняя заработная плата одного рабочего в группе (тыс. руб)	Дисперсия заработной платы
Работающие на одном станке	40	24	18
Работающие на двух станках	60	32	20

Найти общую дисперсию распределения рабочих по заработной плате.

Решение. Найдем общую среднюю по формуле:

$$\bar{x} = \frac{\sum_{i=1}^{\ell} \bar{x}_i n_i}{n} = \frac{24 \cdot 40 + 32 \cdot 60}{100} = 28{,}8(\text{тыс. руб}).$$

Найдем среднюю групповых дисперсий по формуле (7.1):

$$\overline{D}_i = \frac{18 \cdot 40 + 20 \cdot 60}{100} = 19{,}2.$$

Найдем межгрупповую дисперсию по формуле (7.2):

$$\delta^2 = \frac{(24 - 28{,}8)^2 \cdot 40 + (32 - 28{,}8)^2 \cdot 60}{100} = 15{,}36.$$

Используя правило сложения дисперсий, найдем общую дисперсию заработной платы и ее среднее квадратическое отклонение.

$$D = 19{,}2 + 15{,}36 = 34{,}56; \quad \sigma = \sqrt{34{,}56} = 5{,}88(\text{тыс. руб}).$$

7.3. Доверительные интервалы для оценки математического ожидания нормального распределения при известном σ

Пусть количественный признак X генеральной совокупности распределен нормально, среднее квадратическое отклонение σ известно, но неизвестно математическое ожидание.

Доверительным интервалом для параметра $M(X)$ называется интервал $(\bar{x}-\varepsilon, \bar{x}+\varepsilon)$, содержащий истинное значение $M(X)$ с заданной доверительной вероятностью (надежностью) γ.

Доверительный интервал J_ε для математического ожидания $M(X)$ с заданной доверительной вероятностью γ находится по формуле: $J_\varepsilon = (\bar{x}-\varepsilon, \bar{x}+\varepsilon)$, где $\varepsilon = t_\gamma \cdot \sigma[\bar{x}]$. Значение находится из условия: $\Phi(t_\gamma) = \gamma$, $\Phi(x)$ – функция Лапласа (приложение 1), а $\sigma[\bar{x}] = \sqrt{\frac{s^2}{n}}$.

Пример 7. Используя данные примеров 4 и 5, найти доверительный интервал для математического ожидания $M(X)$ при доверительной вероятности $\gamma = 0{,}85$.

Решение. По данным, полученным при решении в примерах 4 и 5, $\bar{x} = 2, s^2 = 1{,}0204,\ s = 1{,}0102$. Для $\sigma[\bar{x}]$ получим оценку:

$$\sigma[\bar{x}] = \sqrt{\frac{1{,}0204}{50}} = 0{,}1429.$$

Чтобы найти t_γ, решим с помощью таблицы функции Лапласа уравнение: $\Phi(t_\gamma) = 0{,}85$. Для этого найдем в таблице (приложение 1) значение функции, наиболее близкое к числу 0,85. Это будет число 0,8501, которому соответствует значение аргумента $t_\gamma = 1{,}44$.

Следовательно, $\varepsilon = t_\gamma \cdot \sigma[\bar{x}] = 1{,}44 \cdot 0{,}1429 = 0{,}2057$. Доверительный интервал имеет вид:

$$J_\varepsilon = (2 - 0{,}2057;\ 2 + 0{,}2057) = (1{,}7943;\ \ 2{,}2057).$$

7.4. Доверительные интервалы для оценки математического ожидания нормального распределения при неизвестном σ

Пусть количественный признак X генеральной совокупности распределен нормально, но математическое ожидание и среднее квадратическое отклонение σ неизвестно.

Доверительный интервал в этом случае имеет вид: $\left(\bar{x} - \frac{t_\gamma s}{\sqrt{n-1}}, \bar{x} + \frac{t_\gamma s}{\sqrt{n-1}}\right)$. Здесь s – исправленная оценка среднего квадратического отклонения, γ – надежность оценки, величина $t_\gamma = t(\gamma, n)$ находится по таблице распределения Стьюдента (приложение 4).

Пример 7. Найти доверительный интервал, содержащий истинное значение $M(X)$, с надежностью 0,99, для выборки в примере 4.

Решение. В примерах 4 и 5 было показано, что $\bar{x} = 2$, а $s = 1{,}01$ и $n = 50$. Чтобы использовать таблицу Стьюдента, нужно указать надежность γ и число степеней свободы $k = n - 1$. Надежность может принимать значения: $\gamma = 0{,}95; 0{,}99; 0{,}999$. Заходим в таблицу при $\gamma = 0{,}99$, $k = 50 - 1 = 49$. Получаем значение $t_\gamma = 2{,}68$. Вычисляя границы интервала, получаем, что истинное значение $M(X)$ с надежностью 0,99 принадлежит интервалу $(1{,}617; 2{,}383)$.

7.5. Доверительный интервал для генеральной дисперсии и среднего квадратического отклонения

Пусть известно, что случайная величина X распределена нормально.

Требуется оценить неизвестные: генеральную дисперсию $D_\text{г}$ и среднеквадратическое отклонение $\sigma_\text{Г}$ по известному выборочному $\sigma_\text{В}$.

Найдем доверительный интервал для генеральной дисперсии с заданной надежностью γ. Для этого используется таблица распределения χ^2. Доверительный интервал для генеральной дисперсии имеет вид:

$$J_x = \left(\frac{nD_B}{\chi_2^2}; \frac{nD_B}{\chi_1^2}\right),$$

а для среднего квадратического отклонения:

$$\left(\frac{\sqrt{nD_B}}{\chi_2}; \frac{\sqrt{nD_B}}{\chi_1}\right).$$

Чтобы найти числа χ_2 и χ_1 , надо воспользоваться таблицей $\chi^2_{\alpha,k}$.

Для входа в таблицу надо вычислить: $\alpha_1 = \frac{1+\gamma}{2}$ и $\alpha_2 = \frac{1-\gamma}{2}$; тогда $\chi_1^2 == \chi^2_{\alpha_1,k}$ и $\chi_1^2 = \chi^2_{\alpha_2,k}$, $k = n - 1$.

Пример 8. На основании выборочных наблюдений производительности труда 20 работниц было установлено, что среднее квадратическое отклонение суточной выработки составляет 15 м ткани в час. Предполагая, что производительность труда работницы имеет нормальное распределение, найти границы, в которых с надежностью 0,9 заключены генеральная дисперсия и среднее квадратическое отклонение суточной выработки рабочих.

Решение. Имеем: $\gamma = 0{,}9; \alpha_1 = \frac{1+\gamma}{2} = 0{,}95$ и $\alpha_2 = \frac{1-\gamma}{2} = 0{,}05$. При числе степеней свободы $k = n - 1 = 19$ определим по таблице приложения 6 значения: $\chi^2_{\alpha_1,k}$ =10,12 и $\chi^2_{\alpha_2,k}$ =30,14. Тогда доверительный интервал для генеральной дисперсии:

$$\frac{20}{30{,}14} \cdot 15^2 < \sigma^2 < \frac{20}{10{,}12} \cdot 15^2$$

или $149{,}29 < \sigma^2 < 444{,}80$ и для σ доверительный интервал:

$\sqrt{149{,}29} < \sigma < \sqrt{444{,}80}$ или $12{,}22 < \sigma < 21{,}09$.

Таблица приложения 6 составлена при числе свободы k от 1 до 30. При $k > 30$ величины χ_1^2 и χ_1^2 вычисляются по формулам:

$$\chi_1^2 = \frac{1}{2}\left(\sqrt{2k-1} - t\right)^2, \quad \chi_1^2 = \frac{1}{2}\left(\sqrt{2k-1} + t\right)^2,$$

где t находим, решая уравнение $\Phi(t) = \gamma$. Здесь $\Phi(t)$ – функция Лапласа (приложение 1) .

Пример 9. Решить задачу в примере 8 при $n = 100$ работницам.

Решение. При $\gamma = 0{,}9$ находим t, решая уравнение $\Phi(t) = 0{,}9$ с помощью таблицы функции Лапласа (приложение 1). Получаем: $t = 1{,}645$, поэтому:

$$\chi_1^2 = \frac{1}{2}\left(\sqrt{2 \cdot 99 - 1} - 1{,}645\right)^2 = 76{,}8\,,$$

$$\chi_2^2 = \frac{1}{2}\left(\sqrt{2 \cdot 99 - 1} + 1{,}645\right)^2 = 122{,}9\,.$$

Доверительный интервал для генеральной дисперсии :

$$\left(\frac{nD_B}{\chi_2^2}; \frac{nD_B}{\chi_1^2}\right) = \left(\frac{100 \cdot 225}{122{,}9}; \frac{100 \cdot 225}{76{,}8}\right) = (183{,}1;\ 293{,}0).$$

Для среднего квадратического отклонения:

$$13{,}5 < \sigma_г < 17{,}1 \text{ (м/ч)}.$$

7.6. Проверка гипотезы о нормальном законе распределения χ^2 – критерий Пирсона

Пусть задан интервальный вариационный ряд, описывающий случайную величину X. По этим данным можно оценить числовые характеристики $a = m_x$ и σ_x величины С. Проверим гипотезу о нормальном распределении величины X. Предположим, что это так. Тогда по формуле:

$$P(x_i \le X \le x_{i+1}) = \frac{1}{2}\left[\Phi\left(\frac{x_{i+1} - a}{\sigma}\right) - \Phi\left(\frac{x_i - a}{\sigma}\right)\right]$$

можно найти p_i – вероятности попадания в интервал $(x_i; x_{i+1})$ и теоретические частоты np_i попадания в каждый интервал. Затем производится сравнение теоретических частот np_i и эмпирических частот n_i. Для этого вычисляется сумма:

$$\chi^2 = \sum_i \frac{(n_i - np_i)^2}{np_i}.$$

Сравниваем величины χ^2 и $\chi^2_{\text{крит}}$. Значения $\chi^2_{\text{крит}}$ находим по таблице критических точек распределения χ^2(приложение 6). Если $\chi^2 < \chi^2_{\text{крит}}$, то гипотеза о нормальном распределении случайной величины согласуется с опытными данными. В противном случае гипотеза отвергается.

