

Кафедра высшей математики

Курносенко Н.М.

Руководство для практических занятий по
курсу «Теория вероятностей и
математическая статистика»
для студентов заочного факультета
экономических специальностей

1 Случайные события

1.1 Вероятностный эксперимент, случайные события, пространство элементарных событий

1.2 Операции над событиями

1.3 Вероятности случайных событий

1.4 Методы вычисления вероятностей

1.5 Свойства вероятностей случайных событий

1.6 Теоремы сложения и умножения вероятностей

1.7 Формулы полной вероятности и Байеса

1.8 Последовательность независимых испытаний

1.1 Вероятностный эксперимент, случайные события, пространство элементарных событий

Любой эксперимент или наблюдение изучаемого физического явления заканчивается некоторым событием (исходом). Если результат эксперимента заранее однозначно непредсказуем, то данный эксперимент называется вероятностным и обозначается символом « E ».

Элементарным событием (элементарным исходом) ω называется любой мысленно возможный неразложимый результат вероятностного эксперимента E .

Пространством элементарных событий Ω называется множество всех мыслимых взаимоисключающих результатов вероятностного эксперимента E .

Случайным событием называется такое событие, о котором нельзя заведомо точно сказать, произойдёт оно или нет.

Случайные события обозначаются заглавными латинскими буквами ($A, B, C, D, \dots$). Случайное событие является некоторым подмножеством пространства элементарных событий ($A \subseteq \Omega$).

Пример 1

E : бросается игральная кость.

Элементарные события: $\omega_1 = \{\text{выпадение на игральной кости «1»}\}$, $\omega_2 = \{\text{выпадение на игральной кости «2»}\}$ и т. д. Пространство элементарных событий $\Omega = \{\text{выпадение на игральной кости числа от «1» до «6»}\} = \{\omega_1, \omega_2, \dots, \omega_6\}$. Тогда случайные события:

$A = \{\text{выпадение чётного числа}\} = \{\omega_2, \omega_4, \omega_6\}$;

$B = \{\text{выпадение нечётного числа}\} = \{\omega_1, \omega_3, \omega_5\}$;

$$C = \{\text{выпадение «5»}\} = \{\omega_5\};$$

$$D = \{\text{невыпадение «3»}\} = \{\omega_1, \omega_2, \omega_4, \omega_5, \omega_6\};$$

$$F = \{\text{выпадение числа от «3» до «5»}\} = \{\omega_3, \omega_4, \omega_5\};$$

$$G = \{\text{выпадение числа } > 4\} = \{\omega_5, \omega_6\};$$

$$I = \{\text{выпадение числа } < 4\} = \{\omega_1, \omega_2, \omega_3\}.$$

В зависимости от размерности множества возможных элементарных событий, различают конечное, счётное и несчётное пространство элементарных событий Ω .

В примере 1 пространство элементарных событий Ω является конечным, поскольку включает лишь 6 элементарных событий. В эксперименте с исследованием числа поездов, прибывающих на станцию в течение суток, пространство элементарных событий Ω счётно, т.к. каждому элементарному событию эксперимента можно поставить в однозначное соответствие число натурального ряда. В эксперименте с исследованием времени обслуживания поезда на станции, пространство элементарных событий Ω несчётно, т. к. время обслуживания может принимать любые положительные значения.

Элементарные события, которые образуют случайное событие A , называются благоприятными событию A .

В примере 1 элементарные события $\omega_2, \omega_4, \omega_6$ являются благоприятными событию A , элементарные события $\omega_1, \omega_3, \omega_5$ — благоприятными событию B и т. д.

В частном случае множество элементарных исходов, благоприятных событию A , может совпадать с пространством элементарных событий Ω или быть пустым множеством $\emptyset$.

Достоверным событием называется событие, которое всегда происходит, т. е. совпадающее с пространством элементарных событий Ω .

Невозможным событием называется событие, которое никогда не произойдёт, т. е. совпадающее с пустым множеством.

В примере 1 достоверным событием является случайное событие $K = \{\text{выпадение числа от «1» до «6»}\} = \Omega$, а невозможным событием является, например, случайное событие $L = \{\text{выпадение числа «7»}\} = \emptyset$.

1.2 Операции над событиями

Пусть имеется пространство элементарных событий Ω . Будем рассматривать в качестве случайных событий подмножества $A, B, C, \dots$ этого пространства.

Суммой (объединением) событий A и B называется третье событие $A+B$ ($A \cup B$), состоящее в осуществлении хотя бы одного из событий A или B . Благоприятными событию $A \cup B$ являются все элементарные события, благоприятные хотя бы одному из событий A или B .

Аналогично определяется сумма любого числа событий $A_1 \cup A_2 \cup A_3 \dots$.

Произведением (пересечением) событий A и B называется третье событие AB ($A \cap B$), состоящее в одновременном осуществлении событий A и B . Благоприятными событию $A \cap B$ являются все элементарные события, благоприятные одновременно событию A и событию B .

Произведение любого числа событий $A_1 \cap A_2 \cap A_3 \dots$ состоит в одновременном осуществлении событий A_1, A_2, A_3 и т. д.

Разностью событий A и B называется третье событие $A-B$ ($A \setminus B$), состоящее в осуществлении события A без осуществления события B . Событие $A \setminus B$ состоит из элементарных событий благоприятных событию A , за исключением элементарных событий благоприятных событию B .

Противоположным событию A называется событие $\bar{A}$, состоящее в ненаступлении события A . Событию $\bar{A}$ благоприятны все возможные элементарные события пространства элементарных событий Ω , кроме тех, которые благоприятны событию A ($\bar{A} = \Omega \setminus A$).

События A и B называются несовместными, если они не могут произойти одновременно, т. е. одновременное осуществление событий A и B есть событие невозможное ($A \cap B = \emptyset$).

События $A_1, A_2, A_3, \dots, A_n$ образуют полную группу событий, если их сумма составляет пространство элементарных событий Ω ($A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n = \Omega$), т. е. в результате эксперимента хотя бы одно из событий произойдет.

Пример 2

Рассмотрим операции над событиями, используя условия из примера 1.

а) $A \cup B = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} = \Omega$;

$$A \cup C = \{\omega_2, \omega_4, \omega_5, \omega_6\};$$

$$C \cup I = \{\omega_1, \omega_2, \omega_3, \omega_5\};$$

$$C \cup D = \{\omega_1, \omega_2, \omega_4, \omega_5, \omega_6\};$$

$$F \cup G = \{\omega_3, \omega_4, \omega_5, \omega_6\};$$

б) $A \cap B = \emptyset$;

$$B \cap C = \{\omega_5\};$$

$$D \cap F = \{\omega_5, \omega_6\};$$

$$G \cap I = \emptyset;$$

$$C \cap G = \{\omega_5\};$$

$$\text{в) } A \setminus B = \emptyset;$$

$$B \setminus C = \{\omega_1, \omega_3\};$$

$$D \setminus F = \{\omega_1, \omega_2, \omega_6\};$$

$$G \setminus I = \{\omega_5, \omega_6\} = G;$$

$$G \setminus C = \{\omega_6\};$$

$$\text{г) } \bar{A} = \Omega \setminus A = \{\omega_1, \omega_3, \omega_5\} = B;$$

$$\bar{B} = \Omega \setminus B = \{\omega_2, \omega_4, \omega_6\} = A;$$

$$\bar{C} = \Omega \setminus C = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_6\};$$

$$\bar{D} = \Omega \setminus D = \{\omega_3\};$$

$$\bar{F} = \Omega \setminus F = \{\omega_1, \omega_2, \omega_6\};$$

д) События A и B образуют полную группу событий, т. к. $A \cup B = \Omega$.

Пример 3

E : в прямоугольник (рисунок 1) наудачу бросается точка.

Элементарное событие данного эксперимента – некоторая точка внутри прямоугольника. Пространство элементарных событий Ω (в данном случае – несчётное) – всё множество точек внутри прямоугольника. На множестве Ω определены два события: $A = \{\text{выбранная точка лежит внутри круга } A\}$ и $B = \{\text{выбранная точка лежит внутри круга } B\}$.

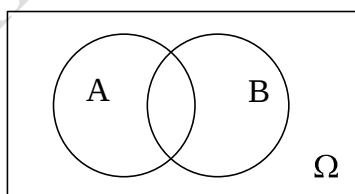
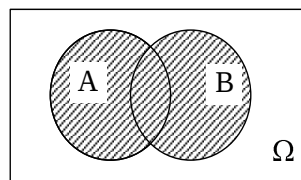
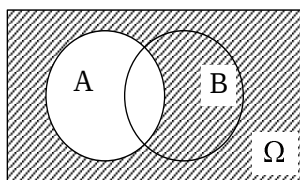
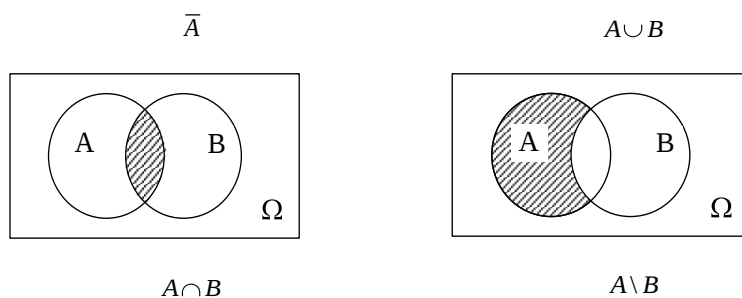


Рисунок 1 – Диаграмма Венна-Эйлера

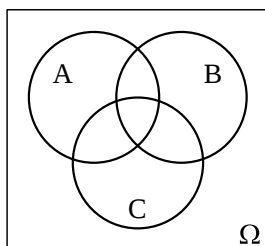
Изобразим области, попадание в которые соответствует осуществлению событий $\bar{A}$, $A \cup B$, $A \cap B$, $A \setminus B$:



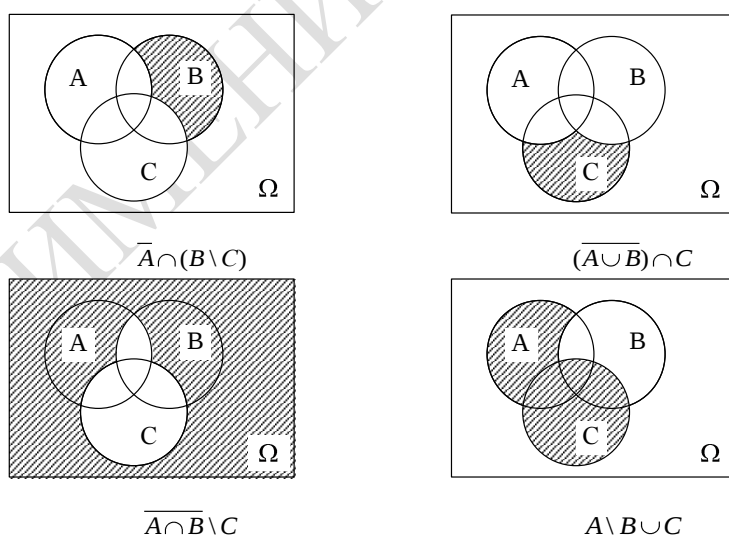


Пример 4

E : в прямоугольник наудачу бросается точка.



На множестве Ω определены три события: $A = \{\text{выбранная точка лежит внутри круга } A\}$, $B = \{\text{выбранная точка лежит внутри круга } B\}$, $C = \{\text{выбранная точка лежит внутри круга } C\}$. Изобразим области, попадание в которые соответствует осуществлению следующих событий $\bar{A} \cap (B \setminus C)$, $(A \cup B) \cap C$, $\overline{A \cap B \setminus C}$, $A \setminus B \cup C$:



1.3 Вероятности случайных событий

1.3.1 Относительная частота случайных событий

Пусть было проведено n вероятностных экспериментов E , при этом случайное событие A произошло m раз.

Число m называется *частотой* появления случайного события A , а

отношение $P^* = \frac{m}{n}$ – *относительной частотой (частотью)* случайного события A .

Относительная частота наступления некоторого случайного события не является постоянной величиной, однако она обладает устойчивостью, стремлением к некоторому постоянному числу и колебания её тем меньше, чем больше проведено экспериментов.

1.3.2 Понятие вероятности случайного события. Аксиомы Колмогорова

Вероятностью случайного события A называется числовая функция $P(A)$, определённая на пространстве элементарных событий Ω , характеризующая меру объективной возможности наступления события A и удовлетворяющая для каждого случайного события аксиомам Колмогорова А. Н.

Аксиома 1. $P(A) \geq 0$, т. е. вероятность наступления произвольного случайного события – неотрицательная функция.

Аксиома 2. $P(\Omega) = 1$, т. е. вероятность наступления достоверного события равна 1.

Аксиома 3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$, $A_i \cap A_j = \emptyset$, $\forall i \neq j$, т. е. вероятность наступления суммы счётного множества попарно несовместных событий A_i , $i = \overline{1, \infty}$ равна сумме вероятностей этих событий.

1.4 Методы вычисления вероятностей

1.4.1 Классический метод вычисления вероятностей

Пусть пространство элементарных событий Ω некоторого вероятностного эксперимента E конечно $(\omega_1, \omega_2, \dots, \omega_n)$ и все элементарные события равновозможны, т. е. $P(\omega_1) = P(\omega_2) = \dots = P(\omega_n)$.

По *классическому (лапласовскому) методу* вероятность случайного события A равна отношению числа элементарных событий N_A , благоприятных событию A , к общему количеству элементарных событий N пространства элементарных событий Ω , т. е. $P(A) = \frac{N_A}{N}$.

Учитывая, что A и Ω – множества (элементарных событий), можно записать:

$$P(A) = \frac{|A|}{|\Omega|}, \quad (1.4.1.1)$$

где $|A|$ – количество элементарных событий, благоприятных событию A ; $|\Omega|$ – общее количество элементарных событий пространства элементарных событий Ω .

Классический метод вычисления вероятностей имеет следующие ограничения:

а) все элементарные события вероятностного эксперимента E должны быть равновероятными, т. е. $P(\omega_i) = P(\omega_j)$, $\forall i, j$;

б) множество элементарных событий пространства Ω должно быть конечным, чтобы отношение $|A|/|\Omega|$ не являлось неопределённостью ∞/∞ .

Пример 5

E : бросается игральная кость.

Найдём вероятности случайных событий из примера 1.

$$P(A) = \frac{1}{2}; P(B) = \frac{1}{2}; P(C) = \frac{1}{6}; P(D) = \frac{5}{6}; P(F) = \frac{1}{2}; P(G) = \frac{1}{3};$$

$$P(I) = \frac{1}{2}.$$

Пример 6

E : разгрузка вагонов на сортировочной станции.

На сортировочную станцию прибывают вагоны из Минска, Гомеля и Бреста. Предполагая равновероятными все варианты очередности разгрузки этих трёх вагонов, найти вероятности следующих случайных событий:

$A = \{\text{вагон из Гомеля будет разгружен первым}\};$

$B = \{\text{вагон из Бреста будет разгружен не ранее, чем вагон из Минска}\}.$

Решение. Пространство элементарных событий Ω в данном эксперименте состоит из шести элементарных событий $\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. Введём условные обозначения элементарных событий по первым буквам названий городов $\Omega = \{\text{ГМБ, ГБМ, МГБ, МБГ, БГМ, БМГ}\}$, где, например, элементарное событие МГБ соответствует такой последовательности разгрузки: из Минска – из Гомеля – из Бреста. Тогда

$$A = \{\text{ГМБ, ГБМ}\}, P(A) = \frac{2}{6} = \frac{1}{3},$$

$$B = \{\text{ГМБ, МГБ, МБГ}\}, P(B) = \frac{3}{6} = \frac{1}{2}.$$

1.4.2 Элементы комбинаторики

Комбинаторика – раздел математики, изучающий количество ком-

бинаций, подчинённых некоторым условиям.

Лемма 1 (основная лемма комбинаторики)

Из m элементов первого множества $\{a_1, a_2, \dots, a_m\}$ и n элементов второго множества $\{b_1, b_2, \dots, b_n\}$ можно составить ровно mn различных упорядоченных пар (a_i, b_j) , содержащих по одному элементу из каждого множества.

Пример 7

Рассмотрим две группы элементов: ♠ – пики, ♣ – трефы, ♥ – черви, ♦ – бубны и 6, 7, 8, 9, 10, валет, дама, король, туз. По лемме 1 число пар $4 \cdot 9 = 36$. Это число равно числу карт в колоде, т. к. каждая карта определяется парой элементов (масть и значение).

Пример 8

На «горном» велосипеде 3 передние и 6 задних звездочек. Сколько скоростей у «горного» велосипеда?

Решение. Так как каждая скорость велосипеда – комбинация одной из 3 передних ($n_1 = 3$) и одной из 6 задних ($n_2 = 6$) звездочек, то количество скоростей на велосипеде равно количеству комбинаций звездочек двух типов и определяется, в соответствии с леммой 1, произведением $n_1 n_2 = 3 \cdot 6 = 18$ скоростей.

Пример 9

Е: бросаются две игральные кости.

Определим элементарное событие как пару $\omega = (i, j)$, где i – число очков, выпавших на первой кости, j – число очков, выпавших на второй кости. Тогда i выбирается из группы 1, 2, 3, 4, 5, 6; j выбирается из этой же группы. По лемме 1 число всех элементарных событий (т. е. всевозможных пар (i, j)) $6 \cdot 6 = 6^2 = 36$.

Лемма 2

Из n_1 элементов первого множества $\{a_1, a_2, \dots, a_{n_1}\}$, n_2 элементов второго множества $\{b_1, b_2, \dots, b_{n_2}\}$ и т. д., n_k элементов k -го множества $\{x_1, x_2, \dots, x_{n_k}\}$ можно составить ровно $n_1 \cdot n_2 \cdot \dots \cdot n_k$ различных упорядоченных комбинаций $(a_i, b_j, \dots, x_s)$, содержащих по одному элементу из каждого множества.

Пример 10

Е: бросаются три игральные кости.

Элементарное событие $\omega = (i, j, k)$, где i – число очков, выпавших на первой кости, j – на второй кости, k – на третьей кости. По лемме 2 число всех элементарных событий (т. е. всевозможных комбинаций (i, j, k)) будет $6 \cdot 6 \cdot 6 = 6^3 = 216$.

Пример 11

Из пункта A в пункт B проходит 10 дорог, из пункта B в пункт C – 5 дорог, из пункта C в пункт D – 6 дорог. При этом все дороги, ведущие из A в D , проходят сначала через B , а затем через C . По лемме 2 из пункта A в пункт D проходит $10 \cdot 5 \cdot 6 = 300$ дорог.

Перестановками называются комбинации n различных элементов и отличающиеся только порядком их расположения. Количество возможных перестановок n различных элементов обозначается $P_n = n!$

Упорядоченными выборками (размещениями) называются комбинации, составленные из n различных элементов по m элементов, различающиеся либо составом элементов, либо их порядком. Количество возможных размещений m элементов из n различных элементов обозначается A_n^m .

Неупорядоченными выборками (сочетаниями) называются комбинации, составленные из n различных элементов по m элементов, различающиеся только составом элементов. Количество возможных сочетаний m элементов из n различных элементов обозначается C_n^m .

Упорядоченные и неупорядоченные выборки, элементы которых могут повторяться, называются соответственно *упорядоченными и неупорядоченными выборками с повторением*. Количество возможных упорядоченных и неупорядоченных выборок m элементов из n различных элементов с повторением обозначается соответственно $\tilde{A}_n^m$ и $\tilde{C}_n^m$.

Таблица 1 – Числа выборок объёма m из множества $\{a_1, a_2, \dots, a_n\}$

Выборки	Упорядоченные (размещения)	Неупорядоченные (сочетания)
С повторением (с возвращением)	$\tilde{A}_n^m = n^m$	$\tilde{C}_n^m = C_{n+m-1}^m$
Без повторения (без возвращения)	$A_n^m = \frac{n!}{(n-m)!}$	$C_n^m = \frac{n!}{(n-m)!m!}$

Пример 12

Сколькими способами можно расположить три шара, пронумерованных цифрами «1», «2», «3»?

Решение. Поскольку комбинации расположения трёх различных шаров отличаются лишь порядком расположения, то данные комбинации называются перестановками. Перечислим все возможные способы перестановками трёх шаров: «1-2-3», «1-3-2», «2-1-3», «2-3-1», «3-1-2», «3-2-1». Таким образом, количество всевозможных перестановок равно 6.

Пример 13

Перечислить все возможные способы выбора двух шаров из урны с тремя шарами, пронумерованными числами «1», «2», «3».

Таблица 2 – Способы выбора 2 шаров из урны с 3 пронумерованными шарами

Выборки	Упорядоченные (размещения)	Неупорядоченные (сочетания)
С повторением (с возвращением)	(1,1) (1,2) (1,3) (2,1) (2,2) (2,3) (3,1) (3,2) (3,3)	(1,1) (1,2) (1,3) (2,2) (2,3) (3,3)
Без повторения (без возвращения)	(1,2) (1,3) (2,1) (2,3) (3,1) (3,2)	(1,2) (1,3) (2,3)

Пример 14

На железнодорожной станции имеются 10 путей. Сколькими способами можно расставить на них три состава?

Решение. Поскольку комбинации расположения трёх различных составов на 10 путях отличаются лишь расположениями, то данные комбинации являются упорядоченными выборками без возвращения. Количество всевозможных размещений в этом случае равно $A_n^m = A_{10}^3 = 720$.

Пример 15

В вагон электрички, делающей 9 остановок, на первой остановке вошли два пассажира. Каждый из них, с одинаковой вероятностью, выходит на любой из остановок, начиная со второй. Найти вероятность того, что оба пассажира выйдут на одной остановке.

Решение. Пусть случайное событие $A = \{\text{два пассажира выйдут на}\}$

одной остановке}. По классическому методу $P(A) = \frac{N_A}{N}$, где $N_A = 8$ – число элементарных событий, благоприятных событию A , а $N = \tilde{A}_n^m = n^m = 8^2 = 64$ – число всевозможных элементарных событий пространства Ω . Таким образом, $P(A) = \frac{8}{64} = \frac{1}{8} = 0,125$.