Чтобы использовать таблицу $\chi^2_{\text{крит}}$, необходимо указать значения $k = m - 3$ и уровень значимости α.

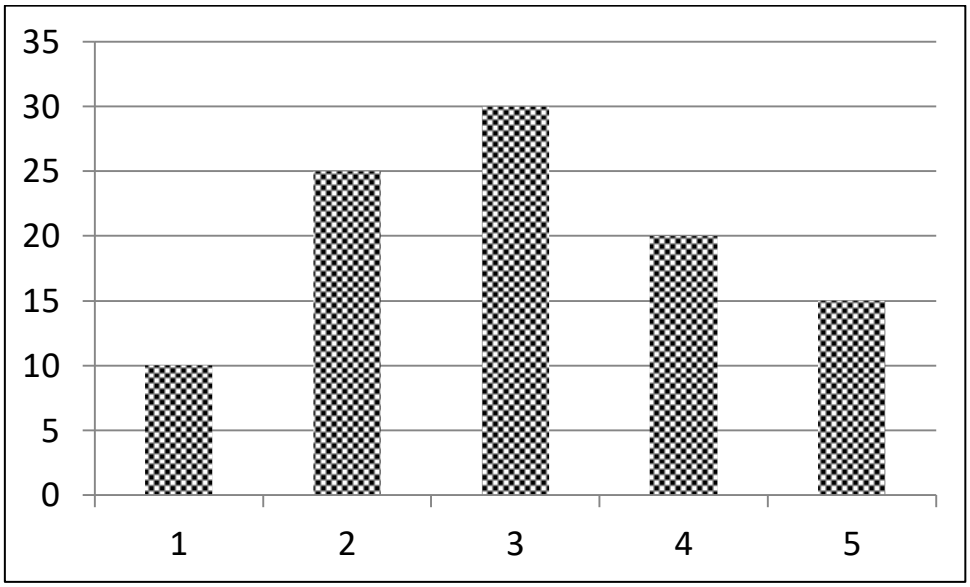
Уровень значимости – это вероятность того, что гипотеза окажется верна, хотя условие $\chi^2 < \chi^2_{\text{крит}}$ не выполняется. Иными словами α – это вероятность того, что на основе данного критерия мы зря отбросим правильную гипотезу.

Пример 10. Дан интервальный вариационный ряд:

i	*Интервалы*	*Частоты*
1	*95 – 105*	*10*
2	*105 – 115*	*25*
3	*115 – 125*	*30*
4	*125 – 135*	*20*
5	*135 – 145*	*15*

Проверить гипотезу о нормальном распределении признака X на уровне значимости $\alpha = 0{,}05$.

Решение. Чтобы выяснить, насколько эта гипотеза правдоподобна, с помощью программы Excel построим гистограмму.



По вертикали здесь отложены частоты, а по горизонтали номер интервала. Полученное изображение примерно соответствует нормальному распределению. Проверим гипотезу с помощью критерия Пирсона.

По формуле: $\bar{x} = \frac{\sum_{i=1}^{m} Z_i n_i}{n}$, где Z_i – середины интервалов, найдем: $\bar{x} = 120{,}5$; $D_B = \bar{x}^2 - [\bar{x}]^2 = 144{,}75$ и $\sigma_{\text{В}} \approx 12{,}03$.

Вычислим вероятности $p_i = P(x_i \leq X \leq x_{i+1})$, теоретические частоты np_i. Сравниваем теоретические и эмпирические частоты. Для этого вычисляем χ^2.

i	Интервалы $[x_i, x_{i+1}]$	Эмпирические частоты n_i	Вероятности p_i	Теоретические частоты np_i	$(n_i - np_i)^2$	$\frac{(n_i - np_i)^2}{np_i}$
1	95 – 105	10	0,082	8,179	3,315	0,405
2	105 – 115	25	0,225	22,497	6,267	0,279
3	115 – 125	30	0,322	32,202	4,851	0,151
4	125 –135	20	0,240	24,013	16,103	0,671
5	135 – 145	15	0,093	9,321	32,255	3,461
	Σ	100	0,962	96,21	–	$\chi^2 = 4{,}97$

Находим $\chi^2_{\text{крит}}$ по таблице «Значения $\chi^2_{\alpha k}$ критерия Пирсона» (приложение 6). Входим в таблицу при $k = m - 3 = 5 - 3 = 2$ и $\alpha = 0{,}05$.

Получается: $\chi^2_{\text{крит}} = 5{,}99 > \chi^2 = 4{,}97$. Следовательно, гипотеза о нормальном распределении случайной величины X с параметрами $a = 120{,}5$; $\sigma = 12{,}03$ не противоречит опытным данным.

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

7.1. Было проделано 15 опытов, в которых измерялась величина X. Получены результаты: 9,12; 9,07; 9,07; 9,10; 9,12; 9,19; 9,12; 9,12; 9,24; 9,19; 9,24; 9,16; 9,19; 9,10; 9,16.

Требуется:

1) составить вариационный ряд, построить полигон относительных частот;

2) найти точечные оценки $M(X), D(X), \sigma_X$;

3) найти доверительный интервал для $M(X)$ с заданной доверительной вероятностью 0,95.

7.2. На предприятие поступила партия из 40 деталей. Положительные отклонения размера (в мм) равны:
17, 21, 9, 20, 23, 18, 22, 20, 17, 12, 20, 11, 9, 19, 20, 9, 19, 17, 21, 13, 17, 22, 22, 10, 20, 20, 15, 19, 20, 20, 13, 21, 21, 9, 14, 11, 19,18, 23, 19.

Требуется:

1) составить группированный вариационный ряд, разбив промежуток $(8{,}5;\ 23{,}5)$ на пять интервалов одинаковой длины, построить гистограмму;

2) найти точечные оценки $M(X), D(X), \sigma_X$.

3) найти доверительный интервал для $M(X)$ с заданной доверительной вероятностью 0,95.

В задачах 7.3–7.4 дано распределение признака X (случайной величины X), полученной по n наблюдениям. Необходимо: 1) построить полигон (гистограмму) и кумуляту X; 2) найти: а) выборочную среднюю $\bar{x}$; б) дисперсию D_B и среднее квадратическое отклонение σ_B; в) генеральную дисперсию s^2 и среднее квадратическое отклонение s.

7.3. X – число сделок на фондовой бирже за квартал: $n = 60$ инвесторов. Найти доверительный интервал для $M(X)$ с надежностью 0,99.

x_i	0	1	2	3	4
n_i	20	17	10	8	5

7.4. X – месячный доход жителя региона (в тыс. руб.); выборка $n = 1000$.

x_i	Менее 12	12 - 24	24 - 36	36 - 48	48 - 60	Свыше 60
n_i	58	96	239	328	147	132

7.5. X – удой коров на молочной ферме за лактационный период (в ц); $n = 100$ (коров)

x_i	4 - 8	8 - 12	12 - 16	16 - 20	20 -24	24 -28
n_i	4	17	35	26	16	2

7.6. В таблице приведено распределение 50 рабочих по производительности труда X (единиц за смену), разделенных на две группы: 30 и 20 человек.

	Прошедшие техническое обучение (группа 1)					Не прошедшие техническое обучение (группа 2)				
x_i	85	94	96	102	103	63	69	83	89	106
n_i	2	5	11	8	4	2	6	8	3	1

Вычислить общие и групповые средние и дисперсии и убедиться в справедливости правила сложения дисперсий.

7.7. Произведено 100 измерений случайной величины X. Получена оценка: $\bar{x} = 24$. Можно считать достоверным, что $\sigma(X) = 4$. Найти доверительный интервал для $M(X)$ с надежностью 0,8.

7.8. Дан интервальный вариационный ряд:

i	Интервалы	Частоты
1	15 – 25	8
2	25 – 35	12
3	35 – 45	18
4	45 – 55	10
5	55 – 65	2

Проверить гипотезу о нормальном распределении признака X на уровне значимости $\alpha = 0{,}1$.

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

7.9. Для выборки объема $n = 9$ найдена выборочная дисперсия $D_B = 72$. Найти исправленную дисперсию s^2.

7.10. Было проделано 15 опытов, в которых измерялась величина X. Получены результаты:

3,6 2,1 2,1 3 3,6 5,7 3,6 3,6
7,2 5,7 7,2 4,8 5,7 3 4,8

Требуется:

1) составить вариационный ряд, построить полигон относительных частот;

2) найти точечные оценки $M(X), D(X), \sigma_X,\ \ s^2, s$;

3) найти доверительный интервал для $M(X)$ с заданной доверительной вероятностью 0,95.

8. Корреляционный метод

8.1. Уравнение регрессии

Если с ростом факторного признака X средние значения результирующего признака Y изменяются равномерно, то уравнение регрессии будет линейным:

$$\overline{y(x)} = ax + b.$$

Система линейных уравнений для определения параметров a и b методом наименьших квадратов имеет вид:

$$\begin{cases} \bar{x}a + b = \bar{y} \\ \overline{x^2}a + \bar{x}b = \overline{xy} \end{cases},$$

где

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \qquad \overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2,$$

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i.$$

Для линейного уравнения регрессии можно использовать в качестве показателя тесноты связи число:

$$r_b = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x \cdot \sigma_y},$$

которое называется выборочным линейным коэффициентом корреляции. Здесь $\sigma_x = \sqrt{\overline{x^2} - (\overline{x})^2}$ – среднее квадратическое отклонение признака X, $\sigma_y = \sqrt{\overline{y^2} - (\overline{y})^2}$– среднее квадратическое отклонение признака Y.