1.4.3 Геометрический метод вычисления вероятностей

Если пространство элементарных событий Ω вероятностного эксперимента E является несчетным, то для вычисления вероятностей случайных событий может применяться геометрический метод.

Пусть пространство Ω эксперимента E содержит несчетное множество элементарных исходов ω (т. е. $|\Omega| = \infty$) и их можно трактовать как точки в евклидовом пространстве, а события эксперимента E – как некоторые ограниченные области этого пространства. Тогда вероятность случайного события A может быть определена выражением

$$P(A) = \frac{\mu(A)}{\mu(\Omega)}, \quad (1.4.3.1)$$

где $\mu(A)$ – геометрическая мера (длина, площадь, объём) области, соответствующая событию A ; $\mu(\Omega)$ – геометрическая мера области, соответствующая пространству элементарных событий Ω .

Пример 16

Простой состава в ожидании осмотра бригадой пункта технического осмотра (ПТО) в парке прибытия сортировочной станции равно возможен в интервале $[0; 50 \text{ мин}]$. Простой в ожидании расформирования состава также равновозможен в интервале $[0; 40 \text{ мин}]$.

Найти вероятность того, что нерегламентированный простой состава в парке прибытия не превысит 20 минут.

Решение. Пространством элементарных событий Ω этого эксперимента является прямоугольник, изображённый на рисунке 5, т. е.

$$\Omega = \{\omega = (x, y) \mid 0 \leq x \leq 50, 0 \leq y \leq 40\}.$$

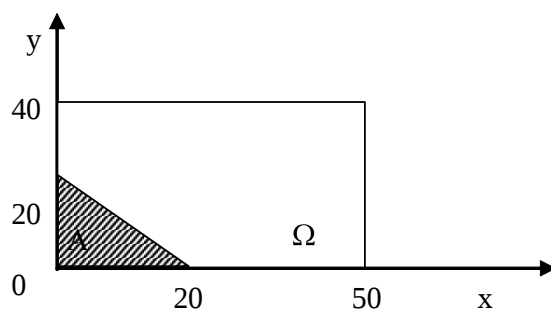


Рисунок 2 – Геометрическая интерпретация вероятностного эксперимента

Событию $A = \{\omega = (x, y) \mid x + y \leq 20\} = \{\text{простой состава не превысит 20 минут}\}$ соответствует заштрихованная на рисунке область.

Следовательно,

$$P(A) = \frac{S_A}{S_\Omega} = \frac{0,5 \cdot 20 \cdot 20}{40 \cdot 50} = 0,1.$$

Пример 17

Два теплохода должны подойти к одному причалу. Моменты прихода обоих теплоходов независимы и равновозможны в течение суток. Определить вероятность того, что одному из теплоходов придётся ожидать на рейде, пока не освободится причал, если время стоянки теплоходов равно одному часу.

Решение. Пространством элементарных событий Ω этого эксперимента является прямоугольник, изображённый на рисунке 3, т. е.

$$\Omega = \{\omega = (x, y) \mid 0 \leq x \leq 24, 0 \leq y \leq 24\}.$$

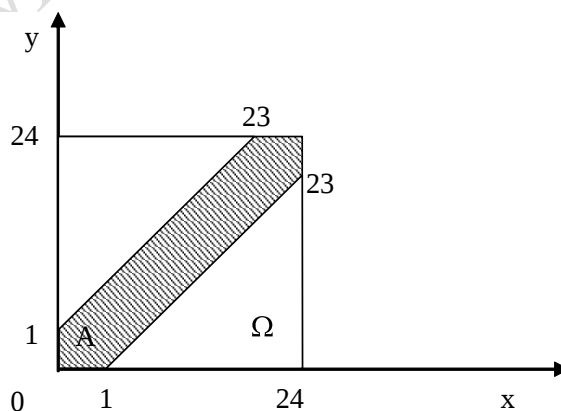


Рисунок 3 – Геометрическая интерпретация вероятностного эксперимента

Событию $A = \{\omega = (x, y) \mid |x - y| \leq 1\} = \{\text{одному из теплоходов придётся ожидать на рейде}\}$ соответствует заштрихованная на рисунке область.

Следовательно,

$$P(A) = \frac{S_A}{S_\Omega} = \frac{24^2 - 23^2}{24^2} = \frac{47}{576} \approx 0,082.$$

1.4.4 Статистический и экспертный методы вычисления вероятностей

Кроме классического и геометрического существуют еще два способа определения вероятностей случайных событий: *статистический* и *экспертный*. Статистический способ заключается в оценке вероятности случайного события по результатам многократного воспроизведения вероятностного эксперимента E , например, по относительной частоте появления случайного события.

Метод экспертных оценок заключается в опросе мнения некоторого количества экспертов о значении вероятности случайного события. Анализируя полученные значения экспертных оценок, можно получить представление о реальном значении вероятности исследуемого случайного события.

Статистический и экспертный способы оценки вероятностей являются универсальными, однако, предоставляемый с их помощью результат не является точным. Для увеличения достоверности и точности оценки вероятности требуется проведение большего количества повторных экспериментов и привлечение большего числа опытных экспертов.

1.5 Свойства вероятностей случайных событий

Вероятности случайных событий обладают следующими важными свойствами:

Свойство 1. $P(\emptyset) = 0$, т. е. вероятность невозможного события равна 0.

Свойство 2. Если в пространстве Ω , содержащем конечное или счётное множество возможных элементарных событий $\omega_1, \omega_2, \dots, \omega_i, \dots$ ($\Omega = \{\omega_1, \omega_2, \dots, \omega_i, \dots\}$), заданы вероятности элементарных событий $P(\omega_1) = p_1, P(\omega_2) = p_2, \dots, P(\omega_i) = p_i, \dots$, то вероятность произвольного события $A = \{\omega_j, \omega_k, \dots, \omega_l\}$ равна сумме вероятностей элементарных событий, благоприятных событию A , т. е. $P(A) = \sum_{\omega_i \in A} P(\omega_i)$.

Говорят, что событие A влечёт событие B ($A \subseteq B$), если все элемен-

тарные события ω_i , благоприятные событию A , благоприятны событию B (рисунок 4).

Свойство 3. Если $A \subseteq B$, то $P(B \setminus A) = P(B) - P(A)$.

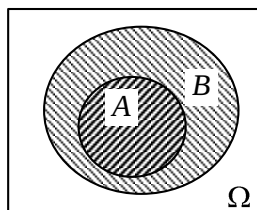


Рисунок 4 – Пример события A , которое влечёт событие B

Свойство 4 (следствие свойства 3). Если $A \subseteq B$, то $P(B) \geq P(A)$.

Свойство 5. $P(\bar{A}) = 1 - P(A)$, т.е. сумма вероятностей противоположных событий равна 1 (рисунок 5).

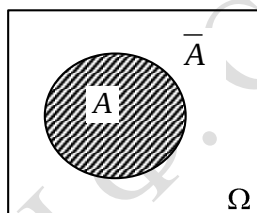


Рисунок 5 – Пример противоположных событий A и $\bar{A}$

Свойство 6. $0 \leq P(A) \leq 1$, т. е. вероятность произвольного случайного события принадлежит отрезку $[0, 1]$.

Свойство 7. Вероятность произведения двух несовместных случайных событий равна нулю. То есть, если события A и B несовместны, то $P(A \cap B) = 0$.

Несовместные события $H_1, H_2, \dots, H_n$, которые образуют полную группу событий, называются *гипотезами*.

Свойство 8. Сумма вероятностей гипотез $H_1, H_2, \dots, H_n$ равна единице.

Гипотезы, вероятности которых равны, называются *шансами*.

1.6 Теоремы сложения и умножения вероятностей

1.6.1 Теорема сложения вероятностей

Теорема сложения вероятностей двух событий: Пусть A и B – произвольные случайные события, принадлежащие пространству Ω , тогда вероятность суммы этих событий равна сумме вероятностей этих событий без вероятности их совместного появления (рисунок 6):

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad (1.6.1.1)$$

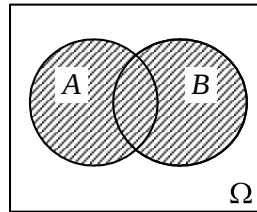


Рисунок 6 – Сумма случайных событий A и B

Следствие. Вероятность суммы двух несовместных событий A и B равна сумме вероятностей этих событий: $P(A \cup B) = P(A) + P(B)$, т. к. вероятность произведения несовместных событий равна нулю по свойству 7.

Теорема сложения вероятностей трёх произвольных событий:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C). \quad (1.6.1.2)$$

Пример 18

В подаче вагонов на контейнерную площадку могут находиться четырёхосная платформа с вероятностью 0,35, четырёхосный полувагон с вероятностью 0,5 и шестиосный полувагон с вероятностью 0,15. Найти вероятность того, что выбранный наудачу вагон окажется четырёхосным.

Решение. Пространство элементарных событий Ω этого эксперимента определяется следующим образом:

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$

где ω_1 – выбранный вагон – четырёхосная платформа; ω_2 – выбранный вагон – четырёхосный полувагон; ω_3 – выбранный вагон – шестиосный полувагон.

Событие $A = \{\text{выбран четырёхосный вагон}\} = \{\omega_1, \omega_2\}$ и, следовательно,

$$P(A) = P(\omega_1) + P(\omega_2) = 0,35 + 0,5 = 0,85.$$

1.6.2 Условная вероятность; независимость событий

Рассмотрим следующий вероятностный эксперимент E . Пусть в пространстве Ω определены случайные события $A, B, C, \dots$ и их вероятности. Предположим, что в ходе эксперимента E событие A уже произошло. Получение дополнительной информации о ходе эксперимента E может привести к желанию пересмотреть вероятности других событий пространства Ω , связанных с событием A . Ведь логично предположить, что появление события A каким-то образом может изменить вероятность появления событий, связанных (зависимых) с ним.

Событие B называется *зависимым* от события A , если появление (или непоявление) события A изменяет вероятность появления события B . Если наступление события A не изменяет вероятности появления B , событие B называется *независимым* от события A .

Условной вероятностью $P(B|A)$ (или $P_A(B)$) называют вероятность наступления события B при условии, что событие A уже произошло.

1.6.3 Теорема умножения вероятностей

Теорема умножения вероятностей двух произвольных событий. Вероятность произведения двух произвольных событий равна произведению вероятности одного из событий на условную вероятность другого события при условии, что первое уже произошло:

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B). \quad (1.6.3.1)$$

Теорема умножения вероятностей трёх произвольных событий:

$$P(A \cap B \cap C) = P(A)P(B|A)P(C|A \cap B). \quad (1.6.3.2)$$

Следствие 1. Если событие A не зависит от B , то и событие B не зависит от A .

Следствие 2. Вероятность произведения двух независимых событий равна произведению вероятностей этих событий. То есть, если события A и B независимы, то

$$P(A \cap B) = P(A)P(B). \quad (1.6.3.3)$$

Случайные события называются *независимыми в совокупности*, если вероятность наступления каждого из них не изменяется с наступлением любой комбинации остальных событий. Для случайных событий $A_1, A_2, \dots, A_n$, независимых в совокупности, справедлива следующая теорема умножения вероятностей (необходимое условие независимости в совокупности n случайных событий):

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2)\dots P(A_n). \quad (1.6.3.4)$$

Замечание – Попарная независимость случайных событий не означает их независимость в совокупности.

Пример 19

Вероятность появления в поезде вагонов на контейнерную площадку – 0,1, на грузовой двор – 0,3, на промышленное предприятие – 0,4. Определить вероятность появления в поезде вагонов на все три направления.

Решение. Обозначим событие $A = \{\text{в поезде появятся вагоны на все три направления}\}$.

$A = A_1 A_2 A_3$, где $A_1 = \{\text{появление в поезде вагона на контейнерную площадку}\}$, $A_2 = \{\text{появление в поезде вагона на грузовой двор}\}$, $A_3 = \{\text{появление в поезде вагона на промышленное предприятие}\}$. Поскольку события A_1 , A_2 и A_3 независимы (вагоны появляются в поезде независимо друг от друга),

$$P(A) = P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3) = 0,1 \cdot 0,3 \cdot 0,4 = 0,012.$$

Пример 20

На пути движения локомотива три светофора. Каждый из них либо разрешает, либо запрещает дальнейшее движение локомотива с вероятностью 0,5. Какова вероятность того, что локомотив сделает три остановки?

Решение. Обозначим событие $A = \{\text{локомотив сделает три остановки}\}$.

$A = A_1 A_2 A_3$, где $A_1 = \{\text{локомотив сделает остановку на первом светофоре}\}$, $A_2 = \{\text{локомотив сделает остановку на втором светофоре}\}$, $A_3 = \{\text{локомотив сделает остановку на третьем светофоре}\}$. Поскольку события A_1 , A_2 и A_3 независимы (остановки локомотива на светофорах не зависят друг от друга),

$$P(A) = P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3) = 0,5 \cdot 0,5 \cdot 0,5 = 0,125.$$

Пример 21

Вероятность прибытия поезда на станцию без опоздания равна 0,95. Найти вероятность того, что четыре последовательно прибывших на станцию поезда опоздали.

Решение. Обозначим событие $A = \{\text{четыре прибывших на станцию поезда опоздали}\}$.

$A = A_1 A_2 A_3 A_4$, где $A_1 = \{\text{опоздает поезд, прибывший первым}\}$, $A_2 = \{\text{опоздает поезд, прибывший вторым}\}$, $A_3 = \{\text{опоздает поезд, прибывший третьим}\}$, $A_4 = \{\text{опоздает поезд, прибывший четвёртым}\}$. Поскольку события A_1, A_2, A_3 и A_4 независимы (поезда опаздывают независимо друг от друга),

$$\begin{aligned} P(A) &= P(A_1 A_2 A_3 A_4) = P(A_1)P(A_2)P(A_3)P(A_4) = \\ &= 0,05 \cdot 0,05 \cdot 0,05 \cdot 0,05 = 6,25 \cdot 10^{-6}. \end{aligned}$$

1.7 Формулы полной вероятности и Байеса

1.7.1 Формула полной вероятности

Формула полной вероятности является следствием теорем сложения и умножения вероятностей.

Пусть требуется определить вероятность некоторого случайного события A , которое может произойти только с одной из гипотез $H_1, H_2, \dots, H_n$. Тогда вероятность указанного события можно вычислить по формуле

$$P(A) = \sum_{i=1}^n (P(H_i)P(A|H_i)), \quad (1.7.1.1)$$

т.е. как сумму произведений вероятности каждой гипотезы на условную вероятность события A при условии гипотезы. Формула Байеса для пересчета вероятностей гипотез, сопутствующих (предшествующих) некоторому случайному событию A , о котором стало известно, что оно произошло.

Пусть некоторое случайное событие A может произойти только с одной из гипотез $H_1, H_2, \dots, H_n$. Причем известны *априорные* (доопытные) вероятности этих гипотез $P(H_i)$, $i = \overline{1, n}$. Пусть известно, что событие A произошло. Требуется найти *апостериорные* (послеопытные) вероятности гипотез H_i , $i = \overline{1, n}$, т. е. пересчитать вероятности гипотез, сопутствующих (предшествующих) случайному событию A при наличии дополнительной информации о нем. В данном случае для вычисления апостериорных вероятностей гипотез используется формула

Байеса:

$$P(H_i | A) = \frac{P(H_i)P(A | H_i)}{P(A)} = \frac{P(H_i)P(A | H_i)}{\sum_{i=1}^n P(H_i)P(A | H_i)}, \quad (1.7.1.2)$$

где $P(H_i | A)$ – апостериорная (послеопытная) вероятность гипотезы H_i при условии, что событие A произошло; $P(H_i)$ – априорная (доопытная) вероятность гипотезы H_i , $i = \overline{1, n}$; $P(A | H_i)$ – условная вероятность события A при условии справедливости гипотезы H_i ; $P(A) > 0$ – безусловная вероятность случайного события A , определяемая по формуле полной вероятности (1.7.1.1).

Пример 22

Из депо прописки вагон, нуждающийся в ремонте, направлен в одно из трёх ремонтных депо. Производительности этих депо соотносятся как 6:5:4. Вероятности бездефектного ремонта вагонов для первого, второго и третьего депо соответственно равны 0,9, 0,95 и 0,85.

а) Найти вероятность того, что направленный на ремонт из депо прописки вагон будет отремонтирован без дефектов.

б) Известно, что направленный на ремонт из депо прописки вагон был отремонтирован без дефектов. Найти вероятность того, что он подвергался ремонту во втором депо.

Решение. Относительно условий рассматриваемого случайного эксперимента, состоящего в направлении неисправного вагона в одно из ремонтных депо, можно выдвинуть три несовместные гипотезы:

$H_1 = \{\text{вагон ремонтировался в первом депо}\};$

$H_2 = \{\text{вагон ремонтировался во втором депо}\};$

$H_3 = \{\text{вагон ремонтировался в третьем депо}\}.$

Причём $H_1 + H_2 + H_3 = \Omega$.

Согласно условию $P(H_1):P(H_2):P(H_3) = 6:5:4$.

Учитывая свойство вероятностей гипотез $P(H_1) + P(H_2) + P(H_3) = \Omega$, определим:

$$P(H_1) = \frac{6}{15}; P(H_2) = \frac{5}{15}; P(H_3) = \frac{4}{15}.$$

Условные вероятности события $A = \{\text{вагон отремонтирован без дефектов}\}$ при осуществлении этих гипотез известны:

$$P(A | H_1) = 0,9; P(A | H_2) = 0,95; P(A | H_3) = 0,85.$$

а) Для определения вероятности события A воспользуемся формулой полной вероятности

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) = \\ = \frac{6}{15} \cdot 0,9 + \frac{5}{15} \cdot 0,95 + \frac{4}{15} \cdot 0,85 \approx 0,903.$$

б) Для определения вероятности того, что вагон подвергался ремонту во втором депо, при условии, что он был отремонтирован без дефектов, воспользуемся формулой Байеса

$$P(H_2|A) = \frac{P(H_2)P(A|H_2)}{\sum_{i=1}^3 P(H_i)P(A|H_i)} = \frac{\frac{5}{15} \cdot 0,95}{0,903} = 0,351.$$

Ответ: а) вероятность того, что направленный на ремонт из депо прописки вагон будет отремонтирован без дефектов, равна 0,903;

б) вероятность того, что вагон подвергался ремонту во втором депо, при условии, что он был отремонтирован без дефектов, равна 0,351.

Пример 23

На сортировочную станцию прибывают полувагоны, платформы и крытые вагоны с вероятностями соответственно 0,25, 0,3, 0,45. Вероятность неисправности полувагона равна 0,02, платформы – 0,015, крытого вагона – 0,01.

а) Найти вероятность того, что поступивший на осмотр в парк приёма вагон окажется неисправным.

б) Поступивший на осмотр в парк приёма вагон оказался неисправным. Найти вероятность того, что этот вагон является платформой.

Решение. Относительно условий рассматриваемого случайного эксперимента, состоящего в направлении вагонов на осмотр в парк приёма, можно выдвинуть три несовместные гипотезы:

$H_1 = \{\text{поступивший в парк вагон является полувагоном}\};$

$H_2 = \{\text{поступивший в парк вагон является платформой}\};$

$H_3 = \{\text{поступивший в парк вагон является крытым вагоном}\}.$

Причём $H_1 + H_2 + H_3 = \Omega$.

Согласно условию $P(H_1) = 0,25; P(H_2) = 0,3; P(H_3) = 0,45$.

Условные вероятности события $A = \{\text{поступивший вагон окажется неисправным}\}$ при осуществлении этих гипотез известны:

$$P(A|H_1) = 0,02; P(A|H_2) = 0,015; P(A|H_3) = 0,01.$$

а) Для определения вероятности события A воспользуемся формулой полной вероятности

$$P(A) = P(H_1)P(A|H_1) + P(H_2)P(A|H_2) + P(H_3)P(A|H_3) = 0,25 \cdot 0,02 + 0,3 \cdot 0,015 + 0,45 \cdot 0,01 = 0,014.$$

б) Для определения вероятности того, что поступивший на осмотр вагон является платформой, при условии, что он был неисправным, воспользуемся формулой Байеса

$$P(H_2|A) = \frac{P(H_2)P(A|H_2)}{\sum_{i=1}^3 P(H_i)P(A|H_i)} = \frac{0,3 \cdot 0,015}{0,014} = 0,3214.$$

Ответ: а) вероятность того, что поступивший на осмотр в парк приёма вагон окажется неисправным, равна 0,014;

б) вероятность того, что поступивший на осмотр в парк приёма вагон оказался неисправным, при условии, что этот вагон является платформой, равна 0,3214.

1.8 Последовательность независимых испытаний

1.8.1 Последовательность независимых испытаний; испытания Бернулли; схема Бернулли

Повторные испытания называются *независимыми*, если вероятности их исходов не зависят от исходов предшествующих испытаний. Например, многократное подбрасывание кубика, стрельба по мишеням (если считать вероятность попадания неизменной), безотказная работа однотипных устройств, эксплуатируемых в одинаковых условиях.

Испытаниями Бернулли называются повторные независимые испытания, в каждом из которых возможны два исхода (условно именуемые “успехом” и “неудачей”), вероятности которых не меняются от испытания к испытанию. Примерами испытаний Бернулли являются: многократное подбрасывание монеты (успех – выпадение герба); стрельба по мишеням в биатлоне (если считать вероятность попадания неизменной); проверка автобусов перед выходом на линию

$$\omega = \{\omega_1, \omega_2, \dots, \omega_n\}, \quad (1.8.1.1)$$

где $\omega_i = \begin{cases} \text{"у"}, & \text{если в } i\text{-м испытании произошёл успех;} \\ \text{"н"}, & \text{если в } i\text{-м испытании произошла неудача.} \end{cases}$

Обозначим через m количество успехов в n испытаниях Бернулли. Тогда вероятность того, что в n испытаниях Бернулли произойдёт ровно k успехов, определяется формулой

$$P_n(k) = C_n^k p^k q^{n-k}, \quad (k = \overline{0, n}). \quad (1.8.1.2)$$

Формула (1.8.1.2) называется *формулой Бернулли*.