Пример 1. Методами корреляционно-регрессивного анализа изучить влияние факторного признака X на результативный признак Y:

1) определить уравнение регрессии;

2) вычислить выборочные показатели тесноты связи.

Зависимость потребления некоторого продукта Y от среднегодового дохода X:

X	1	2	3	4	5	6
Y	2,6	5,7	9,3	11,9	15,1	17,9

Решение. Нарисуем точки, соответствующие значениям X и Y (рис. 11). Как видим, они приблизительно лежат на одной прямой. Можно предположить, что уравнение регрессии будет линейным. Т.е. $\overline{y(x)} = ax + b$. Система нормальных уравнений для определения параметров методом наименьших квадратов имеет вид:

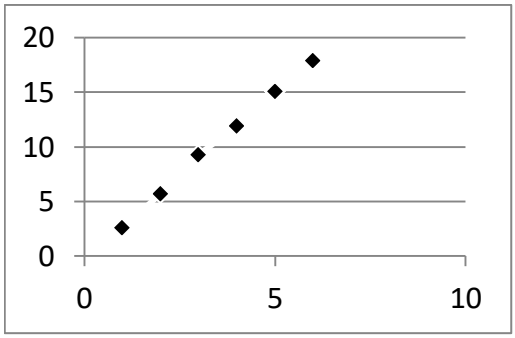


Рис.11

$$\begin{cases} \bar{x}a + b = \bar{y} \\ \overline{x^2}a + \bar{x}b = \overline{xy} \end{cases},$$

где $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$

$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2, \quad \overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i.$

Вычисляя эти параметры, получаем значения: $\bar{x} = 3{,}5$; $\bar{y} = 10{,}4167$,

$\overline{x^2}$ =15,1667; $\overline{xy}$=45,4 $\overline{y^2}$ =135,9617 . Тогда:

$$\begin{cases} 3{,}5a + b = 10{,}4167\ , \\ 15{,}1667a + 3{,}5b = 45{,}4\ , \end{cases}$$

$$\Delta = -2{,}9167;\ \ \Delta a = -8{,}9417;\ \ \Delta b = 0{,}9139 \Rightarrow \begin{cases} a = 3{,}0657 \\ b = -0{,}3133 \end{cases}.$$

Уравнение линейного тренда (линии регрессии):

$y = 3{,}0657\text{x} - 0{,}3133.$

Для оценки тесноты связи вычислим выборочный линейный коэффициент корреляции:

$$r_b = \frac{\overline{xy} - \bar{x}\cdot\bar{y}}{\sigma_x\cdot\sigma_y},$$

где $\sigma_x = \sqrt{\overline{x^2} - (\bar{x})^2}$ и $\sigma_y = \sqrt{\overline{y^2} - (\bar{y})^2}$. Вычислим сигмы: $\sigma_x =$

$$\sqrt{15{,}1667 - (3{,}5)^2} = 1{,}7078;\ \ \sigma_y = \sqrt{135{,}9617 - (10{,}4167)^2} = 5{,}2397.$$

Тогда:

$$r_b = \frac{45{,}4 - 3{,}5\cdot 10{,}4167}{1{,}7078\cdot 5{,}2397} = 0{,}9992.$$

Близость результата к единице говорит о тесной связи между переменными.

8.2. Линейная парная регрессия

Данные о статистической зависимости удобно задавать в виде корреляционной таблицы.

Пример 2. Распределение 50 заводов по основным производственным фондам X (сотни млн. руб.) и суточной выработке продукции Y (тонн) дано в табл. 8.1.

Найти уравнения регрессии Y по X и X по Y, установить степень тесноты связи между переменными с помощью коэффициента корреляции.

Таблица 8.1

Величина ОПФ(сотни млн. руб.) X	Середины интервалов	Суточная выработка продукции Y				Всего n_i	Групповая средняя , $\bar{y}_i$
		4 - 6	6 - 8	8 - 10	10 - 12		
	x_i \ y_j	5	7	9	11		
3 - 5	4	1	4			5	6,6
5 - 7	6		9	8	1	18	29,2
7 - 9	8			10	5	15	29
9 - 11	10			4	8	12	24,8
Всего n_j		1	13	22	14	50	
Групповая средняя $\bar{x}_i$		4	5,385	7,636	9		

Решение. Допустим, что существует линейная корреляционная зависимость Y по X. Уравнение регрессии будем искать в виде :

$$y_x - \bar{y} = b_1(x - \bar{x}).$$

Коэффициент b_1 называют *коэффициентом регрессии* Y по X. Обозначим его b_{yx}. Тогда уравнение регрессии запишется так:

$$y_x - \bar{y} = b_{yx}(x - \bar{x}).$$

Коэффициент регрессии b_{yx} показывает, на сколько единиц в среднем изменится переменная Y при увеличении X на одну единицу.

Коэффициент b_{yx} найдем по формулам $b_{yx} = \frac{\mu}{s_x^2}$, где s_x^2 – выборочная дисперсия переменной X: $s_x^2 = \overline{x^2} - (\bar{x})^2$;

μ – выборочный корреляционный момент : $\mu = \overline{xy} - \bar{x}\bar{y}$.

Вычислим все необходимые суммы в нашем примере:

$\Sigma_{i=1}^4 x_i n_i = 4 \cdot 5 + 6 \cdot 18 + 8 \cdot 15 + 10 \cdot 12 = 368;$

$\Sigma_{i=1}^4 x_i^2 n_i = 16 \cdot 5 + 36 \cdot 18 + 64 \cdot 15 + 100 \cdot 12 = 2888;$

$\Sigma_{i=1}^4 y_j n_j = 5 \cdot 1 + 7 \cdot 13 + 9 \cdot 22 + 11 \cdot 14 = 448;$

$\Sigma_{i=1}^4 y_j^2\, n_j = 25 \cdot 1 + 49 \cdot 13 + 81 \cdot 22 + 121 \cdot 14 = 4138;$

$\Sigma_{i=1}^4 \Sigma_{j=1}^4\, x_i y_j n_{ij} = 4(5 \cdot 1 + 7 \cdot 4) + 6(7 \cdot 9 + 9 \cdot 8 + 11 \cdot 1) +$

$+8(9 \cdot 10 + 11 \cdot 5) + 10(9 \cdot 4 + 11 \cdot 8) = 3408.$

Затем находим параметры уравнения регрессии:

$\bar{x} = \Sigma_{i=1}^4 x_i n_i / n = 368/50 = 7{,}36;\ \bar{y} = \Sigma_{i=1}^4 y_j n_j / n = 448/50 = 8{,}96;$

$s_x^2 = \overline{x^2} - \bar{x}^2 = \Sigma_{i=1}^4 x_i^2 n_i / n - (\bar{x})^2 = 2888/50 - 7{,}36^2 = 3{,}59;$

$s_y^2 = \overline{y^2} - \bar{y}^2 = \Sigma_{j=1}^4 y_i^2 n_j / n - (\bar{y})^2 = \dfrac{4138}{50} - 8{,}96^2 = 2{,}4784;$

$\mu = \overline{xy} - \bar{x}\bar{y} = \Sigma_{i=1}^4 \Sigma_{j=1}^4\, x_i y_j n_{ij} / 50 - \bar{x}\bar{y} = 3408/50 - 7{,}36 \cdot 8{,}96 = 2{,}21.$

Найдем выборочные коэффициенты регрессии b_{yx} и b_{xy}:

$$b_{yx} = \frac{\mu}{s_x^2} = \frac{2{,}21}{3{,}59} = 0{,}617; \quad b_{xy} = \frac{\mu}{s_y^2} = \frac{2{,}21}{2{,}4784} = 0{,}893.$$

Уравнения регрессии а) Y по X, б) X по Y.

а) $y_x - \bar{y} = b_{yx}(x - \bar{x}) \Rightarrow\ y_x - 8{,}96 = 0{,}617(x - 7{,}36) \Rightarrow$

$y_x = 0{,}617x + 4{,}421.$

б) $x_y - \bar{x} = b_{xy}(y - \bar{y}) \Rightarrow\ x_y - 7{,}36 = 0{,}893(y - 8{,}96) \Rightarrow$

$x_y = 0{,}893y - 0{,}646.$

Для проверки тесноты связи между случайными переменными можно использовать выборочный коэффициент корреляции:

$$r = \frac{\overline{xy} - \bar{x}\bar{y}}{s_x s_y} = \frac{\mu}{s_x s_y} = 0{,}742.$$

Это число близко к единице, поэтому связь между переменными прямая и достаточно тесная.

8.3. Нелинейная регрессия

Соотношения между величинами, встречающимися в экономических задачах, далеко не всегда можно выразить линейными функциями. Выбор уравнения регрессии производится на основе предыдущего опыта, литературных источников, а также визуального наблюдения расположения точек корреляционного поля. Наиболее часто встречаются следующие виды уравнений нелинейной регрессии: *полиномиальное* $y_x = b_0 + b_1 x + \cdots + b_k x^k$, *гиперболическое* $y_x = b_0 + \frac{b_1}{x}$, *степенное* $y_x = b_0 \cdot x_1^{b_1} \cdot \ldots \cdot x_p^{b_p}$.

Для подбора зависимости можно использовать метод наименьших квадратов. В программе Excel имеется большой выбор функций (трендов), которые используются для аппроксимации набора данных.

Пример 3. По данным, приведенным в табл. 8.2, исследовать зависимость урожайности зерновых культур Y (ц/га) от количества осадков X (см), выпавших в вегетационный период.