1.8.3 Наиболее вероятное число успехов в схеме Бернулли

Наиболее вероятным числом успехов в схеме Бернулли называется число k_0 , для которого справедливо следующее двойное неравенство:

$$P_n(k_0 - 1) \leq P_n(k_0) \leq P_n(k_0 + 1), \quad (1.8.3.1)$$

т. е. вероятность появления именно k_0 успехов в n испытаниях не меньше вероятности появления меньшего или большего числа успехов, чем k_0 .

Иначе, наиболее вероятное число успехов в схеме Бернулли определяется следующим двойным неравенством:

$$np - q \leq k_0 \leq np + p. \quad (1.8.3.2)$$

Пример 24

На автобазе имеется десять автомашин. Вероятность выхода на линию каждой из них равна 0,8.

Найти: а) вероятность того, что в определенный день на линию выйдут 9 автомашин; б) вероятность нормальной работы автобазы в ближайший день, если для этого необходимо иметь на линии не менее восьми автомашин; в) наивероятнейшее число вышедших на линию автомашин и соответствующую этому числу вероятность.

Решение. Предполагая, что выходы машин на линию осуществляются независимо друг от друга, условие задачи можно рассматривать как серию из $n=10$ независимых испытаний, в каждом из которых вероятность события $A = \{\text{выход автомашины на линию}\}$ равна 0,8. То есть $p = 0,8$, $q = 1 - p = 0,2$.

а) Для определения вероятности того, что в определенный день на линию выйдут 9 из 10 машин автобазы, воспользуемся формулой Бернулли

$$P_{10}(9) = C_{10}^9 p^9 q^1 = \frac{10!}{9!1!} \cdot 0,8^9 \cdot 0,2^1 = 10 \cdot 0,8^9 \cdot 0,2^1 \approx 0,2684.$$

б) Введем в рассмотрение событие $B = \{\text{нормальная работа автобазы}\}$. Тогда

$$P(B) = P_{10}(k \geq 8) = P_{10}(8) + P_{10}(9) + P_{10}(10) = 0,2718 + 0,2684 + 0,1074 = 0,6476;$$

$$P_{10}(8) = C_{10}^8 p^8 q^2 = \frac{10!}{8!2!} \cdot 0,8^8 \cdot 0,2^2 = 10 \cdot 0,8^8 \cdot 0,2^2 \approx 0,2718;$$

$$P_{10}(10) = C_{10}^{10} p^{10} q^0 = \frac{10!}{10!0!} \cdot 0,8^{10} \cdot 0,2^0 = 1 \cdot 0,8^{10} \cdot 1 \approx 0,1074.$$

в) Наивероятнейшее число k_0 вышедших на линию автомашин найдем по формуле $np - q \leq k_0 \leq np + p$. Отсюда $7,8 \leq k_0 \leq 8,8$. Единственное целое число k_0 , удовлетворяющее этому двойному неравенству $k_0 = 8$, $P_{10}(8) \approx 0,2718$. Этому значению k_0 соответствует наибольшее значение вероятности $P_{10}(8)$.

1.8.4 Предельная теорема Пуассона

Формула Бернулли позволяет точно определить вероятность появления k успехов в n испытаниях Бернулли. Однако при $n \rightarrow \infty$ ($n \gg 50$) применение формулы Бернулли осложнено вычислением больших факториалов и значительными вычислительными погрешностями, связанными с возведением в большую степень чисел, близких к нулю. В данном случае для вычисления вероятности появления k успехов в n испытаниях Бернулли следует использовать специальные предельные теоремы, рассматриваемые ниже.

Пусть число экспериментов Бернулли велико ($n \rightarrow \infty$), а вероятность успеха в каждом из них мала ($p \rightarrow 0$, $p < 0,1$) таким образом, что произведение $np = \lambda = \text{const}$ не мало и не велико; тогда вероятность появления ровно k успехов в n испытаниях Бернулли:

$$\lim_{\substack{n \rightarrow \infty \\ np = \lambda}} P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}. \quad (1.8.4.1)$$

Замечание 1 – Предельная теорема Пуассона позволяет приближенно вычислять вероятность появления ровно k маловероятных успехов в большом количестве экспериментов. Она тем точнее, чем меньше вероятность успеха и чем больше проводится испытаний Бернулли.

Замечание 2 – Предельная теорема Пуассона может применяться и в случае, если p велико ($(1-p) \rightarrow 0$). Для этого следует поменять местами понятия “успеха” и “неудачи”. В случае, когда вероятность успеха близка к 0,5, для вычисления вероятности $P_n(k)$ применяется локальная предельная теорема Муавра-Лапласа, рассматриваемая ниже.

Пример 25

В порту каждые сутки может появиться одно большегрузное судно с вероятностью $p = \frac{1}{6}$. Вероятность появления более одного судна в течение суток пренебрежимо мала. Какова вероятность того, что за месяц (30 дней) порт посетят не более 4 судов?

Решение. Предполагая, что суда появляются в порту независимо друг от друга, условие задачи можно рассматривать как серию из $n=30$ независимых испытаний, в каждом из которых вероятность события $A = \{\text{судно прибывает в порт}\}$ равна $\frac{1}{6}$. Для нахождения вероятности $P_n(k \leq 4)$ воспользуемся предельной теоремой Пуассона

$$P_n(k \leq 4) = \sum_{k=0}^4 P_n(k) = \sum_{k=0}^4 \frac{e^{-5} 5^k}{k!} = 0,4405.$$

1.8.5 Предельные теоремы Муавра-Лапласа

Локальная предельная теорема Муавра-Лапласа. Если в схеме Бернулли $0 < p < 1$, то при $n \rightarrow \infty$ справедлива следующая теорема:

$$\lim_{\substack{n \rightarrow \infty \\ 0 < p < 1}} P_n(k) \frac{\sqrt{npq}}{\varphi(x_k)} = 1,$$

где $\varphi(x_k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_k^2}{2}}$ – функция плотности стандартного нормального распределения (рисунок 7), (значения функции $\varphi(x_k)$ определяются по таблице из приложения А); $x_k = \frac{k - np}{\sqrt{npq}}$.

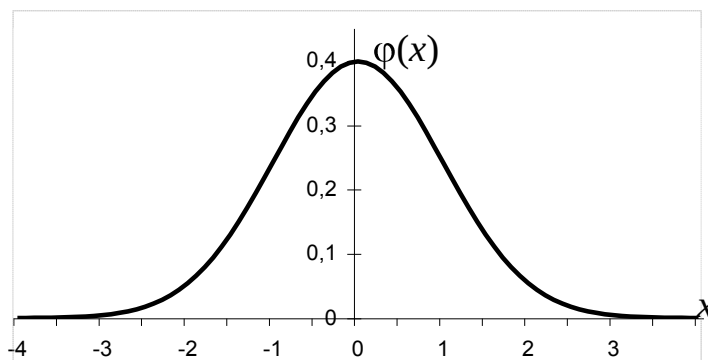


Рисунок 7 – Функция плотности стандартного нормального распределения

Следствие 1. При больших n вероятность появления ровно k успехов в n испытаниях Бернулли определяется приближенным выражением

$$P_n(k) \approx \frac{\varphi(x_k)}{\sqrt{npq}}. \quad (1.8.5.1)$$

Интегральная предельная теорема Муавра-Лапласа. Если в схеме Бернулли $0 < p < 1$, а $n \rightarrow \infty$, тогда для любых k_1 и k_2 , таких, что $0 \leq k_1 < k_2 \leq 1$, справедлива следующая теорема:

$$\lim_{\substack{n \rightarrow \infty \\ 0 < p < 1}} (P(k_1 \leq m \leq k_2) - \Phi(x_2) + \Phi(x_1)) = 0,$$

где $\Phi(x_i) = \frac{1}{\sqrt{2\pi}} \int_0^{x_i} e^{-\frac{z^2}{2}} dz$ – функция Лапласа (рисунок 8); $x_i = \frac{k_i - np}{\sqrt{npq}}$, $i = \overline{1,2}$.

Следствие 2. При больших значениях n вероятность того, что число успехов m в серии n испытаний Бернулли будет принадлежать отрезку $[k_1; k_2]$, определяется приближенным выражением

$$P_n(k_1, k_2) = P(k_1 \leq m \leq k_2) = \Phi(x_2) - \Phi(x_1). \quad (1.8.5.2)$$

Замечание 1 – Непосредственное вычисление значения функции Лапласа затруднено (интеграл является неберущимся), поэтому значения функции Лапласа табулированы и представлены в приложении Б. При использовании таблицы следует учитывать, что функция Лапласа – нечетная, т. е. $\Phi(-x) = -\Phi(x)$; при $x > 4$ функция Лапласа принимает значения, близкие к 0,5 (см. рисунок 8).

Замечание 2 – Теоремы Муавра-Лапласа позволяют получать приближенное значение искомой вероятности, однако оно тем точнее, чем ближе вероятность успеха p к 0,5 и чем больше проводится испытаний Бернулли n .

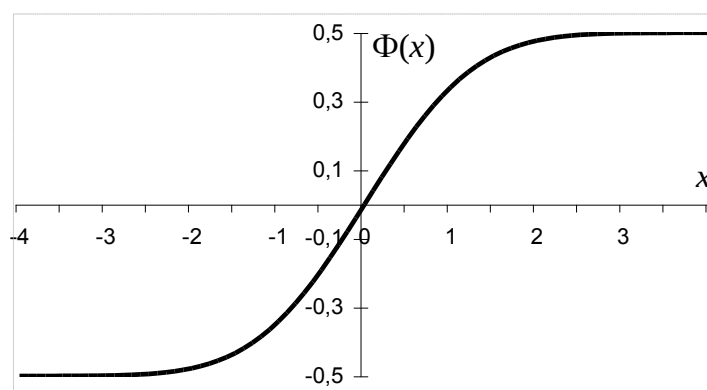


Рисунок 8 – Функция Лапласа

Пример 26

Депо производит ремонт вагонов. Вероятность того, что ремонт будет произведён со сдачей с первого предъявления, равна 0,6. Найти вероятность того, что из 100 вагонов, отремонтированных в депо:

- ровно 80 вагонов будут сданы с первого предъявления;
- от 40 до 60 вагонов будут сданы с первого предъявления.

Решение. Предполагая, что проверка качества ремонта вагонов осуществляется независимо друг от друга, условие задачи можно рассматривать как серию из $n=100$ независимых испытаний, в каждом из которых вероятность “успеха” {ремонт произведён со сдачей с первого предъявления} равна 0,6. То есть $p=0,6$, $q=1-p=0,4$.

а) Для вычисления вероятности события $A = \{\text{ровно 80 вагонов будут сданы с первого предъявления}\}$ воспользуемся локальной теоремой Муавра-Лапласа:

$$P(A) = P_{100}(80) = \frac{1}{\sqrt{100 \cdot 0,6 \cdot 0,4}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x = \frac{80 - 100 \cdot 0,6}{\sqrt{100 \cdot 0,6 \cdot 0,4}} \approx 4,08.$$

По таблицам значений функции стандартного нормального распределения

$$\varphi(x_k) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x_k^2}{2}}$$

находим, что $\varphi(x) \approx 0$. Следовательно, вероятность интересующего нас события $P(A) = P_{100}(80) \approx 0$.

б) Для вычисления вероятности события $B = \{\text{от 40 до 60 вагонов будут сданы с первого предъявления}\}$ воспользуемся интегральной теоремой Муавра-Лапласа:

$$P(B) = P_{100}(40 \leq m \leq 60) = \Phi(x_2) - \Phi(x_1),$$

$$x_1 = \frac{60 - 100 \cdot 0,6}{\sqrt{100 \cdot 0,6 \cdot 0,4}} = 0, \quad x_2 = \frac{40 - 100 \cdot 0,6}{\sqrt{100 \cdot 0,6 \cdot 0,4}} = -4,08.$$

По таблицам значений функции Лапласа

$$\Phi(x_i) = \frac{1}{\sqrt{2\pi}} \int_0^{x_i} e^{-\frac{z^2}{2}} dz$$

находим, что $\Phi(0) = 0$, $\Phi(-4,08) = -0,49997$.

Отсюда $P(B) = P_{100}(40 \leq m \leq 60) = \Phi(0) - \Phi(-4,08) = 0 + 0,49997$.

2 Одномерные случайные величины

2.1 Понятие случайной величины

2.2 Закон распределения случайной величины

2.3 Функция распределения случайной величины; свойства функции распределения

2.4 Функция плотности распределения непрерывной случайной величины; свойства функции плотности распределения

2.5 Числовые характеристики случайных величин

2.6 Законы распределения дискретных случайных величин

2.7 Законы распределения непрерывных случайных величин

2.1 Понятие случайной величины

Случайной величиной называется функция $\xi = \xi(\omega)$, которая каждому элементарному исходу ω пространства элементарных событий Ω вероятностного эксперимента E ставит в соответствие некоторое действительное число x .

Таким образом, областью определения случайной величины ξ является пространство элементарных исходов Ω , а областью значений – множество действительных чисел R (рисунок 9).

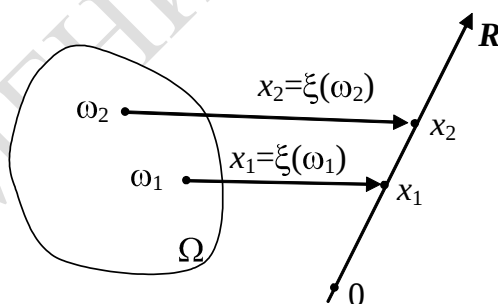


Рисунок 9

Пусть, например, в эксперименте с подбрасыванием монеты определена функция

$$\xi(\omega) = \begin{cases} 0, & \text{если выпал герб,} \\ 1, & \text{если – решка.} \end{cases}$$

Тогда $\xi = \xi(\omega)$ является случайной величиной.

Пример 27

Функция $\xi = \xi(\omega)$, определяющая число очков на верхней грани игральной кости, также является случайной величиной. Здесь элементарными исходами являются: $\omega_i = \{\text{выпадение на кости } i \text{ числа очков}\}$, $i = \overline{1,6}$. Пространство элементарных событий данного вероятностного эксперимента: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$.

В данном случае случайная величина ξ – число точек на верхней грани игральной кости определяется функцией $\xi = \xi(\omega_i) = i$, $i = \overline{1,6}$.

Из приведенных выше определений следует, что *случайная величина* – величина, которая в результате эксперимента обязательно принимает некоторое единственное значение, однако, заведомо неизвестное.

Примерами случайных величин являются: число составов, прибывших в течение суток на станцию; время простоя вагонов в ожидании разгрузки; масса топлива, израсходованного в течение суток и т. д.

Замечание – Условимся обозначать случайные величины малыми греческими буквами: $\xi, \eta, \dots$, а их значения – малыми латинскими буквами: $x_1, x_2, \dots, x_n, \dots; y_1, y_2, \dots, y_m, \dots$.

В зависимости от количества возможных значений случайные величины разделяются на два класса: дискретные и непрерывные.

Дискретной называется случайная величина ξ , которая в результате эксперимента E может принимать только определенные изолированные друг от друга значения. Множество значений дискретной случайной величины, определяемое пространством Ω , конечно или счетно.

Примерами дискретных случайных величин являются: число вагонов, прибывших в течение суток в депо для проверки; число бракованных деталей, изготовленных в течение смены; число успешно сданных экзаменов и т. д.

Непрерывной называется случайная величина ξ , которая в результате эксперимента может принимать все значения из некоторого промежутка или всей числовой оси. Множество значений непрерывной случайной величины, определяемое пространством Ω , несчетно.

Примерами непрерывных случайных величин являются: время простоя вагонов в ожидании разгрузки; масса топлива, израсходованного в течение суток и т. д.

2.2 Закон распределения случайной величины

Законом распределения дискретной случайной величины ξ называется правило, которое каждому возможному значению x величины ξ ставит в соответствие вероятность появления данного значения. Закон распределения полностью характеризует случайную величину ξ с вероятностной точки зрения, т. е. определяет множество значений, которое может принимать величина, и то, с какими вероятностями величина ξ принимает значения $x_1, x_2, \dots$.

Закон распределения случайной величины ξ может быть задан таблично, графически и аналитически (таблица 3).

Таблица 3 – Способы задания законов распределения случайных величин

Табличный	Графический		Аналитический		
Ряд распределения	Столбцовая диаграмма	Многоугольник распределения	Непосредственная формула $P(\xi = x_i)$	Функция распределения $F(x)$	Функция плотности распределения $f(x)$
ДСВ	ДСВ	ДСВ	ДСВ	ДСВ, НСВ	НСВ

Рядом распределения называется таблица, в которой непосредственно указаны возможные значения случайной величины ξ ($x_1, x_2, \dots, x_n, \dots$) и соответствующие им вероятности p_i (таблица 4). Причем $p_i > 0$, $\sum_i p_i = 1$.

Таблица 4 – Ряд распределения дискретных случайных величин

x_i	x_1	x_2	$\dots$	x_n	$\dots$
$P(\xi = x_i)$	p_1	p_2	$\dots$	p_n	$\dots$

Для наглядности ряд распределения случайных величин можно представить графически. По оси абсцисс откладываются возможные значения случайной величины, а по оси ординат – вероятности соответствующих значений (рисунок 10).

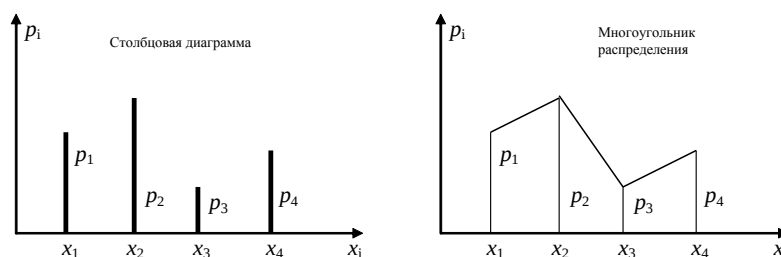


Рисунок 10 – Графические способы задания законов распределения

2.3 Функция распределения случайной величины; свойства функции распределения

Универсальным способом задания закона распределения произвольной случайной величины является функция распределения.

Функцией распределения $F(x)$ случайной величины ξ в точке x называется вероятность того, что величина ξ примет значение меньше x , т. е. функция распределения определяет вероятность события $\{\xi < x\}$:

$$F(x) = P(\xi < x). \quad (2.3.1)$$

Функция распределения произвольной случайной величины ξ обладает следующими свойствами:

Свойство 1. $F(x) \geq 0$, т. е. функция распределения – неотрицательная функция.

Свойство 2. $F(-\infty) = 0$.

Свойство 3. $F(\infty) = 1$.

Свойство 4. Если $x_1 < x_2$, то $F(x_1) \leq F(x_2)$, т. е. функция распределения – неубывающая функция.

Свойство 5. Если $x_1 < x_2$, то $P(x_1 \leq \xi < x_2) = F(x_2) - F(x_1)$, т. е. вероятность того, что случайная величина примет значение, принадлежащее полуинтервалу $[x_1, x_2)$, равна приращению функции распределения на этом полуинтервале.

Свойство 6. $F(x-0) = F(x)$, т. е. функция распределения непрерывна слева.

Замечание – У дискретных случайных величин функция распределения является разрывной ступенчатой функцией (имеет разрывы в точках, соответствующих возможным значениям случайной величины); у непрерывных величин функция распределения непрерывна на всей числовой оси.

Пример 28

В депо для проверки, независимо друг от друга, поступают полува-

гон, платформа и крытый вагон. Вероятности поступления в течение заданного интервала времени t для них соответственно равны: 0,6, 0,7 и 0,75. Рассматривается случайная величина ξ – число вагонов, поступивших на проверку в депо в течение времени t . Построить ряд распределения и вычислить функцию распределения данной случайной величины ξ .

Решение. Возможные значения данной случайной величины ξ : 0, 1, 2, 3. Запишем их в верхней строке ряда распределения. Для определения вероятностей возможных значений данной случайной величины введём в рассмотрение события: $A_i = \{\text{поступление в депо в течение времени } t \text{ } i\text{-го вагона}\}$, ($i = \overline{1,3}$); $B_j = \{\text{поступление в депо } j \text{ вагонов в течение времени } t\}$, ($j = \overline{1,3}$). Событие B_j можно представить в виде:

$$B_0 = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3};$$

$$B_1 = A_1 \cap \overline{A_2} \cap \overline{A_3} \cup \overline{A_1} \cap A_2 \cap \overline{A_3} \cup \overline{A_1} \cap \overline{A_2} \cap A_3;$$

$$B_2 = A_1 \cap A_2 \cap \overline{A_3} \cup A_1 \cap \overline{A_2} \cap A_3 \cup \overline{A_1} \cap A_2 \cap A_3;$$

$$B_3 = A_1 \cap A_2 \cap A_3.$$

Применяя теорему сложения вероятностей несовместных событий и теорему умножения вероятностей независимых событий, вычисляем:

$$P(\xi = 0) = P(B_0) = P(\overline{A_1})P(\overline{A_2})P(\overline{A_3}) = 0,4 \cdot 0,3 \cdot 0,25 = 0,03;$$

$$\begin{aligned} P(\xi = 1) &= P(B_1) = P(A_1)P(\overline{A_2})P(\overline{A_3}) + P(\overline{A_1})P(A_2)P(\overline{A_3}) + \\ &+ P(\overline{A_1})P(\overline{A_2})P(A_3) = 0,6 \cdot 0,3 \cdot 0,25 + 0,4 \cdot 0,7 \cdot 0,25 + 0,4 \cdot 0,3 \cdot 0,75 = \\ &= 0,205; \end{aligned}$$

$$\begin{aligned} P(\xi = 2) &= P(B_2) = P(A_1)P(A_2)P(\overline{A_3}) + P(A_1)P(\overline{A_2})P(A_3) + \\ &+ P(\overline{A_1})P(A_2)P(A_3) = 0,6 \cdot 0,7 \cdot 0,25 + 0,6 \cdot 0,3 \cdot 0,75 + 0,4 \cdot 0,7 \cdot 0,75 = \\ &= 0,45; \end{aligned}$$

$$P(\xi = 3) = P(B_3) = P(A_1)P(A_2)P(A_3) = 0,6 \cdot 0,7 \cdot 0,75 = 0,315.$$

Таким образом, ряд распределения случайной величины ξ имеет вид:

x_i	0	1	2	3
$P(\xi = x_i)$	0,03	0,205	0,45	0,315

Убедимся, что $\sum_{i=1}^4 p_i = \sum_{i=1}^4 P(\xi = x_i) = 1$.

Столбцовая диаграмма и многоугольник распределения, представляющие ряд распределения этой случайной величины, изображены

соответственно на рисунке 11, а, б.

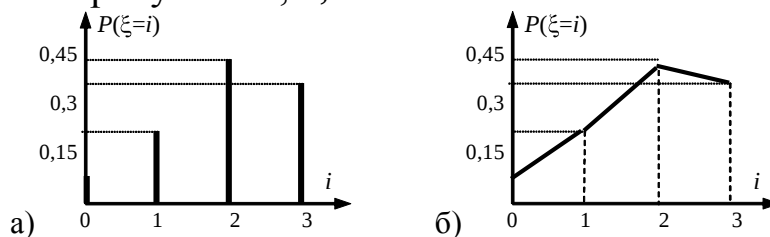


Рисунок 11 – Графические способы задания законов распределения дискретной случайной величины

Определим значение функции распределения $F(x) = P(\xi < x)$ для всех возможных значений x :

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

при $x \in (0; 1]$, $F(x) = P(\xi = 0) = 0,03$;

при $x \in (1; 2]$, $F(x) = P(\xi = 0) + P(\xi = 1) = 0,03 + 0,205 = 0,235$;

при $x \in (2; 3]$, $F(x) = P(\xi = 0) + P(\xi = 1) + P(\xi = 2) = 0,03 + 0,205 + 0,45 = 0,685$;

при $x \in [3; +\infty)$, $F(x) = P(\xi = 0) + P(\xi = 1) + P(\xi = 2) + P(\xi = 3) = 0,03 + 0,205 + 0,45 + 0,315 = 1$.

Таким образом,

$$F(x) = \begin{cases} 0; & \text{при } x \in (-\infty; 0]; \\ 0,03; & \text{при } x \in (0; 1]; \\ 0,235; & \text{при } x \in (1; 2]; \\ 0,685; & \text{при } x \in (2; 3]; \\ 1; & \text{при } x \in (3; +\infty). \end{cases}$$

График функции $F(x)$ изображён на рисунке 12.

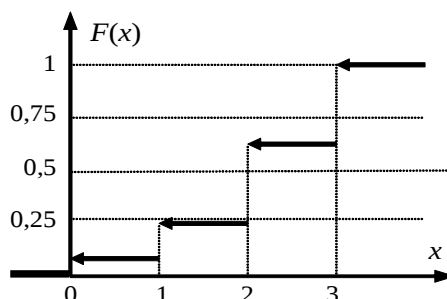


Рисунок 12 – График функции $F(x)$

2.4 Функция плотности распределения непрерывной случайной величины. Свойства функции плотности распределения

Рассмотрим непрерывную случайную величину ξ . На основании

свойства 5 функции распределения (см. подразд. 2.3), найдем вероятность попадания величины ξ в полуинтервал $[x_1, x_1 + \Delta x)$:

$$P(x_1 \leq \xi < x_1 + \Delta x) = F(x_1 + \Delta x) - F(x_1).$$

Разделим обе части равенства на Δx и перейдем к пределу при $\Delta x \rightarrow 0$:

$$\lim_{\Delta x \rightarrow 0} \frac{P(x_1 \leq \xi < x_1 + \Delta x)}{\Delta x} = \lim_{\Delta x \rightarrow 0} \frac{F(x_1 + \Delta x) - F(x_1)}{\Delta x} = F'(x_1) = f(x_1).$$

Предел отношения приращения функции к приращению аргумента при стремлении приращения аргумента к нулю – есть производная функции распределения $F(x)$ в точке x_1 . Эту производную принято называть *функцией плотности распределения* (дифференциальной функцией распределения) случайной величины ξ и обозначать $f(x)$. Таким образом, *функция плотности распределения* непрерывной величины ξ в точке x характеризует вероятность попадания значения случайной величины ξ в окрестность точки x , отнесенную к величине данной окрестности Δx :

$$f(x) = \lim_{\Delta x \rightarrow 0} \frac{P(x \leq \xi < x + \Delta x)}{\Delta x} = F'(x). \quad (2.4.1)$$

Вероятностный смысл функции $f(x)$ заключается в том, что она указывает, насколько вероятно значения непрерывной случайной величины ξ попадут в окрестность точки x .

Функция плотности распределения произвольной случайной величины ξ обладает следующими свойствами:

Свойство 1. $f(x) \geq 0$, т. е. функция плотности распределения – неотрицательная функция.

Свойство 2. $f(x) = F'(x)$.

Свойство 3. $P(x_1 \leq \xi < x_2) = \int_{x_1}^{x_2} f(x) dx$.

Свойство 4. $\int_{-\infty}^{\infty} f(x) dx = 1$.

Свойство 5. $\int_{-\infty}^{x_1} f(x) dx = F(x_1)$.

Замечание 1 – Вероятность того, что *непрерывная* случайная вели-

чина примет значение x_1 , равна нулю, т. е. $P(\xi = x_1) = 0$. Вместе с тем, событие $\{\xi = x_1\}$ не является невозможным. Данное замечание может быть объяснено тем, что множество возможных значений непрерывной случайной величины несчетно, следовательно, принятие одного из них – есть практически невозможное событие.

Замечание 2 – На основании предыдущего замечания, для непрерывной величины ξ события $\{\xi \leq x_1\}$ и $\{\xi < x_1\}$, а также их вероятности будем отождествлять. Аналогично для событий $\{\xi \geq x_1\}$ и $\{\xi > x_1\}$.

Замечание 3 – Приведем еще одно определение понятия непрерывной случайной величины. *Непрерывной* называется случайная величина, функция распределения которой непрерывна и определяется выражением $F(x) = \int_{-\infty}^x f(t)dt$.

Замечание 4 – Функция плотности распределения определена только для непрерывных случайных величин.

2.5 Числовые характеристики случайных величин

2.5.1 Характеристики положения

Случайные величины в вероятностном смысле полностью характеризуются законами распределения. Однако на практике знание закона распределения случайной величины часто оказывается излишним. Иногда бывает достаточным знать лишь отдельные числовые параметры, характеризующие существенные черты закона распределения исследуемой случайной величины или некоторые ее характерные значения.

Характеристики, выражающие в сжатой форме наиболее существенные особенности распределения случайной величины, называются *числовыми характеристиками* случайной величины. Все числовые характеристики случайных величин разделяют на характеристики положения, рассеяния, характеристики асимметрии и эксцесса (таблица 5).

Таблица 5 – Числовые характеристики случайных величин

Характеристика			
положения	рассеяния	асимметрии	эксцесса
▪ математическое ожидание ▪ мода ▪ медиана ▪ квантили	▪ среднее квадратическое отклонение ▪ дисперсия ▪ коэффициент вариации	▪ коэффициент асимметрии	▪ коэффициент эксцесса

Математическим ожиданием случайной величины ξ называется число $M[\xi]$, характеризующее среднее значение случайной величины с учётом вероятностей её значений. Математическое ожидание дискретной случайной величины ξ вычисляется по формуле

$$M[\xi] = \sum_{i=1}^n x_i P(= x_i), \quad (2.5.1.1)$$

а непрерывной случайной величины ξ – по формуле

$$M[\xi] = \int_{-\infty}^{\infty} x f(x) dx, \quad (2.5.1.2)$$

если соответствующая сумма или интеграл сходятся абсолютно. В противном случае говорят, что математическое ожидание у случайной величины ξ отсутствует.

Геометрический смысл математического ожидания заключается в том, что $M[\xi]$ «уравновешивает» систему «балок», соответствующую столбцовой диаграмме дискретной случайной величины (рисунок 13, а), или криволинейную трапецию, образованную функцией плотности распределения непрерывной случайной величины (рисунок 13, б), поэтому математическое ожидание называют также средневзвешенным (взвешенным по вероятностям) значением случайной величины ξ .

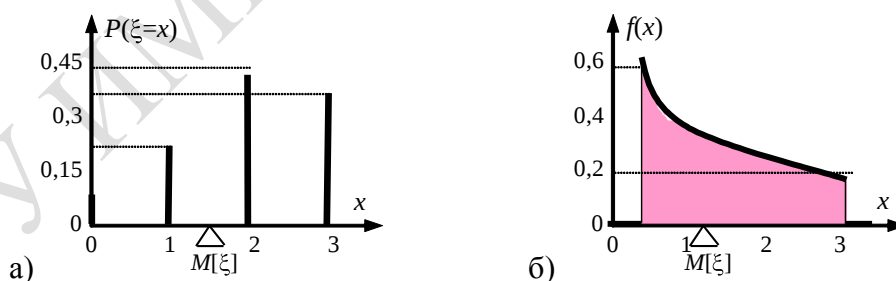


Рисунок 13 – Геометрический смысл математического ожидания

Математическое ожидание произвольной случайной величины обладает следующими свойствами.

Свойство 1. $M[\alpha] = \alpha$, где $\alpha - \text{const}$.

Свойство 2. $M[\alpha\xi] = \alpha M[\xi]$, где $\alpha - \text{const}$, ξ – произвольная случай-

ная величина.

Свойство 3. $M[\xi + \eta] = M[\xi] + M[\eta]$, где ξ и η – произвольные случайные величины.

Свойство 4. $M[\xi - \eta] = M[\xi] - M[\eta]$, где ξ и η – произвольные случайные величины.

Свойство 5. $M[\xi\eta] = M[\xi]M[\eta] + M[(\xi - M[\xi])(\eta - M[\eta])] = M[\xi]M[\eta] + \mu_{\xi\eta}$, где ξ и η – произвольные случайные величины, $\mu_{\xi\eta}$ – корреляционный момент случайных величин ξ и η .

Свойство 6. $M[\xi\eta] = M[\xi]M[\eta]$, где ξ и η – независимые случайные величины.

Свойство 7. Если $P(\alpha \leq \xi \leq \beta) = 1$, то $M[\xi] \in [\alpha, \beta]$, т. е. математическое ожидание произвольной случайной величины ξ принадлежит интервалу между минимальным и максимальным возможными значениями случайной величины ξ .

Мода – наиболее вероятное значение случайной величины ξ , обозначаемое $Mod[\xi]$. Моду дискретной случайной величины ξ определяют графически по столбцовой диаграмме, как абсциссу столбца, имеющего наибольшую высоту (рисунок 14, а). В примере 28 наиболее вероятным числом вагонов, поступивших на проверку в депо, является число 2, следовательно, $Mod[\xi] = 2$. Моду непрерывной величины определяют как значение величины, в котором функция плотности распределения $f(x)$ имеет максимум (рисунок 14, б).

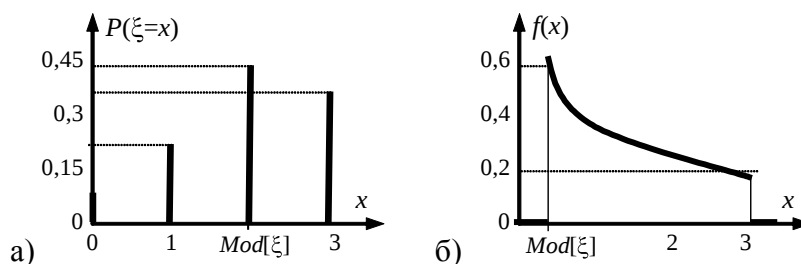


Рисунок 14 – Определение моды случайных величин

Замечание 1 – Некоторые случайные величины могут не иметь моды (все значения равновероятны) или иметь несколько мод. В этом

случае распределение случайной величины называют полимодальным.

Медианой называется значение случайной величины ξ , обозначаемое $Med[\xi]$ и для которого $P(\xi < Med[\xi]) = P(\xi > Med[\xi]) = 0,5$, т. е. величина ξ одинаково вероятно примет значение, меньшее или большее медианы, поэтому медиану называют средневероятным значением случайной величины ξ . Учитывая определение функции распределения (18), $F(Med[\xi]) = P(\xi < Med[\xi]) = 0,5$.

Замечание 2 – Медиана определена лишь для непрерывных случайных величин. Для дискретных случайных величин множество значений x , удовлетворяющих свойству медианы $F(Med[\xi]) = 0,5$, либо бесконечно, либо является пустым.

Геометрический смысл медианы заключается в том, что прямая $x = Med[\xi]$ делит площадь криволинейной трапеции, образованной графиком функции плотности непрерывной случайной величины ξ и осью абсцисс, пополам (рисунок 15). Учитывая, что площадь указанной криволинейной трапеции равна единице (см. свойство 4. $f(x)$), $S_1 = S_2 = 0,5$.

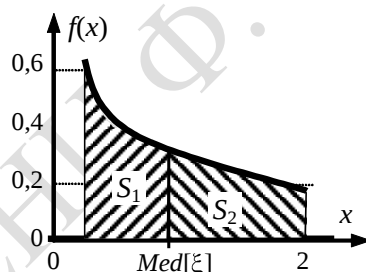


Рисунок 15 – Геометрический смысл медианы

Квантилью распределения случайной величины ξ уровня α называется значение x величины ξ , для которого выполняется равенство $F(x) = 1 - \alpha$, т. е. вероятность того, что случайная величина ξ примет значение, меньшее, чем x , равна $1 - \alpha$. Квантили распределения случайной величины ξ уровней 0,25 и 0,75 называются *квартелями*. Квантиль распределения непрерывной случайной величины ξ уровня 0,5 является медианой данной величины.

2.5.2 Характеристики рассеяния

Дисперсией случайной величины ξ называется число $D[\xi]$, характеризующее меру рассеяния значений случайной величины вокруг ее математического ожидания и равное математическому ожиданию

квадрата отклонения значений случайной величины ξ от $M[\xi]$:

$$D(\xi) = M[(\xi - M[\xi])^2] = M[\xi^2] - M^2[\xi]. \quad (2.5.2.1)$$

Дисперсия дискретной случайной величины определяется по формуле (2.5.2.2), а непрерывной – по формуле (2.5.2.3):

$$D(\xi) = \sum_{i=1}^{\infty} x_i^2 P(\xi = x_i) - M^2[\xi], \quad (2.5.2.2)$$

$$D(\xi) = \int_{-\infty}^{\infty} x^2 f(x) dx - M^2[\xi]. \quad (2.5.2.3)$$

Дисперсия произвольной случайной величины обладает следующими свойствами.

Свойство 1. $D(\xi) \geq 0$, т. е. дисперсия произвольной случайной величины ξ – неотрицательна.

Свойство 2. $D(\alpha) = 0$, где $\alpha - \text{const}$. То есть дисперсия неслучайной величины равна нулю.

Свойство 3. $D[\alpha\xi] = \alpha^2 D[\xi]$, где $\alpha - \text{const}$, ξ – произвольная случайная величина.

Свойство 4. $D[\xi \pm \eta] = D[\xi] + D[\eta] \pm 2\mu_{\xi\eta}$, где ξ и η – произвольные случайные величины, $\mu_{\xi\eta} = M[(\xi - M[\xi])(\eta - M[\eta])]$ – корреляционный момент случайных величин ξ и η .

Свойство 5. $D[\xi \pm \eta] = D[\xi] + D[\eta]$, где ξ и η – независимые случайные величины.

Свойство 6. $D[\alpha \pm \xi] = D[\xi]$, где $\alpha - \text{const}$, ξ – произвольная случайная величина.

Свойство 7. $D[\xi\eta] = D[\xi]D[\eta] + M^2[\xi]D[\eta] + D[\xi]M^2[\eta]$, где ξ и η – независимые случайные величины.

Свойство 8. Дисперсия случайной величины имеет размерность квадрата размерности случайной величины. Так, если случайная величина имеет размерность “час”, то дисперсия данной величины – “час²”.

Средним квадратическим отклонением случайной величины ξ называется число $\sigma[\xi]$, равное положительному значению квадратного корня из дисперсии, т. е. $\sigma[\xi] = +\sqrt{D[\xi]}$. Таким образом, среднее квадратическое отклонение $\sigma[\xi]$ имеет размерность, равную размерности случайной величины ξ . Среднее квадратическое отклонение $\sigma[\xi]$, как и дисперсия, характеризует степень разброса значений случайной величины ξ вокруг $M[\xi]$; чем больше разброс, тем больше $D[\xi]$ и $\sigma[\xi]$ (рисунок 16).

Коэффициентом вариации случайной величины ξ называется число $V[\xi]$, определяемое выражением $V[\xi] = \frac{\sigma[\xi]}{M[\xi]}$, характеризующее, насколько хорошо математическое ожидание $M[\xi]$ представляет ряд возможных значений случайной величины ξ .

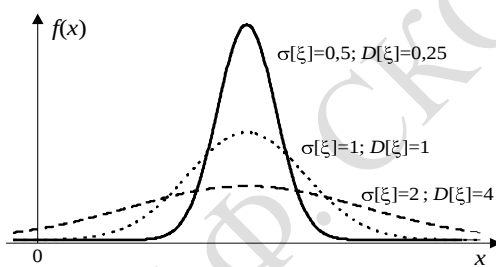


Рисунок 16 – Иллюстрация значений средних квадратических отклонений и дисперсии различных случайных величин

2.5.3 Моменты случайных величин; характеристики асимметрии и эксцесса

Моментом случайной величины ξ k -го порядка называется число $M[(\xi - a)^k]$, где a – произвольное число. Если $a = 0$, то момент случайной величины называется **начальным** ($v_k = M[\xi^k]$), если $a = M[\xi]$, то момент случайной величины ξ называется **центральный** моментом k -го порядка ($\mu_k = M[(\xi - M[\xi])^k]$).

Очевидно, что $v_0 = M[\xi^0] = 1$; $v_1 = M[\xi]$; $v_2 = M[\xi^2]$; $\mu_0 = M[(\xi - M[\xi])^0] = 1$; $\mu_1 = M[\xi - M[\xi]] = M[\xi] - M[M[\xi]] =$

$$= M[\xi] - M[\xi] = 0; \quad \mu_2 = M[(\xi - M[\xi])^2] = D[\xi].$$

При этом центральные и начальные моменты связаны между собой следующими соотношениями:

$$\mu_0 = 1; \quad \mu_1 = 0; \quad \mu_2 = v_2 - (v_1)^2; \quad \mu_3 = v_3 - 3v_2v_1 + 2v_1^3; \quad \mu_4 = v_4 - 4v_3v_1 + 6v_2v_1^2 - 3v_1^4.$$

Рассмотрим несколько важных особенностей центральных моментов старших порядков.

Коэффициентом асимметрии (скошенности) распределения случайной величины ξ называется число, вычисляемое по формуле

$$\beta_1[\xi] = \frac{\mu_3}{\sigma^3} = \frac{M[(\xi - M[\xi])^3]}{\sigma^3}. \quad (2.5.3.1)$$

Если распределение вероятностей случайной величины скошено влево, то $\beta_1[\xi] > 0$ (рисунок 17, а); если вправо, то $\beta_1[\xi] < 0$ (рисунок 17, б), если же распределение вероятностей случайной величины ξ симметрично относительно математического ожидания $M[\xi]$, то $\beta_1[\xi] = 0$.

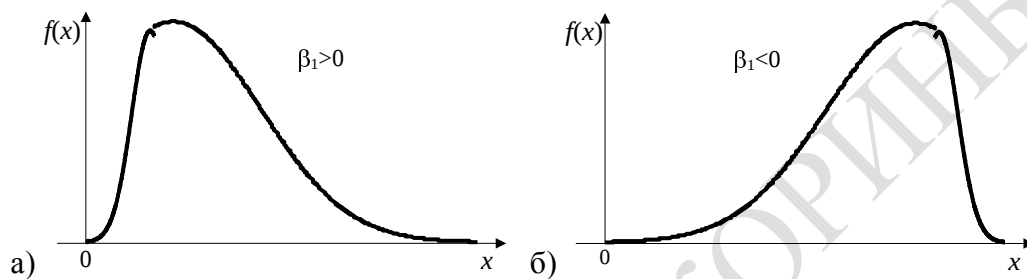


Рисунок 17 – Иллюстрация значений коэффициента асимметрии различных случайных величин

Коэффициентом эксцесса случайной величины ξ называется число $\beta_2[\xi]$, характеризующее островершинность распределения случайной величины ξ по сравнению с нормальным распределением и определяемое по формуле

$$\beta_2[\xi] = \frac{\mu_4}{\sigma^4} - 3. \quad (2.5.3.2)$$

У случайной величины ξ , имеющей нормальный закон распределения коэффициент эксцесса равен нулю, т. е. $\beta_2[\xi] = 0$. У случайных величин с более островершинным распределением $\beta_2[\xi] > 0$, а у величин с менее островершинным – $\beta_2[\xi] < 0$ (рисунок 18).

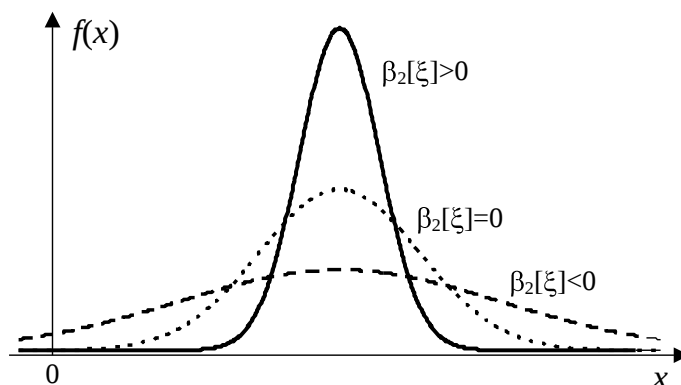


Рисунок 18 – Иллюстрация значений коэффициента эксцесса различных случайных величин

2.6 Законы распределения дискретных случайных величин

2.6.1 Биномиальный закон распределения

Говорят, что случайная величина ξ имеет биномиальный закон распределения (обозначается: $\xi \sim Bi(n, p)$), если данная величина дискретна и определяет число успехов k в схеме n испытаний Бернулли. Очевидно, что случайная величина ξ , имеющая биномиальное распределение, принимает только целые значения на отрезке $[0; n]$ с вероятностями, определяемыми формулой Бернулли

$$P(\xi = k) = P_n(k) = C_n^k p^k q^{n-k}, \quad (k = \overline{0, n}). \quad (2.6.1.1)$$

Биномиальное распределение характеризуется двумя параметрами: числом проводимых экспериментов n и вероятностью успеха p в каждом испытании.