Таблица 8.2

N	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Количество осадков x_i	25	27	30	35	36	38	39	41	42	45	46	47	50	52	53
Урожайность y_i (ц/га)	23	24	27	27	32	31	33	35	34	32	29	28	25	24	25

Решение. Можно предположить, что увеличение количества выпавших осадков приводит к увеличению урожайности до некоторого предела, после чего урожайность будет снижаться. Учитывая также расположение точек на диаграмме (рис. 12), будем считать, что линию регрессии нужно искать в виде па-

раболы с уравнением $y = b_0 + b_1x + b_2x^2$. Это можно сделать с помощью программы Excel (см. 9.5). Результат представлен на рисунке 12. Программа вычислила коэффициенты b_0, b_1, b_2 и получила уравнение регрессии: $y = -0{,}0484x^2 + 3{,}8342x - 43{,}932 = f_2(x)$. Это уравнение совпадает с уравнением регрессии в книге [2] п.13.3. Получим для сравнения линейное уравнение регрессии. Программа вернула уравнение.

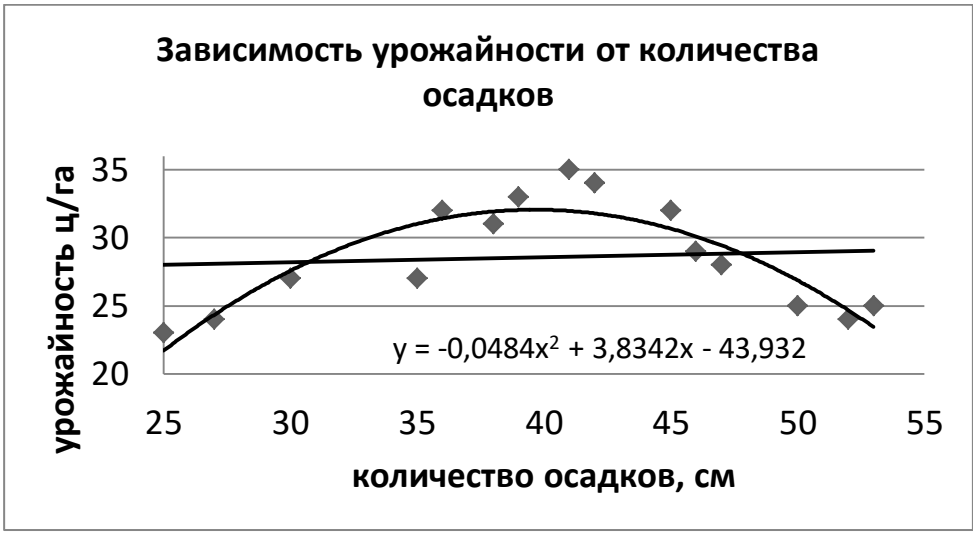


Рис. 12

$y = 0{,}037x + 27{,}106 = f_1(x)$. Из рисунка видно, что график линейной зависимости не отражает главных особенностей распределения экспериментальных данных. Судя по точкам, функция $y = y(x)$ должна сначала возрастать, а затем убывать. Тренд второй степени хорошо демонстрирует эту особенность. Другим способом оценить близость эмпирических данных и теоретической зависимости является сравнение остаточных дисперсий. Предпочтение должно быть отдано тому уравнению, для которого остаточная дисперсия меньше. Остаточная дисперсия вычисляется по формуле:

$$Q_e = \Sigma_{i=1}^{n}\left(y_i - y_{x_i}\right)^2.$$

Для ее вычисления составим таблицу 8.3. В этой таблице второй и третий столбцы совпадают со строками таблицы 8.2. В столбцах 5 и 7 вычислены квадраты разностей $\left(y_i - y_{x_i}\right)^2$, где y_{x_i} вычисляются как значения функций $f_2(x_i)$ и $f_1(x_i)$ в точках x_i. Суммируя эти квадраты разностей, найдем остаточные дисперсии.

Для функции $f_2(x) = -0{,}0484x^2 + 3{,}8342x - 43{,}932$ остаточная дисперсия равна 46,02 , а для $f_1(x) = 0{,}037x + 27{,}106$ остаточная дисперсия равна 222,14, т.е. почти в 5 раз больше. Очевидно, в данном случае нужно выбрать в качестве линии регрессии параболу.

Таблица 8.3

i	x_i	y_i	$f_2(x_i)$	$\left(y_i - f_2(x_i)\right)^2$	$f_1(x_i)$	$\left(y_i - f_1(x_i)\right)^2$
1	25	23	21,67	1,76	28,031	25,31
2	27	24	24,31	0,09	28,105	16,85
3	30	27	27,53	0,29	28,216	1,48
4	35	27	30,98	15,80	28,401	1,96
5	36	32	31,37	0,39	28,438	12,69
6	38	31	31,88	0,77	28,512	6,19
7	39	33	31,99	1,03	28,549	19,81
8	41	35	31,91	9,55	28,623	40,67
9	42	34	31,73	5,17	28,66	28,52
10	45	32	30,60	1,97	28,771	10,43
11	46	29	30,03	1,05	28,808	0,04
12	47	28	29,36	1,85	28,845	0,71
13	50	25	26,78	3,16	28,956	15,65
14	52	24	24,57	0,33	29,03	25,30
15	53	25	23,33	2,81	29,067	16,54
				46,02		222,14

Было бы интересно испробовать уравнение более высокой степени. В принципе можно найти такой многочлен высокой степени, что линия регрессии пройдет прямо через все эмпирические точки. Однако чрезмерное повышение порядка кривой может привести к неоправданному усложнению уравнения регрессии, когда случайные колебания эмпирических точек будут восприниматься как определенные закономерности в поведении линии регрессии. Поэтому на практике обычно редко используются уравнения регрессии степени выше трех.

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

Методами корреляционно-регрессивного анализа изучить влияние факторного признака X на результативный признак Y:

1) определить уравнение регрессии;

2) вычислить выборочные показатели тесноты связи между признаками.

8.1. Зависимость себестоимости единицы продукции Y (тыс. руб./шт.) предприятия от трудоемкости единицы продукции X (чел/час):

X	1	2	3	4	5	6
Y	1,8	4,5	7,1	9,2	11,2	14,5

8.2. Зависимость потребления материала Y (у.е.) от объема производства продукции X (тыс. шт.):

X	1	2	3	4	5	6
Y	2,7	5,9	8,5	11,9	15,1	17,9

8.3. Зависимость среднемесячной производительности труда Y (детали/час) от количества обучений на курсах повышения квалификации X (ед.):

X	1	2	3	4	5	6
Y	2,8	8,1	13,4	19,1	23,8	29,1

В задачах 8.4 и 8.5 зависимость между случайными величинами X и Y задается в виде корреляционной таблицы.

1) Найти уравнения регрессии;

2) вычислить коэффициент корреляции и сделать выводы о тесноте связи между переменными.

8.4.

X	Y			
	5	7	9	11
4	**5**	**1**		
5		**5**	**10**	**1**
6		**3**	**12**	**2**
7		**1**	**2**	**5**

8.5.

X	Y			
	2	4	6	8
10	**4**	**2**		
20		**6**	**8**	
30		**4**	**10**	**2**
40		**2**	**2**	**5**

ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

Методами корреляционно-регрессивного анализа изучить влияние факторного признака X на результативный признак Y:

1) определить уравнение регрессии;

2) вычислить выборочные показатели тесноты связи между признаками.

8.6. Зависимость объема производства Y (тыс. ед.) от численности занятых X (100 чел.):

X	1	2	3	4	5	6
Y	5,1	6,4	7,6	8,9	9,9	10,8

8.7. Зависимость объема продаж Y (тыс. руб.) от расходов на рекламу (тыс. руб.).

X	1	2	3	4	5	6
Y	0,3	1,5	3,3	4,6	6,1	7,5

В задаче 8.8 зависимость между случайными величинами X и Y задается в виде корреляционной таблицы.

1) Найти уравнения регрессии;

2) вычислить коэффициент корреляции и сделать выводы о тесноте связи между переменными.

8.8.

X	Y			
	4	6	8	10
5	**2**	**1**		
10		**6**	**6**	
15		**4**	**8**	**2**
20			**2**	**6**

9. Решение задач математической статистики с помощью программы EXCEL

9.1. Функция БИНОМ. РАСП.

В программе EXCEL имеется много функций, которые можно использовать при решении задач теории вероятностей и математической статистики. Например, при решении задачи о вычислении вероятности того, что в n опытах событие A произойдет k раз, т.е. о поиске $P_n(k)$.

Чтобы найти $P_n(k)$, можно использовать функцию БИНОМ. РАСП. Для этого следует щелкнуть указателем мыши по значку $f(x)$. В появившемся меню выбрать категорию «Статистические», в этом разделе найти функцию БИНОМ. РАСП. При щелчке по этому названию появляется диалоговое окно, т.е.

табличка с надписями и пустыми окнами, в которые мы должны записать данные нашей задачи. В сокращенном виде эта табличка выглядит так:

БИНОМ. РАСП				
Число_успехов	3	=	3	
Число_ испытаний	10	=	10	
Вероятность_успеха	0,15	=	0,15	
Интегральная	Ложь	=	Ложь	
		=	0,129834	
			ОК	Отмена

Чтобы использовать данную функцию, следует заполнить пустые ячейки второго столбца, в котором надо указать, сколько раз произошло событие A (число успехов k); число испытаний n; вероятность успеха p. Ответ появится в четвертом столбце, ниже всех данных. Чтобы использовать его в дальнейших вычислениях, надо нажать клавишу ОК. Мы вернемся в основное рабочее поле программы, и ответ будет в той ячейке, где мы были, когда обратились к функции БИНОМ. РАСП.

Пример 1. В среднем по 15% договоров страховая компания выплачивает страховую сумму. Найти вероятность того, что из десяти договоров будут выплачены страховые суммы по а) трем договорам; б) менее чем по трем договорам, в) от двух до пяти договоров.

Решение. В случае а) нам нужно найти $P_{10}(3)$, при $p = 0{,}15$. Заполним графы диалогового окна **Число_успехов** – 3; **Число_ испытаний** – 10; **Вероятность_успеха** – 0,15; **Интегральная** – ЛОЖЬ. Против последнего равенства появляется ответ: 0,275897.

Замечание. В графе **Интегральная** следует писать ЛОЖЬ, если мы хотим вычислить значение $P_n(k)$ при одном значении k. Если же написать **Интегральная** – ИСТИНА, то будет возвращена сумма $P_n(0)$ +$P_n(1)$ +… +$P_n(k)$. Для ответа на вопрос задачи б) следует набрать окна **Число_успехов** – 2; **Число_ испытаний** – 10; **Вероятность_успеха** – 0,15; **Интегральная** – ИСТИНА. Возвращается ответ: 0,820196.

При решении задачи в) надо вычислять по формуле:

$$P_{10}(2) + P_{10}(3) + P_{10}(4) + P_{10}(5) =$$

$$= (P_{10}(0) + P_{10}(1) + P_{10}(2) + P_{10}(3) + P_{10}(4) + P_{10}(5)) - (P_{10}(0) + P_{10}(1)).$$

Возвращается ответ: 0,998617- 0,5443= 0,454317.

9.2. Вычисление числовых характеристик

Средняя арифметическая $\bar{x}$. Вычисляется с помощью функции СРЗНАЧ. Возвращает среднее арифметическое своих аргументов, которые могут быть числами, именами, массивами или ссылками на ячейки с числами. Например, надо найти среднее арифметическое чисел, находящихся в массиве ячеек A1, A2, A3, A4, A5. Найдите и активизируйте СРЗНАЧ (раздел меню «Статистические»). Возникнет диалоговое окно вида:

СРЗНАЧ				
	Число 1	A1:A5		
	Число 2			
			=	
			ОК	Отмена

Если мы находимся в ячейке A6, т.е. под нужным массивом, то программа сама предложит найти нужное среднее арифметическое. Осталось только нажать ОК. Если это не так, то можно стереть предложение программы и записать в первой графе A1:A5 или выделить нужный массив указателем мыши. При этом пустые ячейки будут игнорироваться. Нажать ОК.

Функция используется в случае, когда значения x_i просто перечислены, и мы считаем среднее по формуле: $\bar{x} = (x_1 + x_2 + \cdots + x_n)/n$. Если же значения x_i повторяются и известны их частоты n_i, то следует считать по формуле: $x = \Sigma x_i n_i / n$. В этом случае можно использовать функцию СУММПРОИЗВ (раздел меню «Математические»), которая почленно перемножает два массива $\{x_i\}$ и $\{n_i\}$, а результаты складывает. Число n находим с помощью функции СУММ (раздел меню «Математические»).

Дисперсия. Можно использовать две функции: ДИСП.Г или ДИСП.В (раздел меню «Статистические»). Первая функция возвращает дисперсию D_x, а вторая – исправленную s_x^2.

В случае если даны частоты x_i, следует вычислять по формуле: $D_x = \overline{x^2} - (\bar{x})^2$. Чтобы вычислить $\overline{x^2}$, надо создать массив x_i^2, а затем применить функцию СУММПРОИЗВ. Эта функция перемножит массивы $\{x_i^2\}$ и $\{n_i\}$ и найдет сумму таких произведений.

Среднее квадратическое отклонение σ. Вычисляется с помощью двух функций: СТАНДОТКЛОН.Г и СТАНДОТКЛОН.В (раздел меню «Статистические») . Первая возвращает $\sigma_x = \sqrt{D_x}$, вторая – $s_x = \sqrt{s_x^2}$.

Медиана. Медиана – это число, которое является серединой множества чисел. Построим числа в порядке возрастания. Если количество чисел нечетное,

то будет возвращено среднее из этих чисел. Если количество – четное, то будет возвращено среднее арифметическое находящихся в середине чисел. Например, для чисел 1; 2; 10 медиана равна 2. Для чисел 1;2;10;100 медиана равна: $6 = \frac{2+10}{2}$. Использовать функцию МЕДИАНА (раздел меню «Статистические»).

Мода. Функция МОДА.НСК возвращает то число, которое чаще всего встречается в списке.

9.3. Нормальное распределение

Для вычисления функции плотности нормального распределения используется функция программы Excel НОРМ.РАСП (раздел меню «Статистические»). Диалоговое окно для этой функии имеет вид:

НОРМ.РАСП			
	X	1	=1
	Среднее	0	=0
	Стандартное_откл	1	=1
	Интегральная	ЛОЖЬ	=ЛОЖЬ
			=0,2419…

Надо заполнить третий столбец. Здесь X – значение аргумента, при котором надо вычислить значение функции; **Среднее** – m_x ; **Стандартное_откл** – Среднее квадратическое отклонение σ; если **Интегральная** – Ложь, будут возвращены значения функции плотности нормального распределения:

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

В случае, если **Интегральная** – Истина, будет возвращена функция распределения для нормального рспределения:

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

Чтобы найти вероятность для случайной величины X попадания в интервал (α, β), надо воспользоваться формулой $P(\alpha < X < \beta) = F(\beta) - F(\alpha)$.

9.4. Выборочный линейный коэффициент корреляции

Коэффициент корреляции r_b может быть вычислен с помощью функции КОРРЕЛ (раздел меню «Статистические»). В диалоговом окне предлагается

выбрать два массива (с одинаковым числом элементов). Функция возвращает значение выборочного линейного коэффициента корреляции.

9.5. Метод наименьших квадратов

С помощью этого метода можно для данного набора значений (x_i, y_i) найти подходящую функцию $y = f(x)$, которая в точках x_i близка к y_i. Предварительно надо выбрать класс функций, например, линейных или квадратичных. Наилучшим приближением считается такая функция, для которой сумма квадратов $\Sigma_{i=1}^{n}\left(y_i - y(x_i)\right)^2 = min.$ Для реализации этого метода в программе EXCEL надо произвести следующие действия: 1) записать данные $\{x_i\}$ и $\{y_i\}$ в виде двух массивов; 2)выделив эти массивы, построить график, состоящий из точек (x_i, y_i) нажимая кнопки меню «Вставить» «Точечная», выбрать тип графика. После щелчка по значку «Точечная» появится диаграмма с рядом точек. Щелкнуть правой клавишей мыши по одной из этих точек, возникнет меню, в котором выбрать надпись «Добавить линию тренда». В возникшем меню выбрать тип тренда. Например, «Полиномиальная» «степень 2». Поставить галочку в окошке «Показать уравнение на диаграмме» «Закрыть». Появится график тренда (аппроксимирующей функции) и его уравнение.

ЗАДАЧИ ДЛЯ РЕШЕНИЯ НА ЗАНЯТИИ

9.1. Батарея сделала 20 выстрелов по военному объекту. Вероятность попадания при каждом выстреле равна 0,7. Найти вероятность, что: а) число попаданий будет равно 10; б) число попаданий будет в пределах от 10 до 15.

9.2. При измерении величины X получен ряд значений :

123,08 121,59 121,01 120,67 120,50 120,47 120,49 120,47

Найти среднее значение $\bar{x}$, дисперсию D_x, исправленное значение s_x^2, среднее квадратическое отклонение σ_x, исправленное значение s_x, медиану выборки.

9.3. Дано распределение 50 данных x_i с частотами n_i.

x_i	1	2	3	4	5	6
n_i	2	3	6	8	22	9

Найти среднее значение $\bar{x}$, дисперсию D_x, исправленное значение s_x^2 и среднее квадратическое отклонение σ_x.

9.4. Случайная величина X распределена по нормальному закону с $a = 25$, $\sigma = 10$. Найти: а) значения функции плотности распределения $f(x)$ в

точках $x = 5; x = 10; x = 15; x = 20; x = 25$; б) вероятность попадания X в интервал $[5; 20]$.

9.5. Найти выборочный коэффициент корреляции между значениями переменных X и Y.

X	1	2	3	4	5	6
Y	3,5	4,2	5,1	5,9	6,9	7,8

9.6. Даны две дискретные функции:

а)

X	1	2	3	4	5	6
Y	3,5	4,2	5,1	5,9	6,9	7,8

б)

X	1,5	2,5	3	4	5	6	8
Y	15	21	13	3	4	3	1

Построить точечные графики для обеих функций, построить линии трендов и найти их уравнения. В случае а) испробовать линейный тренд, в случае б) испробовать полиномиальные тренды третьей и пятой степени. Сравнить результаты.

9.7. Нормально распределенная случайная величина измерялась в 20 испытаниях. Получилось среднее значение $\bar{x} = 24$, исправленное среднее квадратическое отклонение $s = 7$. Найти доверительный интервал для среднего генеральной совокупности, с надежностью $\gamma = 0{,}99$.

Указание. Доверительный интервал для среднего имеет вид:

$$\left(\bar{x} - \frac{t_\gamma s}{\sqrt{n-1}}, \bar{x} + \frac{t_\gamma s}{\sqrt{n-1}}\right),$$

где t_γ находим с помощью таблицы или функции СТЬЮДЕНТ.ОБР.2Х. При использовании функции СТЬЮДЕНТ.ОБР.2Х приходится указать значения следующих величин: вероятности и степени свободы. Вероятность равна $1 - \gamma$, а в графе «степени свободы» следует написать: $n - 1$, где n – число испытаний.