Примерами случайных величин, имеющих биномиальное распределение, являются: число гербов, выпавших при n подбрасываниях монеты; количество бракованных деталей в партии из n штук; количество автобусов, вышедших на линию, и др.

На рисунке 19 представлены столбцовые диаграммы случайных величин, имеющих биномиальный закон распределения с различными значениями параметров n и p . Основные числовые характеристики случайных величин, распределенных по биномиальному закону, определяются следующими выражениями:

$$\begin{aligned} M[\xi] &= np; \quad D[\xi] = np(1-p); \\ \text{Mod}[\xi] &\text{— ближайшее целое к } (n+1)p - 0,5. \end{aligned} \quad (2.6.1.2)$$

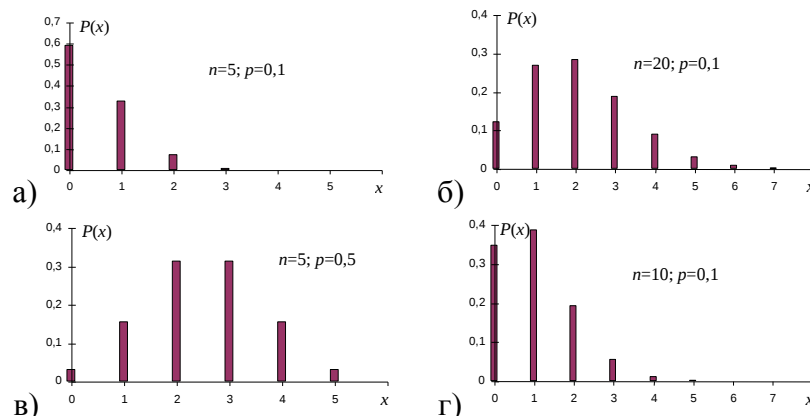


Рисунок 19 – Столбцовые диаграммы случайных величин, которые имеют биномиальное распределение с различными значениями параметров

Пример 29

В автобусном парке имеется пять автобусов. Вероятность выхода на линию каждого из них равна 0,8. Случайная величина ξ – число вышедших на линию машин. Построить ряд распределения и вычислить функцию распределения данной случайной величины. Вычислить её математическое ожидание, дисперсию и среднее квадратическое отклонение непосредственно по ряду распределения и сравнить со значениями, которые получаются при использовании формул (2.6.1.2). Найти вероятность того, что в определенный день на линию выйдут не менее четырёх автобусов.

Решение. Предполагая, что выходы автобусов на линию осуществляются независимо друг от друга, условие задачи можно рассматривать как серию из $n=5$ независимых испытаний, в каждом из которых вероятность события $A = \{\text{выход автобуса на линию}\}$ равна 0,8. Случайная величина ξ , обозначающая число вышедших на линию машин, распределена по биномиальному закону. Возможные значения случайной величины ξ : 0, 1, 2, 3, 4, 5. Вероятности значений определяются по формуле Бернулли:

$$P(\xi = 0) = C_5^0 p^0 q^5 = 1 \cdot 1 \cdot 0,2^5 = 0,00032;$$

$$P(\xi = 1) = C_5^1 p^1 q^4 = 5 \cdot 0,8 \cdot 0,2^4 = 0,0064;$$

$$P(\xi = 2) = C_5^2 p^2 q^3 = 10 \cdot 0,8^2 \cdot 0,2^3 = 0,0512;$$

$$P(\xi = 3) = C_5^3 p^3 q^2 = 10 \cdot 0,8^3 \cdot 0,2^2 = 0,2048;$$

$$P(\xi = 4) = C_5^4 p^4 q^1 = 5 \cdot 0,8^4 \cdot 0,2^1 = 0,4096;$$

$$P(\xi = 5) = C_5^5 p^5 q^0 = 1 \cdot 0,8^5 \cdot 1 = 0,32768.$$

Ряд распределения имеет вид:

x_i	0	1	2	3	4	5
p_i	0,00 032	0,0 064	0,0 512	0,2 048	0,4 096	0,32 768

Убедимся, что $\sum_{i=1}^6 p_i = 1$.

Столбцовая диаграмма и многоугольник распределения, представляющие ряд распределения этой случайной величины, изображены на рисунке 20, а, б.

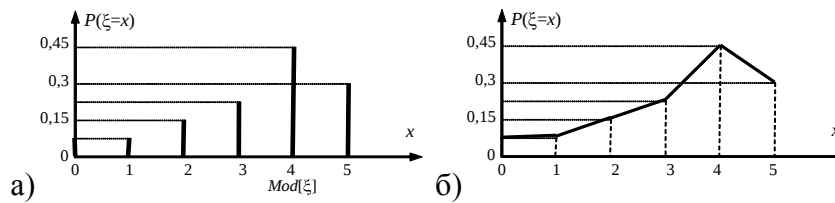


Рисунок 20 – Графические способы задания законов распределения дискретной случайной величины

Определим значение функции распределения $F(x) = P(\xi < x)$ для всех возможных значений x :

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

при $x \in (0; 1]$, $F(x) = P(\xi = 0) = 0,00032$;

при $x \in (1; 2]$, $F(x) = P(\xi = 0) + P(\xi = 1) = 0,00032 + 0,00672 = 0,00704$

при $x \in (2; 3]$, $F(x) = P(\xi = 0) + P(\xi = 1) + P(\xi = 2) =$
 $= 0,00704 + 0,0512 = 0,05824$

при $x \in (3; 4]$, $F(x) = P(\xi = 0) + \dots + P(\xi = 3) =$
 $= 0,05824 + 0,2048 = 0,26304$;

при $x \in (4; 5]$, $F(x) = P(\xi = 0) + \dots + P(\xi = 4) =$
 $= 0,26304 + 0,4096 = 0,67264$;

при $x \in [5; +\infty)$, $F(x) = P(\xi = 0) + \dots + P(\xi = 5) =$
 $= 0,67264 + 0,32736 = 1$.

Таким образом,

$$F(x) = \begin{cases} 0, & \text{при } x \in (-\infty; 0]; \\ 0,00032 & \text{при } x \in (0; 1]; \\ 0,00704 & \text{при } x \in (1; 2]; \\ 0,05824 & \text{при } x \in (2; 3]; \\ 0,26304 & \text{при } x \in (3; 4]; \\ 0,67264 & \text{при } x \in (4; 5]; \\ 1, & \text{при } x \in [5; +\infty). \end{cases}$$

График функции $F(x)$ изображён на рисунке 21.

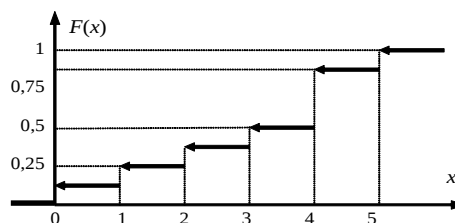


Рисунок 21 – График функции $F(x)$

Вычислим числовые характеристики данной случайной величины:

а) математическое ожидание

$$M[\xi] = \sum_{i=1}^6 x_i p_i = 0 \cdot 0,00032 + 1 \cdot 0,0064 + 2 \cdot 0,0512 + 3 \cdot 0,2048 + \\ + 4 \cdot 0,4096 + 5 \cdot 0,32768 = 4 \text{ (автобуса)};$$

б) дисперсия

$$D[\xi] = \sum_{i=1}^6 x_i^2 p_i - (M[\xi])^2 = 0 \cdot 0,00032 + 1 \cdot 0,0064 + 4 \cdot 0,0512 + \\ + 9 \cdot 0,2048 + 16 \cdot 0,4096 + 25 \cdot 0,32768 - 4^2 = 0,8 \text{ (автобуса}^2\text{)};$$

в) среднее квадратическое отклонение

$$\sigma[\xi] = \sqrt{D[\xi]} = \sqrt{0,8} \approx 0,894 \text{ (автобуса)};$$

г) мода

$$Mod[\xi] = 4 \text{ (автобуса)}.$$

Вычислим числовые характеристики данной случайной величины по формулам (2.6.1.2):

$$M[\xi] = np = 5 \cdot 0,8 = 4; \quad D[\xi] = np(1 - p) = 0,8.$$

Как и следовало ожидать, получены точно такие же значения.

Вероятность того, что в определенный день на линию выйдут не менее четырёх автобусов,

$$P(\xi \geq 4) = P(\xi = 4) + P(\xi = 5) = 0,4096 + 0,32768 = 0,73728.$$

2.6.2 Закон распределения Пуассона

Если в схеме Бернулли число испытаний бесконечно велико ($n \rightarrow \infty$), а вероятность успеха в каждом испытании стремится к нулю ($p \rightarrow 0$) таким образом, что $np = \lambda = \text{const}$, то вероятность появления в схеме n испытаний Бернулли ровно k успехов определяется предельной теоремой Пуассона (1.8.4.1). В данном случае говорят, что случайная величина ξ , определяющая число успехов k в схеме n испытаний Бернулли, имеет пуассоновский закон распределения, т. е. $\xi \sim \Pi(\lambda)$.

Таким образом, закон распределения Пуассона является предельным случаем биномиального закона распределения, при $n \rightarrow \infty$, $p \rightarrow 0$, $np = \lambda = \text{const}$.

Очевидно, что распределение Пуассона характеризуется единственным параметром $\lambda = np$; а случайная величина ξ , имеющая пуассоновское распределение, принимает только целые значения на полуинтервале $[0, +\infty)$ с вероятностями, определяемыми предельной тео-

ремой Пуассона (1.8.4.1):

$$P(\xi = k) = P_n(k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k = \overline{0, \infty}. \quad (2.6.2.1)$$

Примерами случайных величин, имеющих пуассоновское распределение, являются: число железнодорожных составов, поступающих на сортировочную горку в течение суток; число заявок, поступающих на АТС в течение часа, и др.

На рисунке 22 представлены столбцовые диаграммы случайных величин, имеющих пуассоновский закон распределения с различными значениями параметра λ . Основные числовые характеристики случайных величин, которые имеют пуассоновский закон распределения, определяются следующими выражениями:

$$M[\xi] = \lambda; D[\xi] = \lambda; \text{Mod}[\xi] - \text{ближайшее целое к } \lambda - 0,5. \quad (2.6.2.2)$$

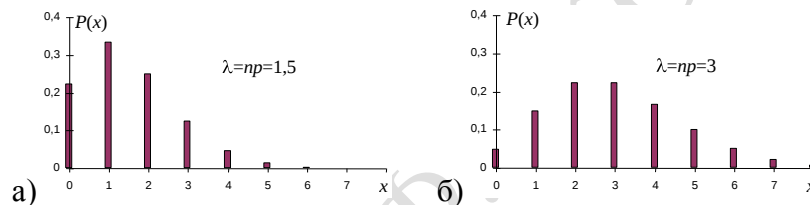


Рисунок 22 – Столбцовые диаграммы случайных величин, которые имеют распределение Пуассона с различными значениями параметров

Замечание – По закону Пуассона распределена случайная величина, описывающая число событий простейшего потока, произошедших в течение промежутка времени t .

Потоком событий называется последовательность однородных событий, наступающих одно за другим в случайные моменты времени.

Интенсивностью потока α называется среднее число событий, происходящих за единицу времени.

Если $\alpha = \text{const}$, то поток называется *стационарным*. Это свойство означает, что вероятность наступления того или иного числа событий в течение отрезка времени длиной t не зависит от расположения на оси времени этого отрезка, а зависит только от его длины.

Поток называется *ординарным*, если вероятность попадания на малый участок Δt двух и более событий пренебрежимо мала по сравнению с вероятностью попадания на него одного события.

Поток событий называется *поток без последствий*, если вероятность попадания того или иного числа событий на какой-то отрезок времени не зависит от того, сколько событий попало на любой другой

непересекающийся с ним участок, то есть предыстория потока не сказывается на вероятности появления событий в ближайшем будущем. Эта независимость физически сводится к тому, что события появляются на оси времени в силу случайных причин, индивидуальных для каждого из них.

Поток, обладающий свойствами стационарности, ординарности и отсутствия последействия, называется *простейшим*.

Доказано, что для простейшего потока число событий, попадающих на каждый отрезок времени длиной t , распределено по закону Пуассона с параметром $\lambda = \alpha t$, где α – интенсивность потока.

Пример 30

На сортировочную горку поступает поток железнодорожных составов с интенсивностью 4 состава в час. Поток составов является простейшим. Найти вероятность того, что в течение 30 минут на горку поступит хотя бы один состав.

Решение. Случайная величина ξ , определяющая число составов, поступивших в течение получаса, может принимать значения 0, 1, 2, 3, 4 и, согласно условию, распределена по закону Пуассона с параметром $\lambda = \alpha t = 4 \cdot 0,5 = 2$ (так как интенсивность потока $\alpha = 4$; $t = 0,5$ [ч]). Обозначим событие: $A = \{\text{в течение получаса поступит хотя бы один состав}\}$. Тогда

$$P(A) = P(\xi \geq 0) = 1 - P(\xi = 0) = 1 - \frac{2^0}{0!} e^{-2} = 0,865.$$

Пример 31

В железнодорожное депо на ремонт поступают вагоны. На основании статистических данных известно, что для некоторого промежутка времени рабочего дня среднее число вагонов, поступающих в течение 1 часа, равно 10. Поток поступлений является простейшим. Для этого промежутка времени найти вероятность того, что: а) в течение часа поступит хотя бы один вагон; б) в течение трех часов произойдет не менее четырех поступлений.

Решение. а) Случайная величина ξ , определяющая число вагонов, поступивших в течение часа, может принимать значения 0, 1, 2, 3, ... и, согласно условию, распределена по закону Пуассона с параметром $\lambda = \alpha t = 10 \cdot 1 = 10$ (так как интенсивность потока $\alpha = 10$; $t = 1$ [ч]). Обозначим событие: $A = \{\text{в течение часа поступит хотя бы один вагон}\}$. Тогда

$$P(A) = P(\xi \geq 0) = 1 - P(\xi = 0) = 1 - \frac{10^0}{0!} e^{-10} = 1 - e^{-10} = 0,99999546.$$

б) Для определения вероятности события $B = \{\text{в течение трех часов поступит не менее четырех вагонов}\}$ введем в рассмотрение случайную величину η , определяющую число вагонов, поступивших в течение трех часов. Эта случайная величина распределена по закону Пуассона с параметром $\lambda = \alpha t = 10 \cdot 3 = 30$.

$$P(B) = P(\eta \geq 4) = 1 - P(\eta < 4) = 1 - [P(\eta = 0) + P(\eta = 1) + P(\eta = 2) + P(\eta = 3)] = 1 - \left[\frac{30^0}{0!} e^{-30} + \frac{30^1}{1!} e^{-30} + \frac{30^2}{2!} e^{-30} + \frac{30^3}{3!} e^{-30} \right] \approx 0,999999.$$

2.6.3 Геометрический закон распределения

Говорят, что случайная величина ξ имеет геометрический закон распределения (обозначается: $\xi \sim G(p)$), если данная величина дискретна и определяет число независимых испытаний Бернулли, предшествующих первому появлению успеха. Множество значений случайной величины, имеющей геометрический закон распределения, – это множество неотрицательных целых чисел $0, 1, 2, \dots$. Геометрическое распределение характеризуется одним параметром – вероятностью успеха p в испытаниях Бернулли.

Пусть случайная величина ξ имеет геометрический закон распределения с параметром p , т. е. $\xi \sim G(p)$. Тогда, вероятности значений случайной величины ξ , имеющей геометрическое распределение, определяются выражением:

$$P(\xi = k) = (1 - p)^k p, \quad (k = \overline{0, n}). \quad (2.6.3.1)$$

Примерами случайных величин, имеющих геометрическое распределение, являются: количество безуспешных попыток установки модемного соединения, число безуспешных попыток спортсмена поразить мишень и некоторые другие.

На рисунке 23 представлены столбцовые диаграммы случайных величин, имеющих геометрическое распределение с различными значениями параметра p . Основные числовые характеристики случайных величин, которые имеют геометрический закон распределения, определяются выражениями:

$$M[\xi] = \frac{1}{p}; \quad D[\xi] = \frac{1-p}{p^2}; \quad \text{Mod}[\xi] = 0. \quad (2.6.3.2)$$

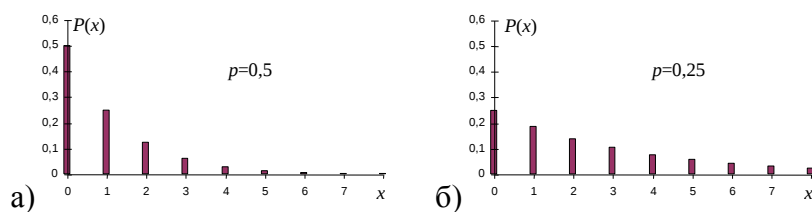


Рисунок 23 – Столбцовые диаграммы случайных величин, которые имеют геометрическое распределение с различными значениями параметров

Замечание – В некоторой литературе указывается, что геометрическое распределение имеет случайная величина η , определяющая номер испытания Бернулли, в котором впервые произошел успех. Очевидно, что указанная величина η может принимать значения из множества натуральных чисел $1, 2, \dots$, при этом $\eta = \xi + 1$, где $\xi \sim G(p)$, поэтому будем обозначать такую величину $\eta \sim G^{+1}(p)$.

2.7 Законы распределения непрерывных случайных величин

2.7.1 Равномерный закон распределения

Говорят, что непрерывная случайная величина ξ имеет равномерный (прямоугольный) закон распределения, т. е. $\xi \sim R(a, b)$, если она может принимать значения только на отрезке $[a; b]$, причём равновозможно. Таким образом, плотность распределения случайной величины ξ , имеющей равномерное распределение, постоянна на указанном отрезке (рисунок 24, б), т. е. вероятность попасть в окрестность любой точки от a до b одинакова.

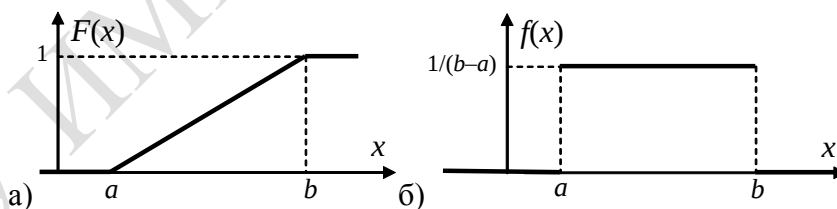


Рисунок 24 – Равномерный закон распределения:

а – функция распределения $F(x)$ случайной величины ξ ;

б – функция плотности распределения $f(x)$ случайной величины ξ

Равномерный закон распределения характеризуется двумя параметрами: минимальным a и максимальным b возможными значениями случайной величины ξ . Функция распределения $F(x)$ и функция плотности распределения $f(x)$ случайной величины ξ определяются выражениями (2.7.1.1) и (2.7.1.2), а их графики представлены на ри-

сунке 24.

$$F(x) = \begin{cases} 0, & \text{если } x < a; \\ \frac{x-a}{b-a}, & \text{если } a \leq x \leq b; \\ 1, & \text{если } x > b; \end{cases} \quad (2.7.1.1)$$

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{если } x \in [a, b]; \\ 0, & \text{если } x \notin [a, b]. \end{cases} \quad (2.7.1.2)$$

Примером случайной величины, которая имеет равномерный закон распределения, является время ожидания регулярных событий, например, время ожидания поезда в метрополитене. Другим примером являются ошибки округления чисел при арифметических вычислениях.

Основные числовые характеристики равномерно распределенной случайной величины ξ определяются следующими выражениями:

$$M[\xi] = Med[\xi] = \frac{a+b}{2}; \quad D[\xi] = \frac{(b-a)^2}{12}; \quad \beta_1[\xi] = 0; \\ Mod[\xi] - \text{отсутствует (т. к. значения равновероятны)}. \quad (2.7.1.3)$$

Пример 32

Интервал движения поездов в метрополитене составляет 3 мин. Определить вероятность, с которой пассажир, подошедший на платформу в случайный момент времени, будет ждать поезда более двух минут, а также среднее время ожидания поезда пассажиром.

Решение. Рассмотрим случайную величину ξ , равную времени ожидания поезда пассажиром. Очевидно, что время ожидания поезда – неотрицательная величина, не превышающая 3 минут, т. е. $\xi \in [0; 3]$. Предполагая, что время прихода пассажира равновозможно в интервале между прибытием поездов, будем считать, что $\xi \sim R(0, 3)$.

По существу, необходимо определить вероятность события $\{\xi > 2\}$, а также математическое ожидание величины ξ , т. е. $M[\xi]$.

Заметим, что $P(\xi > 2) = P(2 < \xi \leq 3) = P(2 \leq \xi < 3)$, т. к. ξ – непрерывная величина. Для определения указанной вероятности воспользуемся свойством 5 функции распределения и выражением функции распределения $F(x)$ случайной величины, имеющей равномерное распределение (2.7.1.1):

$$P(2 \leq \xi < 3) = F(3) - F(2) = \frac{3-0}{3-0} - \frac{2-0}{3-0} = 1 - \frac{2}{3} = \frac{1}{3}.$$

Математическое ожидание величины ξ определим по формуле (2.7.1.3):

$$M[\xi] = \frac{a+b}{2} = \frac{0+3}{2} = 1,5 \text{ мин.}$$

Ответ: вероятность, с которой пассажир будет ожидать поезда более 2 минут, равна $\frac{1}{3}$; среднее время ожидания поезда составляет 1,5 минуты.

2.7.2 Показательный закон распределения

Говорят, что случайная величина ξ имеет показательное (экспоненциальное) распределение, т. е. $\xi \sim E(\lambda)$, если она непрерывна, принимает только положительные значения и имеет функцию распределения

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ 1 - e^{-\lambda x}, & \text{если } x > 0, \end{cases} \quad (2.7.2.1)$$

и, следовательно, функцию плотности распределения

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ \lambda e^{-\lambda x}, & \text{если } x > 0, \end{cases} \quad (2.7.2.2)$$

где λ – единственный параметр показательного распределения, $\lambda > 0$.