Приложение 1

Функция Лапласа $\Phi(x) = \frac{2}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}}dt$

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0	0	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0,080	0,088	0,096	0,103	0,111	0,119	0,127	0,135	0,143	0,151
0,2	0,1585	0,1663	0,1741	0,1819	0,1897	0,1974	0,2051	0,2128	0,2205	0,2282
0,3	0,2358	0,2434	0,2510	0,2586	0,2661	0,2737	0,2812	0,2886	0,2961	0,3035
0,4	0,3108	0,3182	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,5	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,6	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,7	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,8	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,9	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6680	0,6729	0,6778
1	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	0,7287	0,7330	0,7373	0,7415	0,7457	0,7499	0,7540	0,7580	0,7620	0,7660
1,2	0,7699	0,7737	0,7775	0,7813	0,7850	0,7887	0,7923	0,7959	0,7995	0,8029
1,3	0,8064	0,8098	0,8132	0,8165	0,8198	0,8230	0,8262	0,8293	0,8324	0,8355
1,4	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,5	0,8664	0,8690	0,8715	0,8740	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,6	0,8904	0,8926	0,8948	0,8969	0,8990	0,9011	0,9031	0,9051	0,9070	0,9090
1,7	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,8	0,9281	0,9297	0,9312	0,9328	0,9342	0,9357	0,9371	0,9385	0,9399	0,9412
1,9	0,9426	0,9439	0,9451	0,9464	0,9476	0,9488	0,9500	0,9512	0,9523	0,9534
2	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9615	0,9625	0,9634
2,1	0,9643	0,9651	0,9660	0,9668	0,9676	0,9684	0,9692	0,9700	0,9707	0,9715
2,2	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,9780
2,3	0,9786	0,9791	0,9797	0,9802	0,9807	0,9812	0,9817	0,9822	0,9827	0,9832
2,4	0,9836	0,9840	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,5	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,6	0,9907	0,9909	0,9912	0,9915	0,9917	0,9920	0,9922	0,9924	0,9926	0,9929
2,7	0,9931	0,9933	0,9935	0,9937	0,9939	0,9940	0,9942	0,9944	0,9946	0,9947
2,8	0,9949	0,9950	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,9960	0,9961
2,9	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972
3	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9980
3,1	0,9981	0,9981	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,2	0,9986	0,9987	0,9987	0,9988	0,9988	0,9988	0,9989	0,9989	0,9990	0,9990
3,3	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,4	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	0,9995
3,5	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997
3,6	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	0,9998	0,9998
3,7	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

Приложение 2

Значения функции $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	0,3970	0,3965	0,3961	0,3956	0,3951	0,3945	0,3939	0,3932	0,3925	0,3918
0,2	0,3910	0,3902	0,3894	0,3885	0,3876	0,3867	0,3857	0,3847	0,3836	0,3825
0,3	0,3814	0,3802	0,3790	0,3778	0,3765	0,3752	0,3739	0,3725	0,3712	0,3697
0,4	0,3683	0,3668	0,3653	0,3637	0,3621	0,3605	0,3589	0,3572	0,3555	0,3538
0,5	0,3521	0,3503	0,3485	0,3467	0,3448	0,3429	0,3410	0,3391	0,3372	0,3352
0,6	0,3332	0,3312	0,3292	0,3271	0,3251	0,3230	0,3209	0,3187	0,3166	0,3144
0,7	0,3123	0,3101	0,3079	0,3056	0,3034	0,3011	0,2989	0,2966	0,2943	0,2920
0,8	0,2897	0,2874	0,2850	0,2827	0,2803	0,2780	0,2756	0,2732	0,2709	0,2685
0,9	0,2661	0,2637	0,2613	0,2589	0,2565	0,2541	0,2516	0,2492	0,2468	0,2444
1	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	0,2179	0,2155	0,2131	0,2107	0,2083	0,2059	0,2036	0,2012	0,1989	0,1965
1,2	0,1942	0,1919	0,1895	0,1872	0,1849	0,1826	0,1804	0,1781	0,1758	0,1736
1,3	0,1714	0,1691	0,1669	0,1647	0,1626	0,1604	0,1582	0,1561	0,1539	0,1518
1,4	0,1497	0,1476	0,1456	0,1435	0,1415	0,1394	0,1374	0,1354	0,1334	0,1315
1,5	0,1295	0,1276	0,1257	0,1238	0,1219	0,1200	0,1182	0,1163	0,1145	0,1127
1,6	0,1109	0,1092	0,1074	0,1057	0,1040	0,1023	0,1006	0,0989	0,0973	0,0957
1,7	0,0940	0,0925	0,0909	0,0893	0,0878	0,0863	0,0848	0,0833	0,0818	0,0804
1,8	0,0790	0,0775	0,0761	0,0748	0,0734	0,0721	0,0707	0,0694	0,0681	0,0669
1,9	0,0656	0,0644	0,0632	0,0620	0,0608	0,0596	0,0584	0,0573	0,0562	0,0551
2	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0,0440	0,0431	0,0422	0,0413	0,0404	0,0396	0,0387	0,0379	0,0371	0,0363
2,2	0,0355	0,0347	0,0339	0,0332	0,0325	0,0317	0,0310	0,0303	0,0297	0,0290
2,3	0,0283	0,0277	0,0270	0,0264	0,0258	0,0252	0,0246	0,0241	0,0235	0,0229
2,4	0,0224	0,0219	0,0213	0,0208	0,0203	0,0198	0,0194	0,0189	0,0184	0,0180
2,5	0,0175	0,0171	0,0167	0,0163	0,0158	0,0154	0,0151	0,0147	0,0143	0,0139
2,6	0,0136	0,0132	0,0129	0,0126	0,0122	0,0119	0,0116	0,0113	0,0110	0,0107
2,7	0,0104	0,0101	0,0099	0,0096	0,0093	0,0091	0,0088	0,0086	0,0084	0,0081
2,8	0,0079	0,0077	0,0075	0,0073	0,0071	0,0069	0,0067	0,0065	0,0063	0,0061
2,9	0,0060	0,0058	0,0056	0,0055	0,0053	0,0051	0,0050	0,0048	0,0047	0,0046
3	0,0044	0,0043	0,0042	0,0040	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0,0033	0,0032	0,0031	0,0030	0,0029	0,0028	0,0027	0,0026	0,0025	0,0025
3,2	0,0024	0,0023	0,0022	0,0022	0,0021	0,0020	0,0020	0,0019	0,0018	0,0018
3,3	0,0017	0,0017	0,0016	0,0016	0,0015	0,0015	0,0014	0,0014	0,0013	0,0013
3,4	0,0012	0,0012	0,0012	0,0011	0,0011	0,0010	0,0010	0,0010	0,0009	0,0009
3,5	0,0009	0,0008	0,0008	0,0008	0,0008	0,0007	0,0007	0,0007	0,0007	0,0006
3,6	0,0006	0,0006	0,0006	0,0005	0,0005	0,0005	0,0005	0,0005	0,0005	0,0004
3,7	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0003	0,0003	0,0003	0,0003
3,8	0,0003	0,0003	0,0003	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002	0,0002

Приложение 3

Значения функции Пуассона $P(X = m) = \frac{\lambda^m}{m!} e^{-\lambda}$

	λ									
m	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679
1	0,0905	0,1637	0,2222	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659	0,3679
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1217	0,1438	0,1647	0,1839
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494	0,0613
4	0,0000	0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111	0,0153
5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0004	0,0007	0,0012	0,0020	0,0031
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0003	0,0005
7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001

	λ								
m	2	3	4	5	6	7	8	9	10
0	0,1353	0,0498	0,0183	0,0067	0,0025	0,0009	0,0003	0,0001	0,0000
1	0,2707	0,1494	0,0733	0,0337	0,0149	0,0064	0,0027	0,0011	0,0005
2	0,2707	0,2240	0,1465	0,0842	0,0446	0,0223	0,0107	0,0050	0,0023
3	0,1804	0,2240	0,1954	0,1404	0,0892	0,0521	0,0286	0,0150	0,0076
4	0,0902	0,1680	0,1954	0,1755	0,1339	0,0912	0,0573	0,0337	0,0189
5	0,0361	0,1008	0,1563	0,1755	0,1606	0,1277	0,0916	0,0607	0,0378
6	0,0120	0,0504	0,1042	0,1462	0,1606	0,1490	0,1221	0,0911	0,0631
7	0,0034	0,0216	0,0595	0,1044	0,1377	0,1490	0,1396	0,1171	0,0901
8	0,0009	0,0081	0,0298	0,0653	0,1033	0,1304	0,1396	0,1318	0,1126
9	0,0002	0,0027	0,0132	0,0363	0,0688	0,1014	0,1241	0,1318	0,1251
10	0,0000	0,0008	0,0053	0,0181	0,0413	0,0710	0,0993	0,1186	0,1251
11	0,0000	0,0002	0,0019	0,0082	0,0225	0,0452	0,0722	0,0970	0,1137
12	0,0000	0,0001	0,0006	0,0034	0,0113	0,0263	0,0481	0,0728	0,0948
13	0,0000	0,0000	0,0002	0,0013	0,0052	0,0142	0,0296	0,0504	0,0729
14	0,0000	0,0000	0,0001	0,0005	0,0022	0,0071	0,0169	0,0324	0,0521
15	0,0000	0,0000	0,0000	0,0002	0,0009	0,0033	0,0090	0,0194	0,0347
16	0,0000	0,0000	0,0000	0,0000	0,0003	0,0014	0,0045	0,0109	0,0217
17	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0021	0,0058	0,0128
18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0029	0,0071
19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0014	0,0037
20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0006	0,0019
21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0009
22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004
23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002

Приложение 4

Распределение Стьюдента. Таблица значений $t_\gamma = t(\gamma, n)$

n	k	γ			n	k	γ		
		0,95	0,99	0,999			0,95	0,99	0,999
5	4	2,78	4,60	8,61	20	19	2,09	2,86	3,88
6	5	2,57	4,03	6,87	25	24	2,06	2,80	3,75
7	6	2,45	3,71	5,96	30	29	2,05	2,76	3,66
8	7	2,36	3,50	5,41	35	34	2,03	2,73	3,60
9	8	2,31	3,36	5,04	40	39	2,02	2,71	3,56
10	9	2,26	3,25	4,78	45	44	2,02	2,69	3,53
11	10	2,23	3,17	4,59	50	49	2,01	2,68	3,50
12	11	2,20	3,11	4,44	60	59	2,00	2,66	3,46
13	12	2,18	3,05	4,32	70	69	1,99	2,65	3,44
14	13	2,16	3,01	4,22	80	79	1,99	2,64	3,42
15	14	2,14	2,98	4,14	90	89	1,99	2,63	3,40
16	15	2,13	2,95	4,07	100	99	1,98	2,63	3,39
17	16	2,12	2,92	4,01	120	119	1,98	2,62	3,37
18	17	2,11	2,90	3,97	∞	∞	1,96	2,58	3,29
19	18	2,10	2,88	3,92					

$k = n - 1$ – число степеней свободы

Приложение 5

x	e^x	x	e^x	x	e^x	x	e^x	x	e^x	x	e^x
-5	0,007	-3,8	0,022	-2	0,135	-0,2	0,819	1,7	5,474	3,4	29,964
-4,9	0,007	-3,7	0,025	-1,9	0,150	-0,1	0,905	1,8	6,050	3,5	33,115
-4,8	0,008	-3,6	0,027	-1,8	0,165	0	1,000	1,9	6,686	3,6	36,598
-4,7	0,009	-3,5	0,030	-1,7	0,183	0,1	1,105	2	7,389	3,7	40,447
-4,6	0,010	-3,4	0,033	-1,6	0,202	0,2	1,221	2,1	8,166	3,8	44,701
-4,5	0,011	-3,3	0,037	-1,5	0,223	0,3	1,350	2,2	9,025	3,9	49,402
-4,4	0,012	-3,2	0,041	-1,4	0,247	0,4	1,492	2,3	9,974	4	54,598
-4,3	0,014	-3,1	0,045	-1,3	0,273	0,5	1,649	2,4	11,023	4,1	60,340
-4,2	0,015	-2,9	0,055	-1,2	0,301	0,6	1,822	2,5	12,182	4,2	66,686
-4,1	0,017	-2,8	0,061	-1,1	0,333	0,7	2,014	2,6	13,464	4,3	73,700
-4	0,018	-2,7	0,067	-0,9	0,407	0,8	2,226	2,7	14,880	4,4	81,451
-3,9	0,020	-2,6	0,074	-0,8	0,449	0,9	2,460	2,8	16,445	4,5	90,017
		-2,5	0,082	-0,7	0,497	1,2	3,320	2,9	18,174	4,6	99,484
		-2,4	0,091	-0,6	0,549	1,3	3,669	3	20,086	4,7	109,947
		-2,3	0,100	-0,5	0,607	1,4	4,055	3,1	22,198	4,8	121,510
		-2,2	0,111	-0,4	0,670	1,5	4,482	3,2	24,533	4,9	134,290
		-2,1	0,122	-0,3	0,741	1,6	4,953	3,3	27,113	5	148,413

Приложение 6

Значения $\chi^2_{\alpha k}$ критерия Пирсона

Число степеней свободы k	Вероятность α							
	0,99	**0,98**	**0,95**	**0,9**	**0,1**	**0,05**	**0,02**	**0,01**
1	0,00	0,00	0,00	0,02	2,71	3,84	5,41	6,63
2	0,02	0,04	0,10	0,21	4,61	5,99	7,82	9,21
3	0,11	0,18	0,35	0,58	6,25	7,81	9,84	11,34
4	0,30	0,43	0,71	1,06	7,78	9,49	11,67	13,28
5	0,55	0,75	1,15	1,61	9,24	11,07	13,39	15,09
6	0,87	1,13	1,64	2,20	10,64	12,59	15,03	16,81
7	1,24	1,56	2,17	2,83	12,02	14,07	16,62	18,48
8	1,65	2,03	2,73	3,49	13,36	15,51	18,17	20,09
9	2,09	2,53	3,33	4,17	14,68	16,92	19,68	21,67
10	2,56	3,06	3,94	4,87	15,99	18,31	21,16	23,21
11	3,05	3,61	4,57	5,58	17,28	19,68	22,62	24,72
12	3,57	4,18	5,23	6,30	18,55	21,03	24,05	26,22
13	4,11	4,77	5,89	7,04	19,81	22,36	25,47	27,69
14	4,66	5,37	6,57	7,79	21,06	23,68	26,87	29,14
15	5,23	5,98	7,26	8,55	22,31	25,00	28,26	30,58
16	5,81	6,61	7,96	9,31	23,54	26,30	29,63	32,00
17	6,41	7,26	8,67	10,09	24,77	27,59	31,00	33,41
18	7,01	7,91	9,39	10,86	25,99	28,87	32,35	34,81
19	7,63	8,57	10,12	11,65	27,20	30,14	33,69	36,19
20	8,26	9,24	10,85	12,44	28,41	31,41	35,02	37,57
21	8,90	9,91	11,59	13,24	29,62	32,67	36,34	38,93
22	9,54	10,60	12,34	14,04	30,81	33,92	37,66	40,29
23	10,20	11,29	13,09	14,85	32,01	35,17	38,97	41,64
24	10,86	11,99	13,85	15,66	33,20	36,42	40,27	42,98
25	11,52	12,70	14,61	16,47	34,38	37,65	41,57	44,31
26	12,20	13,41	15,38	17,29	35,56	38,89	42,86	45,64
27	12,88	14,13	16,15	18,11	36,74	40,11	44,14	46,96
28	13,56	14,85	16,93	18,94	37,92	41,34	45,42	48,28
29	14,26	15,57	17,71	19,77	39,09	42,56	46,69	49,59
30	14,95	16,31	18,49	20,60	40,26	43,77	47,96	50,89

ОТВЕТЫ

Глава I

1.1. $w = 0{,}05$. **1.2**. 0,36. **1.3.** 0,061101. **1.4**. 7!=5040. **1.5**. $A_{30}^{3} = 24360$. **1.6**. $C_{16}^{2} = 120$. **1.7.** 0,1258. **1.8.** $P(A) = 0{,}004$; $P(B) = 0{,}126$; $P(C) = 0{,}122$; $P(D) = 0{,}874$; $P(E) = 0{,}996$. **1.9.** 0,664. **1.10.** $P(A) = 0{,}24$; $P(B) = 0{,}04$; $P(C) = 0{,}96$. **1.11**. $P(A) = 0{,}008$; $P(B) = 0{,}564$; $P(C) = 0{,}436$; $P(D) = 0{,}116$. **1.12.** $\frac{15}{16}$. **1.13.** а) $C_{10}^{4}/C_{25}^{4} \approx 0{,}0166$; б) $C_{15}^{1}\,C_{10}^{3}/C_{25}^{4} \approx 0{,}142$.**1.14.** 6! =720. **1.15** 0,82. 1.16. а) 0,03; б) 0,34; в) 0,97. 1.17. $P(A) = \frac{1}{60}$; $P(B) = \frac{1}{4}$; $P(C) = \frac{7}{30}$; $P(D) = \frac{3}{4}$.

2.1. 0,86;0,7. **2.2.** 0,77; 0,45. **2.3.** 0,6. **2.4**. 1) 0,4096, 2) 0,9421, 3)0,0067. **2.5.** а) $P_{10}(3) = C_{10}^{3}p^{3}q^{7} = 0{,}1298$; б)$P_{10}(m < 2) = P_{10}(0) + P_{10}(1) = 0{,}544$. **2.6.** а) 2 партии из 4, ибо $P_4(2) = 0{,}375$, а $P_6(3) = 0{,}312$; б) не менее 2 партий из 4, ибо $P_4(m \geq 2) = P_4(2) + P_4(3) + P_4(4) = 0{,}688$, а $P_6(m \geq 3) = 1 - P_6(m < 3) = 1 - \big(P_6(0) + P_6(1) + P_6(2)\big) = 0{,}656$. **2.7.** $P_{10}(5) = 0{,}246$; $P_{10}(6) = 0{,}205$. **2.8.** а) $P_{4000}(3) = 0{,}00715$; б) $P_{4000}(m \leq 3) = 0{,}9992$. Использовать приближенную формулу Пуассона. **2.9.** $P_{400}(180) = 0{,}0054$; б) $P_{400}(180 \leq m \leq 400) = 0{,}977$. **2.10**. а) 0,03; б) 0,167. **2.11.** $P(A) = 0{,}62$; $P(H_1/A) = 0{,}56$. **2.12.** $P_5(4) = 5/32$ **2.13**. $P_{800}(8)$=0,1396. **2.14**. а) $P_{100}(50) = 0{,}07978$; б) $P_{100}(40 \leq m \leq 60) = 0{,}9545$.

Глава II

3.1. $m_x = 5, D_x = 7{,}8$, $\sigma_x = 2{,}79$, $P\{2 \leq X \leq 6\} = 0{,}8$. **3.2.** $m_x = 2{,}7$, $D_x = 16{,}01$, $\sigma_x = 4$, $P\{0 \leq X \leq 5\} = 0{,}7$. **3.3.** $p_1 = 0{,}7$; $p_2 = 0{,}3$.

3.4.

X	0	1	2
$P\{X = x_k\}$	0,08	0,44	0,48

$m_x = 1{,}4$; $D_x = 0{,}4$; $\sigma_x = 0{,}63$.

3.5.

X	0	1	2	3
$P\{X = x_k\}$	0,064	0,288	0,432	0,216

$m_x = 1{,}8;\ D_x = 0{,}72;\ \sigma_x = 0{,}85,\ P\{X \geq 2\} = 0{,}648, Mo[X] = 2.$
3.6. $m_0 = Mo[X] = 16.$ **3.7.** $m_x = 0{,}1; D_x = 2{,}49;\ \sigma_x = 1{,}58;$
$P\{0 \leq X \leq 2\} = 0{,}6.$ **3.8.** $x_2 = 4.$
3.9.