Из графика функции плотности распределения $f(x)$ видно (рисунок 25, б), что случайная величина ξ , имеющая показательное распределение, наиболее вероятно принимает малые положительные значения, менее вероятно – большие положительные значения.

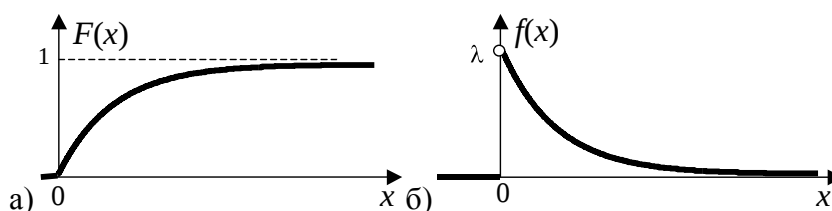


Рисунок 25 – Показательный закон распределения:
а – функция распределения $F(x)$ случайной величины ξ ;

$\bar{b}$ – функция плотности распределения $f(x)$ случайной величины ξ

Основные числовые характеристики случайной величины ξ , имеющей показательный закон распределения, определяются следующими выражениями:

$$M[\xi] = \sigma[\xi] = \frac{1}{\lambda}; \quad D[\xi] = \frac{1}{\lambda^2}; \quad \text{Mod}[\xi] = 0; \quad \text{Med}[\xi] = \frac{\ln 2}{\lambda}. \quad (2.7.2.3)$$

Примерами случайных величин, имеющих показательный закон распределения, являются: время простоя вагона в ожидании ремонта, интервалы времени между поездами, прибывающими на станцию, время наработки на отказ электронных систем тепловоза и другие, поэтому показательное распределение имеет важное значение в теории надежности и теории массового обслуживания.

Случайная величина, распределенная по показательному закону, обладает важным свойством, называемым «отсутствием памяти».

Лемма об «отсутствии памяти» у показательного распределения. Пусть ξ имеет показательное распределение с параметром λ (т. е. $\xi \sim E(\lambda)$). Тогда для любых $t > 0$ и $\tau > 0$ вероятность того, что величина ξ примет значение меньше, чем $(\tau + t)$ при условии, что ξ приняла значение не меньше, чем τ , равна безусловной вероятности того, что случайная величина ξ примет значение меньшее, чем t :

$$P(\xi < \tau + t \mid \xi \geq \tau) = P(\xi < t). \quad (2.7.2.4)$$

Замечание – Если случайная величина ξ – время до наступления некоторого события – имеет показательное распределение, то информация о том, что к моменту времени τ событие еще не наступило, не изменяет шансы на его наступление в дальнейшем. Этим свойством, называемым также отсутствием последствия, обладает только показательное распределение.

Пример 33

Время простоя вагона в ожидании ремонта является случайной величиной, распределенной по показательному закону с математическим ожиданием, равным 1,5 часа. Определить вероятность того, что в ожидании ремонта вагон простоит три часа.

Решение. Согласно условию, математическое ожидание случайной

величины ξ , обозначающей время простоя вагона в ожидании ремонта, равно 1,5. Учитывая, что для случайной величины, распределенной по показательному закону, $M[\xi] = \frac{1}{\lambda}$, определяем значение параметра: $\lambda = \frac{1}{M[\xi]} = \frac{1}{1,5} = \frac{2}{3}$. Функция плотности распределения данной случайной величины ξ имеет вид

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ \frac{2}{3} e^{-\frac{2}{3}x}, & \text{если } x > 0. \end{cases}$$

Определим вероятность простоя вагона в ожидании ремонта три часа:

$$P(\xi \geq 3) = \int_3^{\infty} f(x) dx = \int_3^{\infty} \frac{2}{3} e^{-\frac{2}{3}x} dx = -e^{-\frac{2}{3}x} \Big|_3^{\infty} = 0 + e^{-\frac{2}{3}} = 0,135.$$

Пример 34

Время безотказной работы электронного оборудования тепловоза является случайной величиной, распределенной по показательному закону. Определить вероятность безотказной работы оборудования в течение десяти часов эксплуатации, если среднее время безотказной работы по статистическим данным составляет 200 часов.

Решение. Согласно условию, математическое ожидание случайной величины ξ , обозначающей время безотказной работы оборудования, равно 200. Учитывая, что для случайной величины, распределенной по показательному закону, $M[\xi] = \frac{1}{\lambda}$, определяем значение параметра:

$\lambda = \frac{1}{M[\xi]} = \frac{1}{200} = 0,005$. Функция плотности распределения данной случайной величины ξ имеет вид

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ 0,005 e^{-0,005x}, & \text{если } x > 0. \end{cases}$$

Определим вероятность безотказной работы оборудования в течение десяти часов эксплуатации:

$$P(\xi \geq 10) = \int_{10}^{\infty} f(x) dx = \int_{10}^{\infty} 0,005 e^{-0,005x} dx = -e^{-0,005x} \Big|_{10}^{\infty} = 0 + e^{-0,005} = 0,95123$$

Пример 35

Известно, что для некоторой сортировочной станции интервалы времени между поездами, прибывающими в разборку, представляют собой случайную величину ξ , имеющую усечённое в точке x_0 показательное распределение:

$$f(x) = \begin{cases} 0, & \text{если } x < x_0; \\ A\lambda e^{-\lambda x}, & \text{если } x > x_0. \end{cases}$$

Требуется: 1) вычислить параметр A и построить график функции плотности распределения $f(x)$; 2) вычислить функцию распределения и построить её график; 3) вычислить числовые характеристики случайной величины ξ (математическое ожидание, моду, медиану, дисперсию, среднее квадратическое отклонение); 4) вычислить вероятность того, что случайная величина ξ примет значение в интервале $(1; 2)$.

Решение. 1) Неизвестный параметр A плотности распределения вероятностей найдём из соотношения

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

Поскольку в данном примере плотность $f(x)$ на $(-\infty; x_0)$ равна нулю, то можно записать:

$$\int_{x_0}^{+\infty} A\lambda e^{-\lambda x} dx = 1; \quad A \int_{x_0}^{+\infty} \lambda e^{-\lambda x} dx = 1; \quad A\lambda \left(-\frac{1}{\lambda} \right) e^{-\lambda x} \Big|_{x_0}^{+\infty} = A e^{-\lambda x_0} = 1,$$

отсюда $A = e^{\lambda x_0}$.

Следовательно, плотность распределения вероятностей имеет вид

$$f(x) = \begin{cases} 0, & \text{если } x < x_0; \\ \lambda e^{-\lambda(x-x_0)}, & \text{если } x > x_0. \end{cases}$$

Параметры распределения λ и x_0 имеют следующий смысл: λ — среднее число поездов, прибывающих на станцию в единицу времени; x_0 — минимальный допустимый интервал между последовательно прибывающими поездами.

Допустим $\lambda = 2$ поезда/ч, $x_0 = 0,017$ ч, тогда

$$f(x) = \begin{cases} 0, & \text{если } x < 0,017; \\ 2e^{-2(x-0,017)}, & \text{если } x > 0,017. \end{cases}$$

График функции $f(x)$ изображён на рисунке 26, а.

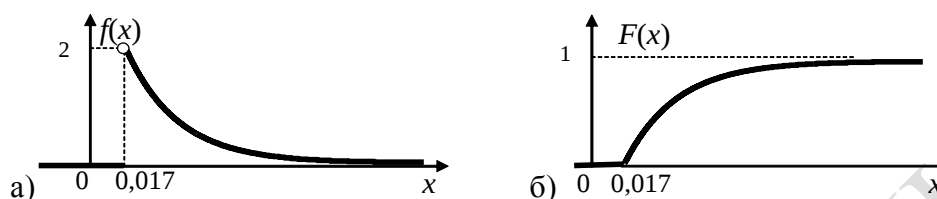


Рисунок 26 – Графики функций $f(x)$ и $F(x)$

2) Вычислим функцию распределения $F(x)$:

при $x \in (-\infty; 0,017]$, $F(x) = P(\xi < x) = 0$;

при $x \in (0,017; +\infty)$, $F(x) = P(\xi < x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^{0,017} 0dx +$

$$+ \int_{0,017}^x 2e^{-2(t-0,017)} dt = 2 \left(-\frac{1}{2} \right) e^{-2(t-0,017)} \Big|_{0,017}^x = -e^{-2(x-0,017)} + e^{-2 \cdot 0} =$$

$$= 1 - e^{-2(x-0,017)}.$$

Следовательно,

$$F(x) = \begin{cases} 0, & \text{если } x < 0,017; \\ 1 - e^{-2(x-0,017)}, & \text{если } x > 0,017. \end{cases}$$

График функции $F(x)$ изображён на рисунке 26, б.

3) Числовые характеристики исследуемой случайной величины:

а) математическое ожидание

$$M[\xi] = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{0,017} x \cdot 0 \cdot dx + \int_{0,017}^{+\infty} 2xe^{-2(x-0,017)} dx =$$

$$= \left[x = u; \quad v = -\frac{1}{2} e^{-2(x-0,017)} \right. \\ \left. e^{-2(x-0,017)} dx = dv; \quad dx = du \right] =$$

$$= 2 \left[-\frac{1}{2} x e^{-2(x-0,017)} \Big|_{0,017}^{+\infty} + \frac{1}{2} \int_{0,017}^{+\infty} e^{-2(x-0,017)} dx \right] =$$

$$= 0,017 - \frac{1}{2} e^{-2(x-0,017)} \Big|_{0,017}^{+\infty} = 0,517 \text{ ч};$$

б) мода случайной величины ξ равна 0,017 ч;

в) медиану найдём из уравнения $F(x) = 0,5$:

$$1 - e^{-2(x-0,017)} = 0,5;$$

$$e^{-2(x-0,017)} = 0,5;$$

$$-2(x - 0,017) = \ln 0,5;$$

$$2(x - 0,017) = 0,693;$$

$$x = 0,363 \text{ ч};$$

г) дисперсия

$$D[\xi] = \int_{-\infty}^{+\infty} x^2 f(x) dx - (M[\xi])^2 = \int_{-\infty}^{0,017} x^2 \cdot 0 \cdot dx + 2 \int_{0,017}^{+\infty} x^2 e^{-2(x-0,017)} dx -$$

$$- (0,517)^2 = 2e^{0,034} \int_{0,017}^{+\infty} x^2 e^{-2x} dx - (0,517)^2 =$$

$$= 2e^{0,034} \left[e^{-2x} \left(\frac{x^2}{2} - \frac{2x}{4} - \frac{2}{8} \right) \right] \Big|_{0,017}^{+\infty} - (0,517)^2 =$$

$$= 0,517 + (0,017)^2 - (0,517)^2 = 0,25 \text{ ч}^2;$$

д) среднее квадратическое отклонение

$$\sigma[\xi] = \sqrt{D[\xi]} = 0,5 \text{ ч.}$$

4) Вероятность того, что $\xi \in (1; 2)$, вычислим по формуле

$$P(1 < \xi < 2) = \int_1^2 2e^{-2(x-0,017)} dx = 2e^{0,034} \int_1^2 e^{-2x} dx =$$

$$= 2e^{0,034} \left(-\frac{1}{2} \right) e^{-2x} \Big|_1^2 = -e^{0,034} (e^{-4} - e^{-2}) = 0,121.$$

Вывод. Средний интервал времени между поездами, прибывающими на сортировочную станцию, равен 0,517 ч; наиболее вероятное значение интервала равно 0,017 ч; среднеквадратическое значение интервала между поездами равно 0,363 ч.

2.7.3 Закон распределения Эрланга

Пусть случайные величины $\xi_1, \xi_2, \dots, \xi_k$ независимы и имеют показательный закон распределения с одинаковым параметром $\lambda_i = \lambda$, т. е. $\xi_i \sim E(\lambda)$, $i = \overline{1, k}$. Тогда случайная величина $\xi = \xi_1 + \xi_2 + \dots + \xi_k$ имеет закон *распределения Эрланга* k -го порядка с параметром λ , т. е. $\xi \sim ER(\lambda, k)$. Очевидно, что величина ξ непрерывна и принимает лишь положительные значения.

Функция плотности распределения случайной величины $\xi \sim ER(\lambda, k)$ определяется выражением (2.7.3.1), а ее графики (для различных значений параметра k) представлены на рисунке 27.

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0; \\ \frac{\lambda^k x^{k-1}}{(k-1)!} e^{-\lambda x}, & \text{если } x > 0, \lambda > 0, k = 1, 2, 3, \dots \end{cases} \quad (2.7.3.1)$$

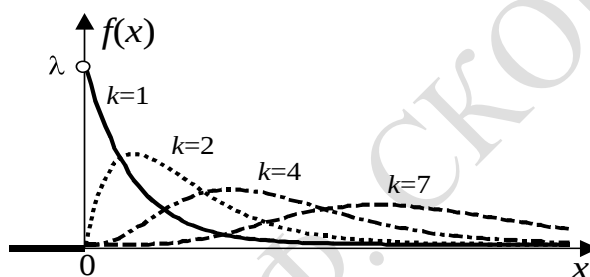


Рисунок 27 – Функция плотности распределения Эрланга для различных значений параметров

Очевидно, что при $k=1$ распределение Эрланга совпадает с показательным распределением. Основные числовые характеристики случайной величины ξ , имеющей закон распределения Эрланга, определяются следующими выражениями:

$$M[\xi] = \frac{k}{\lambda}; \quad D[\xi] = \frac{k}{\lambda^2}; \quad \text{Mod}[\xi] = \frac{k-1}{\lambda}. \quad (2.7.3.2)$$

2.7.4 Нормальный закон распределения

Говорят, что случайная величина ξ имеет нормальный (гауссовский) закон распределения, т. е. $\xi \sim N(m, \sigma)$, если она непрерывна, имеет функцию

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt \quad (2.7.4.1)$$

и, следовательно, имеет функцию плотности распределения

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad (2.7.4.2)$$

где параметр m численно равен математическому ожиданию величины ξ , а параметр $\sigma > 0$ – среднеквадратическому отклонению (т. е. $m = M[\xi]$, $\sigma = \sigma[\xi]$).

Максимальная ордината $\frac{1}{\sigma\sqrt{2\pi}}$ функции плотности распределения (рисунок 31, б) достигается в точке $x = m = M[\xi]$; при $x \rightarrow \pm\infty$ кривая плотности нормального распределения асимптотически стремится к нулю. Следовательно, наиболее вероятно, что нормально распределенная случайная величина примет значение близкое к m и практически никогда не будет принимать бесконечно больших значений.

Из графиков функции $F(x)$ и функции плотности распределения $f(x)$ (рисунок 28) видно, что коэффициент асимметрии случайной величины ξ , имеющей нормальный закон распределения, равен нулю (т. е. $\beta_1[\xi] = 0$); а математическое ожидание, медиана и мода случайной величины ξ совпадают. Основные числовые характеристики случайной величины ξ , имеющей нормальный закон распределения, определяются выражениями:

$$M[\xi] = Med[\xi] = Mod[\xi] = m; D[\xi] = \sigma^2; \beta_1[\xi] = \beta_2[\xi] = 0. \quad (2.7.4.3)$$

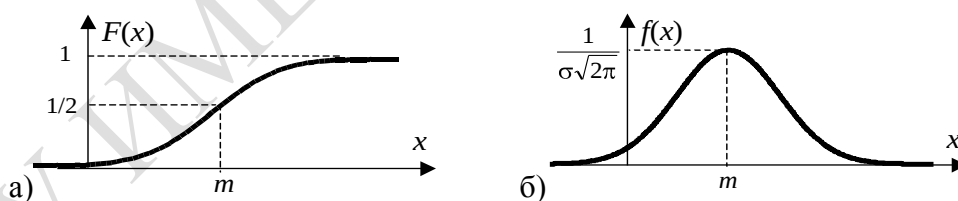


Рисунок 28 – Нормальный закон распределения:

а – функция распределения $F(x)$ случайной величины ξ ;

б – функция плотности распределения $f(x)$ случайной величины ξ

Случайная величина, которая имеет нормальный закон распределения с параметрами $m = 0$ и $\sigma = 1$, называется *стандартной нормальной случайной величиной*. Функция распределения стандартной нормальной случайной величины имеет вид:

$$\begin{aligned}
N(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \\
&= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt = \frac{1}{2} + \Phi(x),
\end{aligned} \tag{2.7.4.4}$$

где $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt = \frac{1}{2}$ (интеграл Пуассона), $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$ – функция Лапласа (см. рисунок 8), значения которой табулированы и представлены в приложении Б.

Вычисление значения функции распределения случайной величины, имеющей нормальный закон распределения, затруднено тем, что интеграл в выражении (2.7.4.4) не выражается через элементарные функции. Для выполнения указанной задачи функцию распределения случайной величины ξ , имеющей нормальный закон распределения с параметрами m и σ , выражают через функцию Лапласа следующим образом. Сделаем в выражении (2.7.4.4) замену переменной $z = \frac{t-m}{\sigma}$, следовательно,

$$\begin{aligned}
F(x) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-m)^2}{2\sigma^2}} dt = \left| \frac{t-m}{\sigma} = z \right|_{dt=\sigma \cdot dz} = \frac{\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{z^2}{2}} dz = \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-m}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-m}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{2} + \\
&\quad + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-m}{\sigma}} e^{-\frac{z^2}{2}} dz = \frac{1}{2} + \Phi\left(\frac{x-m}{\sigma}\right).
\end{aligned} \tag{2.7.4.5}$$

Для **определения вероятности попадания** случайной величины, подчиненной нормальному закону, **в интервал** воспользуемся свойством 5 функции распределения $F(x)$ и выражением функции нормального распределения через функцию Лапласа (2.7.4.5):

$$\begin{aligned}
P(\alpha < \xi < \beta) &= F(\beta) - F(\alpha) = \\
&= \left[\frac{1}{2} + \Phi\left(\frac{\beta-m}{\sigma}\right) \right] - \left[\frac{1}{2} + \Phi\left(\frac{\alpha-m}{\sigma}\right) \right] = \Phi\left(\frac{\beta-m}{\sigma}\right) - \Phi\left(\frac{\alpha-m}{\sigma}\right).
\end{aligned} \tag{2.7.4.6}$$

Примерами случайных величин, имеющих нормальный закон распределения, являются измерение длины, массы, времени, ошибки из-

мерения т. д.

Пример 36

Известно, что для некоторого измерительного устройства систематическая ошибка измерения дальности до объекта равна +20 м. Ошибки измерения распределены по нормальному закону со средним квадратическим отклонением, равным 75 м. Найти вероятность того, что полученное в результате измерения значение будет отличаться от истинного значения не более чем на 100 м.

Решение. Рассмотрим случайную величину ξ , характеризующую ошибку измерения дальности. Согласно условию эта случайная величина распределена по нормальному закону с параметрами $M[\xi] = 20$, $\sigma[\xi] = 75$. Определим искомую вероятность с помощью формулы (2.7.4.6):

$$\begin{aligned} P(|\xi| \leq 100) &= P(-100 < \xi < 100) = \Phi\left(\frac{100 - 20}{75}\right) - \\ &- \Phi\left(\frac{-100 - 20}{75}\right) = \Phi(1,071) - \Phi(-1,6) = 0,3577 + 0,4452 = \\ &= 0,8029. \end{aligned}$$

Пример 37

Случайное отклонение размера детали от номинала при изготовлении ее на данном станке является случайной величиной, распределенной по нормальному закону со средним квадратическим отклонением, равным 5 мк. Систематические отклонения размера изготовленной детали от номинала отсутствуют. Сколько необходимо изготовить деталей, чтобы с вероятностью не менее 0,9 среди них была хотя бы одна годная, если для годной детали допустимо отклонение от номинала не более чем на 2 мк?

Решение. Рассмотрим случайную величину ξ – отклонение размера детали от номинала. Согласно условию, $M[\xi] = 0$ [ммк], $\sigma[\xi] = 5$ [ммк]. Найдем вероятность события $A = \{\text{изготовление годной детали}\}$:

$$\begin{aligned} P(|\xi| \leq 2) &= P(-2 < \xi < 2) = \Phi\left(\frac{2 - 0}{5}\right) - \Phi\left(\frac{-2 - 0}{5}\right) = \\ &= \Phi(0,4) - \Phi(-0,4) = 0,15542 + 0,15542 = 0,31084 \approx 0,31. \end{aligned}$$

Теперь условие задачи можно рассматривать как последовательность n независимых испытаний, в каждом из которых с вероятностью $p = 0,31$ происходит событие A , и с вероятностью $q = 1 - p = 0,69$ со-

бытие A не происходит. Вероятность того, что среди n изготовленных деталей будет хотя бы одна годная, представляет собой вероятность наступления события A хотя бы один раз в серии из n независимых испытаний:

$$1 - q^n \geq 0,9; \quad q^n \leq 0,1; \quad n \ln q \geq \ln 0,1; \quad n \geq \frac{\ln 0,1}{\ln q} = \frac{\ln 0,1}{\ln 0,69} = 6,2.$$

Следовательно, $n \geq 7$, т. е. необходимо изготовить не менее 7 деталей, чтобы с вероятностью не менее 0,9 среди них была хотя бы одна годная.

Пример 38

Размер изготавливаемой на станке детали является случайной величиной, распределённой по нормальному закону с математическим ожиданием, равным 15 мм. Определить среднее квадратическое отклонение размера детали, если известно, что 95,44 % деталей имеют размер от 14 до 16 мм.

Решение. Согласно условию случайная величина ξ , определяющая размер изготавливаемой на станке детали, распределена по нормальному закону, причём $M[\xi] = 15$ [мм], $P(14 \leq \xi \leq 16) = 0,9544$. Определим искомую вероятность с помощью формулы (2.7.4.6):

$$P(14 < \xi < 16) = \Phi\left(\frac{16-15}{\sigma}\right) - \Phi\left(\frac{14-15}{\sigma}\right) = \Phi\left(\frac{1}{\sigma}\right) - \Phi\left(-\frac{1}{\sigma}\right).$$

Учитывая, что $\Phi(x)$ – нечётная функция, получим

$$P(14 < \xi < 16) = 2\Phi\left(\frac{1}{\sigma}\right) = 0,9544$$

$$\text{Отсюда } \Phi\left(\frac{1}{\sigma}\right) = 0,4772$$

По таблицам значений функции $\Phi(x)$ определяем, что $\Phi(2) = 0,4772$. Следовательно, $\sigma[\xi] = 0,5$, т. е. среднее квадратическое отклонение размера изготавливаемых на станке деталей равно 0,5.

3 Многомерные случайные величины

3.1 Понятие многомерной случайной величины

3.2 Функция распределения двумерной случайной величины

3.3 Функция плотности распределения непрерывной
двумерной случайной величины

3.4 Понятие независимости случайных величин

3.5 Числовые характеристики двумерной случайной величины

3.1 Понятие многомерной случайной величины

Результат вероятностного эксперимента может иногда характеризоваться не одним, а одновременно несколькими числами. Например, местоположение корабля в море – пара величин (ξ, η) , указывающая значения широты ξ и долготы η . Если при этом учитывать время τ , тогда мы имеем дело с трехмерной случайной величиной (ξ, η, τ) . Успеваемость студента в семестре – n -мерная случайная величина $(\xi_1, \xi_2, \dots, \xi_n)$, компоненты которой – оценки по каждой из n дисциплин.

Многомерной (n -мерной) случайной величиной называется функция $X(\omega)$, определенная на множестве элементарных событий Ω , которая каждому элементарному исходу ω ставит в соответствие n действительных чисел. Таким образом многомерная (n -мерная) случайная величина является совокупностью n одномерных величин (компонентов).

Все компоненты многомерной *дискретной* случайной величины – одномерные *дискретные* случайные величины. Все компоненты многомерной *непрерывной* случайной величины – одномерные непрерывные случайные величины. Многомерные *смешанные* случайные величины содержат как дискретные, так и непрерывные компоненты.

Основной характеристикой многомерной случайной величины является закон распределения, который (как и для одномерных величин) может быть задан таблично, графически или аналитически (функция распределения, функция плотности распределения и т. д.).

Пример 39

В ящике находится 5 шаров, пронумерованных цифрами «1», «1», «2», «2», «2». Последовательно извлекают два шара. Пусть случайная величина ξ – число на 1-м выбранном шаре, η – число на 2-м шаре. Требуется найти табличный закон распределения (матрицу распределения) двумерной случайной величины (ξ, η) . Определить вероятность

того, что второй шар будет иметь метку «1».

Решение. Определим вероятности возможных значений двумерной случайной величины (ξ, η) и заполним матрицу распределения (таблица 6):

$(\xi, \eta) = (1; 1)$:

$$P(\xi = 1; \eta = 1) = P(\{\xi = 1\} \cap \{\eta = 1\}) = P(\xi = 1)P(\eta = 1 | \xi = 1) = \\ = \left(\frac{2}{5}\right)\left(\frac{1}{4}\right) = 0,1$$

$(\xi, \eta) = (1; 2)$:

$$P(\xi = 1; \eta = 2) = P(\{\xi = 1\} \cap \{\eta = 2\}) = P(\xi = 1)P(\eta = 2 | \xi = 1) = \\ = \left(\frac{2}{5}\right)\left(\frac{3}{4}\right) = 0,3;$$

$(\xi, \eta) = (2; 1)$:

$$P(\xi = 2; \eta = 1) = P(\{\xi = 2\} \cap \{\eta = 1\}) = P(\xi = 2)P(\eta = 1 | \xi = 2) = \\ = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right) = 0,3;$$

$(\xi, \eta) = (2; 2)$:

$$P(\xi = 2; \eta = 2) = P(\{\xi = 2\} \cap \{\eta = 2\}) = P(\xi = 2)P(\eta = 2 | \xi = 2) = \\ = \left(\frac{3}{5}\right)\left(\frac{2}{4}\right) = 0,3.$$

Таблица 6 – Матрица распределения двумерной случайной величины

(ξ, η)	$\eta = 1$	$\eta = 2$
$\xi = 1$	0,1	0,3
$\xi = 2$	0,3	0,3

Очевидно, что сумма вероятностей всех значений многомерной случайной величины равна единице.

Вероятность того, что второй шар будет иметь метку «1», определим как сумму вероятностей в столбце « $\eta = 1$ » матрицы распределения. Таким образом, $P(\eta = 1) = 0,1 + 0,3 = 0,4$.

3.2 Функция распределения двумерной случайной величины

Функцией распределения двумерной случайной величины (ξ, η) или совместной функцией распределения случайных величин ξ и η назы-

вается функция $F_{\xi\eta}(x, y)$, равная вероятности того, что компонент ξ примет значение меньшее, чем x , а компонент η – значение меньшее, чем y ,

$$F_{\xi\eta}(x, y) = P(\xi < x, \eta < y) = P(\{\xi < x\} \cap \{\eta < y\}). \quad (3.2.1)$$

Таким образом, функция распределения двумерной случайной величины $F_{\xi\eta}(x, y)$ в точке (x, y) определяет вероятность, с которой двумерная случайная величина примет значение в нижнем левом квадранте относительно точки (x, y) (рисунок 29).

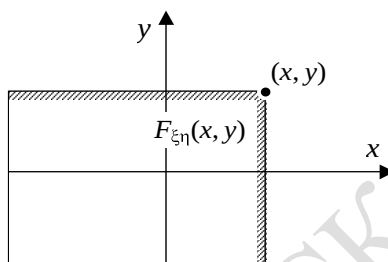


Рисунок 29 – Иллюстрация вероятностного смысла функции распределения двумерной случайной величины

Функция распределения двумерной случайной величины (ξ, η) обладает следующими свойствами (рисунок 30).

Свойство 1. $F_{\xi\eta}(x, y) \geq 0$, т. е. функция распределения двумерной случайной величины (ξ, η) – неотрицательная функция.

Свойство 2. Если $x_1 < x_2$, то $F_{\xi\eta}(x_1, y_1) \leq F_{\xi\eta}(x_2, y_1)$; если $y_1 < y_2$, то $F_{\xi\eta}(x_1, y_1) \leq F_{\xi\eta}(x_1, y_2)$.

Таким образом, функция $F_{\xi\eta}(x, y)$ – неубывающая функция каждого аргумента при условии, что другие аргументы фиксированы.

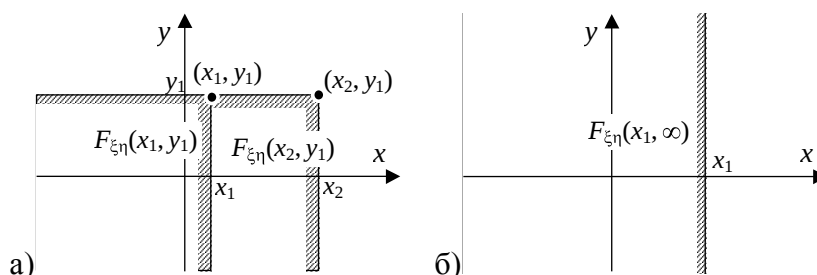


Рисунок 30 – Иллюстрация свойств функции распределения двумерной случайной величины

Свойство 3. $F_{\xi\eta}(x, +\infty) = F_{\xi}(x)$; $F_{\xi\eta}(+\infty, y) = F_{\eta}(y)$.

Таким образом, если один из аргументов функции распределения двумерной случайной величины (ξ, η) равен $+\infty$ (см. рисунок 30, б), то $F_{\xi\eta}(x, y)$ становится равной функции распределения одномерной случайной величины, соответствующей другому аргументу.

Свойство 4. $F_{\xi\eta}(-\infty, y) = 0$; $F_{\xi\eta}(x, -\infty) = 0$.

Таким образом, если хоть один из аргументов функции $F_{\xi\eta}(x, y)$ равен $-\infty$, то $F_{\xi\eta}(x, y) = 0$.

Свойство 5. $F_{\xi\eta}(+\infty, +\infty) = 1$

Таким образом, если все аргументы функции $F_{\xi\eta}(x, y)$ равны $+\infty$, то $F_{\xi\eta}(x, y) = 1$.

Свойство 6. По каждому из аргументов функция $F_{\xi\eta}(x, y)$ непрерывна слева.

Таким образом, $F_{\xi\eta}(x-0, y) = F_{\xi\eta}(x, y)$; $F_{\xi\eta}(x, y-0) = F_{\xi\eta}(x, y)$.

Замечание 1 – Функция распределения двумерной случайной величины представляет собой поверхность в пространстве. Причем в точке $(-\infty; -\infty)$ она равна нулю, а в точке $(\infty; \infty)$ – равна единице.

Замечание 2 – Функция распределения дискретной двумерной случайной величины – разрывная ступенчатая функция. Функция распределения непрерывной двумерной случайной величины непрерывна на всей числовой оси.

3.3 Функция плотности распределения непрерывной двумерной случайной величины

Функция распределения $F_{\xi\eta}(x, y)$ – наиболее универсальная форма закона распределения многомерных случайных величин как дискретных, так непрерывных и смешанных. Кроме этого, закон распределения непрерывных многомерных случайных величин может быть задан с помощью функции плотности распределения.

Двумерная случайная величина называется непрерывной, если ее функция распределения $F_{\xi\eta}(x, y)$ – непрерывная функция, дифференцируемая по каждому из аргументов, у которой существует вторая

смешанная производная $\frac{\partial^2 F_{\xi\eta}(x, y)}{\partial x \partial y}$.

Аналогично тому, как была определена функция плотности распределения одномерной случайной величины, определим *функцию плотности распределения* $F_{\xi\eta}(x, y)$ *непрерывной двумерной случайной величины* как предел отношения вероятности попадания значения случайной величины (ξ, η) в элементарный прямоугольник, примыкающий к точке (x, y) , к площади этого прямоугольника, когда оба его размера стремятся к нулю:

$$\begin{aligned} f_{\xi\eta}(x, y) &= \lim_{\substack{\Delta x \rightarrow 0 \\ \Delta y \rightarrow 0}} \frac{P(\{x < \xi < x + \Delta x\} \cap \{y < \eta < y + \Delta y\})}{\Delta x \cdot \Delta y} = \\ &= \frac{\partial^2 F_{\xi\eta}(x, y)}{\partial x \cdot \partial y}. \end{aligned} \quad (3.3.1)$$

При этом данный предел есть не что иное, как вторая смешанная производная функции распределения двумерной случайной величины (ξ, η) .

Функция плотности распределения непрерывной двумерной случайной величины (ξ, η) обладает следующими свойствами.

Свойство 1. $f_{\xi\eta}(x, y) \geq 0$, т. е. функция плотности распределения непрерывной двумерной случайной величины (ξ, η) — неотрицательная функция.

Свойство 2. $P(\{x_1 \leq \xi < x_2\} \cap \{y_1 \leq \eta < y_2\}) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f_{\xi\eta}(x, y) dx dy.$

Таким образом, вероятность попадания непрерывной двумерной случайной величины в произвольный прямоугольник, ограниченный точками (x_1, y_1) , (x_1, y_2) , (x_2, y_1) и (x_2, y_2) , определяется двойным интегралом функции плотности распределения $f_{\xi\eta}(x, y)$ по каждой из переменных на интервалах (x_1, x_2) и (y_1, y_2) соответственно.

Свойство 3. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\xi\eta}(x, y) dx dy = 1.$

Свойство 4. $\int_{-\infty}^x \int_{-\infty}^y f_{\xi\eta}(x, y) dx dy = F_{\xi\eta}(x, y).$

Свойство 5. $P(\{\xi = x\} \cap \{\eta = y\}) = 0$.

Таким образом, вероятность того, что непрерывная двумерная случайная величина (ξ, η) попадёт в точку (x, y) , равна нулю.

3.4 Понятие независимости случайных величин

При рассмотрении нескольких случайных величин (компонентов многомерной случайной величины) часто встречается задача установления факта зависимости величин. Например, влияет ли величина ξ – «масса поезда» – на величину η – «расход топлива локомотивом».

Независимыми называются случайные величины, закон распределения каждой из которых не зависит (не изменяется) от того, какое значение приняла другая случайная величина.

Зная совместный закон распределения многомерной случайной величины (в частности $F_{\xi\eta}(x, y)$), можно найти законы распределения ее компонентов. Однако совместный закон распределения многомерной случайной величины можно определить через законы распределения компонентов (т. е. одномерных случайных величин), только если эти компоненты независимы.

Рассмотрим определение функции распределения двумерной случайной величины (3.2.1)

$$F_{\xi\eta}(x, y) = P(\{\xi < x\} \cap \{\eta < y\}).$$

Если величины ξ и η независимы, то независимыми являются события $\{\xi < x\}$ и $\{\eta < y\}$. Следовательно, по теореме умножения вероятностей независимых событий:

$$F_{\xi\eta}(x, y) = P(\{\xi < x\} \cap \{\eta < y\}) = P(\xi < x)P(\eta < y) = F_{\xi}(x)F_{\eta}(y),$$

$$F_{\xi\eta}(x, y) = F_{\xi}(x)F_{\eta}(y). \quad (3.4.1)$$

Тождество (3.4.1) является необходимым и достаточным условием независимости двух случайных величин и называется теоремой умножения функций распределения независимых случайных величин.

Если ξ и η – непрерывные независимые случайные величины, то, дифференцируя левую и правую части равенства (3.4.1) по x и по y , получим

$$\frac{\partial^2 F_{\xi\eta}(x,y)}{\partial x \partial y} = \frac{\partial F_{\xi}(x)}{\partial x} \frac{\partial F_{\eta}(y)}{\partial y}.$$

Учитывая определения функции плотности распределения непрерывной двумерной случайной величины (3.3.1) и функции плотности распределения непрерывной одномерной величины (2.4.1), получаем равенство, называемое *теоремой умножения функций плотности распределения независимых величин* (необходимое и достаточное условие независимости непрерывных случайных величин):

$$f_{\xi\eta}(x,y) = f_{\xi}(x) f_{\eta}(y). \quad (3.4.2)$$

Если компоненты двумерной случайной величины (ξ, η) зависимы, то для нахождения совместного закона распределения недостаточно знать законы распределения компонентов: требуется знать так называемый условный закон распределения одной из них.

Условным законом распределения величины ξ называется закон ее распределения, вычисленный в предположении, что другая случайная величина (η) приняла определенное значение.

Функция распределения $F_{\xi\eta}(x,y)$ системы зависимых случайных величин может быть записана в виде, называемом *теоремой умножения функций распределения величин*:

$$\begin{aligned} F_{\xi\eta}(x,y) &= P(\{\xi < x\} \cap \{\eta < y\}) = \\ &= P(\xi < x) P(\eta < y | \xi < x) = F_{\xi}(x) F_{\eta}(y | \xi < x), \end{aligned} \quad (3.4.3)$$

где $F_{\eta}(y | \xi < x) = P(\eta < y | \xi < x)$ — условная функция распределения величины η при условии наступления события $\{\xi < x\}$.

На практике чаще используют другую форму условного закона распределения $F_{\eta}(y | \xi = x)$ или $f_{\eta}(y | \xi = x)$, т. е. когда величина ξ принимает фиксированное значение x . Тем более, что для непрерывных случайных величин справедлива следующая *теорема умножения плотностей*:

$$f_{\xi\eta}(x,y) = f_{\xi}(x) f_{\eta}(y | \xi = x), \quad (3.4.4)$$

т. е. функция плотности распределения непрерывной двумерной величины равна произведению функции плотности распределения од-

ной из них на условную плотность распределения другой при заданном значении первой.

3.5 Числовые характеристики двумерной случайной величины

Случайные величины полностью характеризуются законами распределения, но часто достаточным бывает знать лишь некоторые характерные значения, которые может принимать случайная величина, т. е. ее числовые характеристики. Для описания многомерных случайных величин используются числовые характеристики ее составляющих, а также параметры, характеризующие зависимость между компонентами многомерной величины. Одна из таких характеристик – корреляционный момент (ковариация).

Корреляционным моментом η двух случайных величин ξ и η называют математическое ожидание произведения отклонений этих величин от своих математических ожиданий:

$$\mu_{\xi\eta} = M[(\xi - M[\xi])(\eta - M[\eta])]. \quad (3.5.1)$$

Корреляционный момент имеет размерность, равную произведению размерностей величин ξ и η . Часто пользуются безразмерной характеристикой – *коэффициентом корреляции* случайных величин, который определяется по формуле

$$r_{\xi\eta} = \frac{\mu_{\xi\eta}}{\sigma[\xi]\sigma[\eta]}. \quad (3.5.2)$$

Коэффициент корреляции может принимать значения из отрезка $[-1; 1]$. Корреляционный момент и коэффициент корреляции характеризуют степень линейной зависимости между двумя величинами. Нулевое значение данных характеристик указывает на *отсутствие линейной зависимости* между исследуемыми величинами (при этом может существовать *нелинейная зависимость*). Равенство коэффициента корреляции $r_{\xi\eta}$ единице указывает на наличие *положительной линейной функциональной зависимости* между величинами ξ и η (с увеличением ξ величина η также увеличивается). Если же $r_{\xi\eta} = -1$, то между величинами ξ и η существует *отрицательная* (с увеличением ξ величина η уменьшается) *линейная функциональная зависимость*. Промежуточные значения коэффициента корреляции ($r_{\xi\eta} \in (-1; 0)$ или

$r_{\xi\eta} \in (0;1)$) указывают на тенденцию к наличию линейной зависимости между величинами ξ и η .

Пример 40

Рассмотрим систему двух дискретных случайных величин (ξ, η) , где случайная величина ξ – число поездов (в сутки), задержанных станцией A , случайная величина η – число поездов (в сутки), задержанных станцией B . Известна матрица распределения системы случайных величин (ξ, η) (таблица 7).

Таблица 7 – Матрица распределения двумерной случайной величины (ξ, η)

(ξ, η)	$\eta=0$	$\eta=1$	$\eta=2$
$\xi=0$	0,4	0,05	0,05
$\xi=1$	0,2	0,02	0,03
$\xi=2$	0,05	0,2	0

Требуется: 1) найти законы распределения случайных величин ξ и η ; 2) вычислить числовые характеристики (математическое ожидание, дисперсию и среднее квадратическое отклонение случайных величин ξ и η , математическое ожидание произведения $\xi\eta$, корреляционный момент η и коэффициент корреляции $r_{\xi\eta}$); 3) выяснить, являются ли случайные величины ξ и η зависимыми.

Решение. 1) Найдём законы распределения случайных величин.

а) Вероятности значений случайной величины ξ найдём по формуле $P(\xi = x_i) = \sum_{j=1}^3 p_{ij}, i = \overline{1,3}$.

$$P(\xi = 0) = 0,4 + 0,05 + 0,05 = 0,5;$$

$$P(\xi = 1) = 0,2 + 0,02 + 0,03 = 0,25;$$

$$P(\xi = 2) = 0,05 + 0,2 + 0 = 0,25.$$

Следовательно, ряд распределения случайной величины ξ можем записать в виде таблицы 8.

Таблица 8 – Матрица распределения случайной величины ξ

x_i	0	1	2
$P(\xi = x_i)$	0,5	0,25	0,25

б) Вероятности значений случайной величины η найдём по формуле

$$P(\eta = y_i) = \sum_{j=1}^3 p_{ij}, j = \overline{1,3}.$$

$$P(\eta = 0) = 0,4 + 0,2 + 0,05 = 0,65;$$

$$P(\eta = 1) = 0,05 + 0,02 + 0,02 = 0,27;$$

$$P(\eta = 2) = 0,05 + 0,03 + 0 = 0,08.$$

Следовательно, ряд распределения случайной величины η можем записать в виде таблицы 9.

Таблица 9 – Матрица распределения случайной величины η

y_j	0	1	2
$P(\eta = y_i)$	0,65	0,27	0,08

2) Вычислим числовые характеристики:

а) вычислим математическое ожидание случайной величины ξ , используя ряд распределения из таблицы 8:

$$M[\xi] = \sum_{i=1}^3 x_i p_i = 0 \cdot 0,5 + 1 \cdot 0,25 + 2 \cdot 0,25 = 0,75 \text{ поездов/сут};$$

б) вычислим математическое ожидание случайной величины η , используя ряд распределения из таблицы 9:

$$M[\eta] = \sum_{j=1}^3 y_j p_j = 0 \cdot 0,65 + 1 \cdot 0,27 + 2 \cdot 0,08 = 0,43 \text{ поездов/сут};$$

в) вычислим дисперсию случайной величины ξ , используя ряд распределения из таблицы 8:

$$D[\xi] = \sum_{i=1}^3 x_i^2 p_i - (M[\xi])^2 = 0^2 \cdot 0,5 + 1^2 \cdot 0,25 + 2^2 \cdot 0,25 - (0,75)^2 = 0,6875$$

(поездов/сут)²;

г) вычислим дисперсию случайной величины η , используя ряд распределения из таблицы 9:

$$D[\eta] = \sum_{j=1}^3 y_j^2 p_j - (M[\eta])^2 = 0^2 \cdot 0,65 + 1^2 \cdot 0,27 + 2^2 \cdot 0,08 - (0,43)^2 = 0,675$$

(поездов/сут)²;

д) средние квадратические отклонения случайных величин ξ и η :

$$\sigma[\xi] = \sqrt{D[\xi]} = \sqrt{0,6875} \approx 0,829 \text{ поездов/сут};$$

$$\sigma[\eta] = \sqrt{D[\eta]} = \sqrt{0,675} \approx 0,822 \text{ поездов/сут};$$

е) вычислим математическое ожидание произведения случайных величин ξ и η , пользуясь заданной матрицей распределения:

$$\begin{aligned} M[\xi \eta] &= \sum_{i=1}^3 \sum_{j=1}^3 x_i y_j p_{ij} = 0 \cdot 0 \cdot 0,4 + 0 \cdot 1 \cdot 0,05 + 0 \cdot 2 \cdot 0,05 + \\ &+ 1 \cdot 0 \cdot 0,2 + 1 \cdot 1 \cdot 0,02 + 1 \cdot 2 \cdot 0,03 + 2 \cdot 0 \cdot 0,05 + 2 \cdot 1 \cdot 0,2 + \\ &+ 2 \cdot 2 \cdot 0 = 0,48 \text{ (поездов/сут)}^2; \end{aligned}$$

ж) корреляционный момент двух случайных величин ξ и η

$$\mu_{\xi \eta} = M[\xi \eta] - M[\xi]M[\eta],$$

следовательно,

$$\mu_{\xi \eta} = 0,48 - 0,75 \cdot 0,43 = 0,1575 \text{ (поездов/сут)}^2;$$

з) коэффициент корреляции двух случайных величин ξ и η

$$r_{\xi \eta} = \frac{\mu_{\xi \eta}}{\sigma[\xi]\sigma[\eta]} = \frac{0,1575}{0,829 \cdot 0,822} \approx 0,231.$$

3) Поскольку коэффициент корреляции $r_{\xi \eta} \neq 0$, то случайные величины ξ и η зависимы.