X	0	1	2	3	4
$P\{X = x_k\}$	0,1296	0,3456	0,3456	0,1536	0,0256

3.10. $m_x = 300,\ D_x = 210.$

4.1. $F(x) = 0,$ если $x \leq 0;\ F(x) = x - x^2/4,$ если $0 < x \leq 2;\ F(x) = 1,$ если $x > 2;\ m_x = 2/3;\ D_x = 2/9;\ \sigma_x = \frac{2}{\sqrt{3}};\ P(1 < X < 2) = \frac{1}{4}.$
4.2. $P(X \leq 3) = \frac{3}{5}; m_x = 2{,}5; D_x = \frac{25}{12};\ \sigma_x = 5/2\sqrt{3}$ **. 4.3.** а) $P = 0{,}08;$ б) $P = 0{,}1.$
4.4. 0,6853. **4.5**. 0,023. **4.6.** 0,9545. **4.7.** (85, 115). **4.8.** 0,1079.
4.9. 1) а) $P(X \leq 15{,}3) = F(15{,}3) = 0{,}4332;$ б) $P(X \geq 15{,}4) = 1 - F(15{,}4) = 0{,}0228;$ в) $P(14{,}9 \leq X \leq 15{,}4) = 0{,}6246;$ 2) $14{,}4 \leq X \leq 15{,}6.$
4.10. а) $P(X \leq 470) = 0{,}002;$ б) $P(500 \leq X \leq 550) = 0{,}613;$
в) $P(X > 550) = 0{,}341;$ г) $P(|X - 540| \leq 30) = 0{,}613.$
4.11. $f(x) = \begin{cases} 4, & x \in [0,\ 0{,}25] \\ 0, & x \notin [0,\ 0{,}25] \end{cases}$ $P(0 \leq X \leq 0{,}1) = 0{,}4;\ P_{2,5} = 0{,}35.$
4.12. $F(x) = \begin{cases} 0, & x \in (-\infty, 0] \\ \frac{1}{2}(x^2 + x), & x \in (0, 1] \\ 1, & x \in (1, \infty) \end{cases}$ $m_x = 0{,}58;\ D_x = 0{,}08;\ \sigma_x = 0{,}28;$
$P(1/2 < X < 1) = 0{,}625.$
4.13. 0,84. **4.14**. 0,0876. **4.15**. 0,954

5.1. $P(A) = 0{,}26; P(B) = 0{,}5;\ P(C) = 0{,}32.$
5.2. $P(A) = 0{,}27; P(B) = 0{,}86;\ P(C) = 0{,}25.$
5.3. $P(A) = 0{,}05; P(B) = 0{,}8;\ P(C) = 0{,}25.$
5.4. $P(A) = 0{,}21; P(B) = 0{,}92;\ P(C) = 0{,}29.$
5.5. 0,1681. **5.6**. 0,4135. **5.7**. 0,1353. **5.8.** а) $P(X = 1) = 0{,}8, P(X = 2) = 0{,}2;$ $P(Y = -1) = 0{,}2, P(Y = 0) = 0{,}3, P(Y = 1) = 0{,}3, P(Y = 2) = 0{,}2.$ б) При

$Y = 2$, $p(x_1) = 0{,}75$, $p(x_2) = 0{,}25$; при $X = 1$, $p(y_1) = 0{,}125$, $p(y_2) = 0{,}3125$, $p(y_3) = 0{,}375$, $p(y_4) = 0{,}1875$. в) $P(Y < X) = 0{,}5$. г) $K_{xy} = -0{,}1$, $\rho = -0{,}244$. **5.9.** $P(A) = 0{,}18$; $P(B) = 0{,}86$; $P(C) = 0{,}53$. **5.10.** 0,368.

5.11. $M(X) = 0{,}24$, $M(Y) = 1{,}48$, $\sigma(X) = 0{,}69455$, $\sigma(Y) = 0{,}93252$, $K_{xy} = -0{,}1352$, $\rho_{xy} = -0{,}20874$.

6.1. 1/6. **6.2.** $P \geq 0{,}86$. **6.3.** $P \geq 8/9$. **6.4.** $P \geq 0{,}52$. **6.5.** $P \geq 0{,}6$. **6.6**. $P \geq 0{,}264$. **6.7.** $n \geq 50000$. **6.8.** $P \geq 0{,}95$. **6.9**. $P \geq 0{,}973$. **6.10**. 0,8245. **6.11**. $P \leq 0{,}1$. **6.12**. а) $P \geq 0{,}5$; б) $P \leq 2/3$. **6.13**. $P \geq 0{,}8578$. **6.14.** $P \geq 0{,}9344$.

Глава III

7.1. $\bar{x} = 9{,}146$; $D_B = 19{,}2948$; $\sigma_B = 4{,}3926$; $s^2 = 20{,}6730$; $s = 4{,}5468$; доверительный интервал $(6{,}546; 11{,}746)$.
7.2. $\bar{x} = 16{,}9$; $D_B = 18{,}54$; σ_B =4,30581; $s^2 = 19{,}02$; $s = 4{,}361$; доверительный интервал $(15{,}4895; 18{,}3105)$.
7.3. $\bar{x} = 2{,}025$; $D_B = 1{,}124$; σ_B =1,060; $s^2 = 1{,}143$; $s = 1{,}069$; доверительный интервал $(1{,}433; 2{,}617)$.
7.4. $\bar{x} = 39{,}672$; $D_B = 256{,}66$; σ_B =16,02; $s^2 = 256{,}91$; $s = 16{,}03$.
7.5. $\bar{x} = 15{,}56$; $D_B = 20{,}1264$; σ_B =4,4862; $s^2 = 20{,}3297$; $s = 4{,}5088$.
7.6. Общая средняя 90,02; средняя групповых дисперсий 59,00033; межгрупповая дисперсия 83,17927; общая дисперсия 142,1796. **7.7**. доверительный интервал $(23{,}49; 24{,}51)$. 7.8. $\chi^2 = 1{,}73 < \chi^2_{\text{крит}} = 4{,}61$. Гипотеза о нормальном распределении сл. величины X принимается. **7.9**. 81. **7.10**. $\bar{x}$ =4,38; D_B =2,5416; σ_B =1,5942; s^2 =2,7231; s =1,6502; доверительный интервал $(3{,}4362; 5{,}3238)$. **7.10**. $M(X) = 17{,}13$; $D_x = 17{,}86$; $\sigma_x = 4{,}23$; $J_\varepsilon = (9{,}126; 9{,}166)$.

8.1. $\bar{y}(x) = 2{,}45x - 0{,}52$; $r = 0{,}9977$. **8.2**. $\bar{y}(x) = 3{,}06x - 0{,}37$; $r = 0{,}9995$. **8.3**. $\bar{y}(x) = 5{,}27x - 2{,}38$; $r = 0{,}9998$. **8.4**. $(y_x - 8{,}489) = 1{,}261(x - 5{,}574)$; $x_y - 5{,}574$ =0,356$(y_x - 8{,}489)$; r =0,670. **8.5.** $y_x - 5{,}333 = 0{,}115(x - 26{,}222)$; $x_y - 26{,}222 = 0{,}893(y_x - 5{,}333)$; $r = 0{,}648$.
8.6. $\bar{y}(x) = 1{,}15x + 4{,}09$; $r = 0{,}9978$. **8.7**. $\bar{y}(x) = 1{,}46x - 1{,}23$; $r = 0{,}9991$. **8.8.** $y_x - 7{,}622 = 0{,}275(x - 13{,}649)$; $x_y - 13{,}649 =$
$= 1{,}958(y_x - 5{,}333)$; $r = 0{,}733$.

9.1. $P_{20}(10)$ =0,030817, $P(10 \leq X \leq 15)$ =0,7453. **9.2**. $\bar{x}$ =121,035;

D_x= 0,8351; $s_x^2 = 0{,}7307$; $\sigma_x =$ 0,9138; s_x=0,8548. **9.3**. $\bar{x}$=4,44; D_x= 1,6864; s_x^2 =1,6527; σ_x =1,2986. $(5 \leq X \leq 20)$ =0,2858.

9.4.

x_i	5	10	15	20	25
$f(x_i)$	0,0054	0,0130	0,0242	0,0352	0,0399

$P(5 \leq X \leq 20)$ =0,2858.

9.5. r_b =0,9988. **9.6**.а) Уравнение линейного тренда y = 0,8686x + 2,5267; б) $y = 0{,}2003x^3 - 2{,}4352x^2 + 5{,}043x + 14{,}546$ – уравнение тренда третьей степени; $y = 0{,}2398x^5 - 5{,}7422x^4 + 51{,}966x^3 - 219{,}66x^2 + 421{,}3x - 270{,}95$ – уравнение тренда пятой степени. **9.7**. Доверительный интервал с надежностью $\gamma = 0{,}99$ имеет вид: $[21{,}32;\ 26{,}68\ ]$.

Литература

1. Салимов Р.Б. Математика для инженеров и технологов.– М.: ФИЗМАТЛИТ, 2009. – 484 с.
2. Кремер Н.Ш. Теория вероятностей и математическая статистика. – М.: Юнити , 2001. – 543 с.
3. Марданов Р.Ш., Хасанова А.Ю., Султанов Р.А. Сборник задач по математике для экономистов: Учеб. пособ. – Казань.: КГУ, 2009. – 576 с.
4. Гмурман В.Е. Теория вероятностей и математическая статистика. Т. II. – М.: Наука, 1976. – 672 с.

Содержание

Крепкогорский В.Л.

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие

Редактор В.Н. Сластникова

Издательство
Казанского государственного архитектурно-строительного университета
Подписано в печать 27.03.18 Формат 60×84/16
Заказ № 56 Печать ризографическая Усл. печ. л. 5,0
Тираж 60 экз. Бумага офсетная № 1 Уч.-изд. л. 5,0

Отпечатано в полиграфическом секторе
Издательства КГАСУ.
420043, г. Казань, ул. Зеленая, д. 1.