Наличие зависимости между случайными величинами ξ и η можно проверить и на основании определения независимости: если $P(\xi = x_i, \eta = y_j) = P(\xi = x_i)P(\eta = y_j)$, то случайные величины ξ и η независимы. Если данное равенство нарушается, то случайные величины ξ и η зависимы.

Проверим выполнение нескольких неравенств для заданной матрицы распределения:

$$P(\xi = 0, \eta = 0) = 0,43; \quad P(\xi = 0) = 0,5; \quad P(\eta = 0) = 0,65;$$

$$0,43 \neq 0,5 \cdot 0,65;$$

$$P(\xi = 0, \eta = 1) = 0,05; \quad P(\xi = 0) = 0,5; \quad P(\eta = 1) = 0,27;$$

$$0,05 \neq 0,5 \cdot 0,27;$$

$$P(\xi = 0, \eta = 2) = 0,05; \quad P(\xi = 0) = 0,5; \quad P(\eta = 2) = 0,08;$$

$$0,05 \neq 0,5 \cdot 0,08;$$

$$P(\xi = 1, \eta = 0) = 0,2; \quad P(\xi = 1) = 0,25; \quad P(\eta = 0) = 0,65;$$

$$0,2 \neq 0,25 \cdot 0,65;$$

и т. д.

Выполненные вычисления подтверждают наличие зависимости между случайными величинами ξ и η .

Вывод. Среднее число поездов в сутки, задержанных станцией A , равно 0,75; среднее число поездов в сутки, задержанных станцией B , равно 0,43; корреляционный момент, характеризующий разброс точек (ξ, η) вокруг точки $(M[\xi], M[\eta])$, равен 0,1575; коэффициент корреляции равен 0,231, что говорит об очень слабой линейной зависимости между случайными величинами ξ и η .

4 Математическая статистика

4.1. Выборочный метод; эмпирическая функция распределения; статистическое оценивание параметров распределения методом моментов

4.2 Проверка статистических гипотез

4.3 Элементы теории корреляции; линейная корреляция

4.1. Выборочный метод; эмпирическая функция распределения; статистическое оценивание параметров распределения методом моментов

Предположим, что для изучения некоторого количественного признака ξ из генеральной совокупности извлечена выборка объема n

$$x_1, x_2, \dots, x_n \quad (4.1.1)$$

Эмпирической (или выборочной) функцией распределения называется функция действительного переменного

$$F^*(x) = \frac{v(x)}{n}, \quad (4.1.2)$$

где $v(x)$ – число элементов выборки, меньших x .

Если признак ξ имеет непрерывное распределение, то по выборке (4.1.1) строят **интервальный статистический ряд**, разбивая интервал, содержащий все элементы выборки (4.1.1) на ряд частичных интервалов шириной h и, подсчитывая n_i – частоту элементов выборки, попавших i -й интервал

Таблица 10 – Интервальный статистический ряд

Интервалы	$[\alpha_1, \alpha_2)$	$[\alpha_2, \alpha_3)$	...	$[\alpha_r, \alpha_{r+1})$
Частоты	n_1	n_2	...	n_r

где $\alpha_{i+1} - \alpha_i = h, \quad i = 1, 2, \dots, r; \quad \sum_{i=1}^r n_i = n$

Гистограммой интервального статистического ряда (таблица 10) называется ступенчатая фигура, состоящая из прямоугольников, построенных в прямоугольной системе координат так, что их основаниями служат интервалы $[\alpha_i, \alpha_{i+1})$, $i=1, 2, \dots, r$, отложенные на оси абсцисс, а высоты равны $r_i/(nh)$ соответственно i -му основанию $i=1, 2, \dots, r$ (рисунок 31).

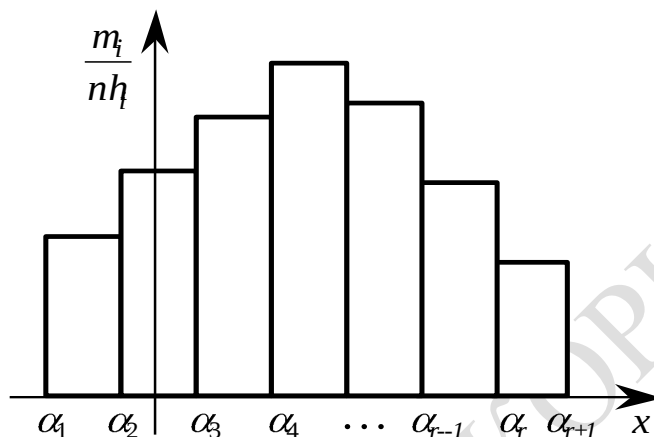


Рисунок 31

Кривая 1, проходящая через середины ступеней гистограммы дает приближенное представление о кривой распределения признака ξ и называется эмпирической кривой распределения. Площадь гистограммы равна единице. Часто из некоторых предпосылок бывает известен тип закона распределения признака ξ , но неизвестны параметры этого распределения. Например, известно, что плотность распределения признака ξ задана функцией $f(x, \theta_1, \dots, \theta_k)$, где $\theta_1, \dots, \theta_k$ — неизвестные параметры. Ставится задача нахождения статистических оценок $\theta_1^*, \theta_2^*, \dots, \theta_k^*$ параметров $\theta_1, \theta_2, \dots, \theta_k$ по выборке (4.1.1) наблюдений над ξ . Таким образом, оценки — функции от выборочных значений.

Одним из методов построения таких оценок является **метод моментов**, который основывается на близости теоретических и эмпирических моментов. Он состоит в следующем:

По формуле плотности распределения $f(x, \theta_1, \dots, \theta_k)$ находим k первых начальных теоретических моментов признака ξ

$$\mu_m = M[\xi^m] = \int_{-\infty}^{+\infty} x^m f(x, \theta_1, \dots, \theta_k) dx, \quad m=1, 2, \dots, k, \quad (4.1.3)$$

затем по выборке (4.1.1) вычисляем соответственные выборочные моменты

$$\mu_m^* = \frac{1}{2} \sum_{i=1}^m x_i^m, \quad m=1,2,\dots,k, \quad (4.1.4)$$

Приравнивая каждый теоретический момент соответствующему выборочному, получаем систему из k уравнений

$$\mu_m(\theta_1, \dots, \theta_k) = \mu_m^*, \quad m=1,2,\dots,k, \quad (4.1.5)$$

с неизвестными $\theta_1, \dots, \theta_k$.

Решение $\theta_1^*, \dots, \theta_k^*$ системы (4.1.5), зависящее от элементов выборки (4.1.1), принимают в качестве статистических оценок параметров $\theta_1, \dots, \theta_k$.

Величина
$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4.1.6)$$

называется **выборочной средней** и служит статистической оценкой для $M[\xi]$.

Величина
$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (4.1.7)$$

называется **несмещенной выборочной дисперсией** и служит статистической оценкой для $D[\xi]$.

$\bar{x}$ и s^2 – точечные оценки параметров $M[\xi]$ и $D[\xi]$.

В математической статистике используются еще и интервальные оценки. **Доверительным** называют интервал, концы которого зависят от выборочных значений и которые с заданной **доверительной вероятностью** покрывает оцениваемый параметр.

Предположим, что выборка (4.1.1) взята из нормально распределенной генеральной совокупности признака ξ с неизвестными параметрами $a = M[\xi]$ и $\sigma^2 = D[\xi]$.

Связь между доверительным интервалом, покрывающим параметр a , с доверительной вероятностью $1-\varepsilon$ задается при этих условиях формулой

$$P\left\{\bar{x} - t \frac{s}{\sqrt{n}} < a < \bar{x} + t \frac{s}{\sqrt{n}}\right\} = 1 - \varepsilon, \quad (4.1.8)$$

где t – критическая точка распределения Стьюдента с $n-1$ степенью свободы, соответствующая уровню значимости $\frac{\varepsilon}{2}$. Для построения доверительного интервала, покрывающего параметр σ^2 с доверительной вероятностью $1-\varepsilon$, используется другая формула

$$P\left\{\frac{(n-1)S^2}{\gamma_2} < \sigma^2 < \frac{(n-1)S^2}{\gamma_1}\right\} = 1 - \varepsilon \quad (4.1.9)$$

где числа γ_1 и γ_2 находятся из уравнений

$$\int_{\gamma_1}^{+\infty} k_{n-1}(x) dx = 1 - \frac{\varepsilon}{2} \quad (4.1.10)$$

$$\int_{\gamma_2}^{+\infty} k_{n-1}(x) dx = 1 - \frac{\varepsilon}{2} \quad (4.1.11)$$

по таблицам χ^2 – распределения с $(n-1)$ степенью свободы, то есть γ_1 и γ_2 – критические точки χ^2 – распределения с $(n-1)$ соответствующие уровням значимости $1 - \frac{\varepsilon}{2}$ и $\frac{\varepsilon}{2}$.

Пример 41

Приведенные ниже данные о ценах на 100 видов товаров (в у. е.) записаны в случайном порядке

126	93	114	126	81	140	129	114	138	140
151	171	152	139	97	163	117	158	125	129
116	129	108	124	105	137	106	140	137	116
120	122	145	136	169	122	232	97	123	112
144	101	148	126	124	125	117	142	133	119
125	170	138	100	80	124	108	90	83	86
163	109	100	125	160	138	144	137	111	128
87	111	130	99	109	165	56	152	115	104
111	107	131	124	162	88	94	92	132	125
112	150	102	82	113	158	107	134	157	101

Используя эти данные, необходимо:

- 1) получить выборку, выбрав 20 значений,
- 2) записать эмпирическую функцию распределения;
- 3) построить интервальный вариационный ряд с шириной интервала 20 у.е;
- 4) построить гистограмму и эмпирическую кривую распределения;
- 5) предполагая, что генеральная совокупность имеет нормальное

распределение с плотностью распределения $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}$ найти

методом моментов оценки a и σ^2 .

Решение.

1) Произведем выборку. Получаем 125, 170, 151, 195, 173, 117, 190, 133, 102, 151, 94, 94, 114, 153, 109, 148, 101, 139, 110, 144.

2) Для построения $F^*(x)$ запишем элементы выборки в порядке возрастания: 94, 94, 101, 102, 109, 110, 114, 117, 125, 133, 139, 144, 148, 151, 151, 153, 170, 173, 190, 196.

$$F^*(x) = \frac{\nu(x)}{n} = \begin{cases} 0, & \text{если } x \leq 94 \\ 2/20, & \text{если } 94 < x \leq 101 \\ 3/20, & \text{если } 101 < x \leq 102 \\ 4/20, & \text{если } 102 < x \leq 109 \\ 5/20, & \text{если } 109 < x \leq 110 \\ 6/20, & \text{если } 110 < x \leq 114 \\ 7/20, & \text{если } 114 < x \leq 117 \\ 8/20, & \text{если } 117 < x \leq 125 \\ 9/20, & \text{если } 125 < x \leq 133 \\ 10/20, & \text{если } 133 < x \leq 139 \\ 11/20, & \text{если } 139 < x \leq 144 \\ 12/20, & \text{если } 144 < x \leq 148 \\ 13/20, & \text{если } 148 < x \leq 151 \\ 15/20, & \text{если } 151 < x \leq 153 \\ 16/20, & \text{если } 153 < x \leq 170 \\ 17/20, & \text{если } 170 < x \leq 173 \\ 18/20, & \text{если } 173 < x \leq 190 \\ 19/20, & \text{если } 190 < x \leq 196 \\ 1, & \text{если } 196 < x \end{cases}$$

3) Составляем вариационный ряд с шириной интервала $h = 20$ (y.e)

Интер-валы	[90,110)	[110,130)	[130,150)	[150,170)	[170,190)	[190,210]
Часто-ты	5	4	4	3	2	2

$$\sum_{i=1}^6 n_i = 5 + 4 + 3 + 2 + 2 = 20.$$

4) Строим гистограмму и эмпирическую (выборочную) кривую распределения (рисунок 32), откладывая на оси абсцисс интервалы, а по оси координат $\frac{n_i}{(nh)}$:

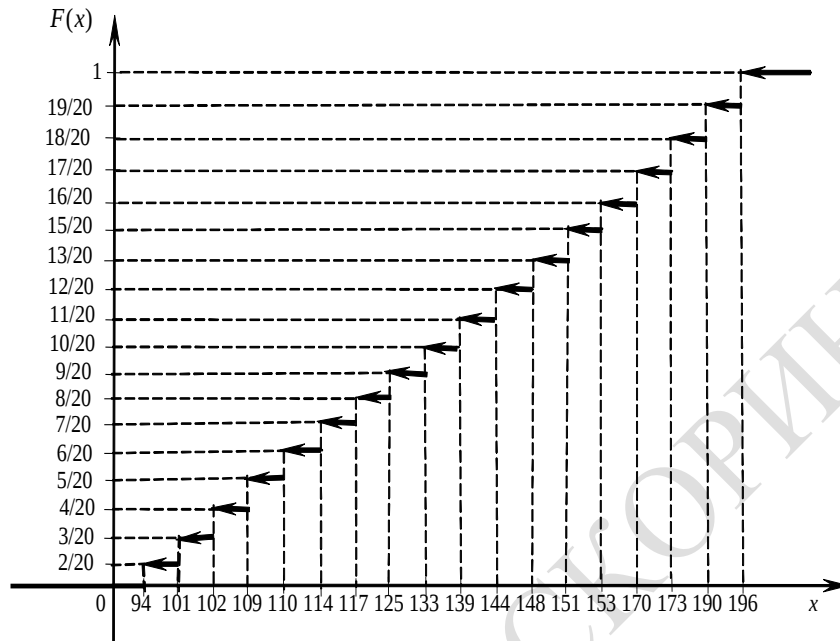


Рисунок 32

5) При изучении нормального закона доказывается, что в плотности $f(x, a, \sigma^2)$ параметр a есть математическое ожидание, а параметр σ^2 — дисперсия:

$$a = M\xi = \mu_1;$$

$$\sigma^2 = D\xi = M(\xi - M\xi)^2 = M\xi^2 - (M\xi)^2 = \mu_2 - \mu_1^2.$$

Приравнивая теоретические моменты μ_1 и μ_2 выборочным μ_1^* и μ_2^* , получаем:

$$a^* = \frac{1}{n} \sum_{i=1}^n x_i; \quad (4.1.12)$$

$$(\sigma^2)^* = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \quad (4.1.13)$$

Используя обозначение (4.1.6), имеем

$$a^* = \bar{x} \quad (4.1.14)$$

Преобразовывая (4.1.9), находим

$$(\sigma^2)^* = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - 2(\bar{x})^2 + (\bar{x})^2 =$$

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

Итак, $(\sigma^2)^* = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$ (4.1.15)

Полученная нами методом моментов оценка (4.1.15) называется смещенной выборочной дисперсией.

По выборке вычислим оценки (4.1.14) и (4.1.15):

$$\begin{aligned} \bar{x} &= \frac{1}{20} - (125 + 170 + 151 + 196 + 173 + 117 + 190 + 133 + 102 + 151 + \\ &+ 94 + 94 + 114 + 153 + 109 + 148 + 101 + 139 + 110 + 144) = \\ &= \frac{1}{20} \cdot 2714 = 135,7 \end{aligned}$$

$$\begin{aligned} (\sigma^2)^* &= -\frac{1}{20} (10,7^2 + 34,3^2 + 15,3^2 + 60,3^2 + 37,3^2 + 18,7^2 + 54,3^2 + \\ &+ 2,7^2 + 33,7^2 + 15,3^2 + 41,7^2 + 41,7^2 + 21,7^2 + 17,3^2 + 24,7^2 + 12,3^2 + \\ &+ 34,7^2 + 3,3^2 + 45,7^2 + 8,3^2) = \frac{1}{20} \cdot 19634,45 = 981,72 \approx 981,7 \end{aligned}$$

6) По формуле (4.1.8):

$$P \left\{ \bar{x} - t \frac{S}{\sqrt{n}} < a < \bar{x} + t \frac{S}{\sqrt{n}} \right\} = 1 - \varepsilon.$$

По условию задачи $n=20$, $1-\varepsilon=0,95$.

В 5) мы вычислили $\bar{x}=135,7$. Для вычисления несмещенной выборочной дисперсии S^2 , сравним формулы (4.1.13) и (4.1.15). Видим, что

$$S^2 = \frac{n}{n-1} (\sigma^2)^*.$$

Подставляя $(\sigma^2)^*$ в эту формулу, имеем

$$S^2 = \frac{20}{19} \cdot 981,7 \approx 1033,4.$$

Отсюда $S \approx 32,14$.

По таблицам распределения Стьюдента с $n-1=19$ степенью свободы находим t при доверительной вероятности 0,95.

$$t = 2,0930 \approx 2,1.$$

Выписываем из (4.1.8) доверительный интервал:

$$J_a = \left(135,7 - 2,1 \cdot \frac{32,14}{\sqrt{20}} \quad 135,7 + 2,1 \cdot \frac{32,14}{\sqrt{20}} \right) = (140,71; 150,69),$$

покрывающий параметр a с вероятностью 0,95.

По формуле (4.1.9)

$$P\left\{\frac{(n-1)S^2}{\gamma_2} < \sigma^2 < \frac{(n-1)S^2}{\gamma_1}\right\} = 1 - \varepsilon.$$

γ_1 и γ_2 находим по таблице распределения χ^2 с $n-1=19$ степенью свободы, используя формулы (4.1.10) и (4.1.11) с $\varepsilon = 1 - 0,95 = 0,05$: $\gamma_1 = 8,907 \approx 8,9$ $\gamma_2 = (32,852 \approx 32,8$.

Выписываем доверительный интервал

$$J_{\sigma^2} = \left(\frac{19 \cdot 10334}{32,8}; \frac{19 \cdot 10334}{8,9} \right) = (598,6; 22061),$$

покрывающий параметр σ^2 с вероятностью 0,95.

4.2 Проверка статистических гипотез

При изучении генеральной совокупности часто необходимо знать закон ее распределения. Если закон распределения неизвестен, но имеются основания предположить, что он имеет определенный вид, выдвигают гипотезу: генеральная совокупность имеет функцию распределения $F_0(x)$. Таким образом, в этой гипотезе речь идет о виде предполагаемого распределения.

Возможен случай, когда закон распределения генеральной совокупности известен, а его параметры неизвестны. Если имеются основания предположить, что неизвестный параметр θ равен определенному значению θ_0 , выдвигают гипотезу: $\theta = \theta_0$. Таким образом, в этой гипотезе речь идет о предполагаемой величине неизвестного параметра известного распределения.

Наряду с выдвинутой гипотезой обычно рассматривают противоречащую ей гипотезу. Если выдвинутая гипотеза будет отвергнута, то имеет место противоречащая гипотеза. По этой причине эти гипотезы целесообразно различать. Нулевой (основной) называют выдвинутую гипотезу H_0 . Конкурирующей (альтернативной) называют гипотезу H_1 , которая противоречит нулевой гипотезе. Например, если нулевая гипотеза состоит в предположении, что математическое ожидание a нормального распределения равно 10, то конкурирующая гипотеза, в частности, может состоять в предположении того, что $a \neq 10$. Коротко это записывается так: $H_0: a = 10$, $H_1: a \neq 10$. Различают гипотезы, которые содержат только одно и более одного предположений.

Простой называют гипотезу, содержащую только одно предположение. Например, $H_0: a = 10$ — простая гипотеза.

Сложной называют гипотезу, которая состоит из конечного или бесконечного числа простых гипотез. Например, $H_0: F(x) \in F_0(x) | F_0(a)=0, F_0(b)=1$. Эта гипотеза состоит в том, что ф.р. генеральной совокупности принадлежит множеству ф.р. $F_0(x)$, таких, что $F_0(a)=0$, а $F_0(b)=1$.

Выдвинутая нулевая гипотеза может быть правильной или неправильной. В итоге статистической проверки нулевой гипотезы в двух случаях может быть принято неправильное решение, т.е. могут быть допущены ошибки двойного рода.

Ошибка первого рода состоит в том, что нулевая гипотеза будет отвергнута, в то время как в действительности она правильная.

Ошибка второго рода состоит в том, что нулевая гипотеза будет принята, в то время как в действительности она неправильная.

Для проверки нулевой гипотезы используют специально подобранную случайную величину, точное или приближенное распределение которой известно. Эту величину обозначают различными буквами в зависимости от закона ее распределения и называют **статистическим критерием** или просто критерием.

После выбора определенного критерия множество всех возможных его значений разбивают на два непересекающихся подмножества; одно из них содержит значения критерия, при которых нулевая гипотеза отвергается, а другое – при которых она принимается.

Критической областью называют совокупность значений критерия, при которых нулевую гипотезу отвергают.

Основной принцип проверки статистических гипотез таков: по данным выборки вычисляется значение критерия, которое называется **наблюдаемым**; если наблюдаемое значение критерия принадлежит критической области, нулевую гипотезу отвергают; в противном случае нулевую гипотезу принимают.

4.2.1 Критерий χ^2 Пирсона

На практике часто возникает следующая задача. Пусть в результате какого-либо эксперимента получена выборка

$$x_1, x_2, \dots, x_n \quad (4.2.1.1)$$

объема n с функцией распределения $F_\xi(x) = P_{x_i < x}$. Нас интересует гипотеза, состоящая в том, что функция распределения $F_\xi(x)$ совпадает с некоторой фиксированной ф.р. $F_0(x)$. Задача проверки статистиче-

ской гипотезы $H_0: F_\xi(x) = F_0(x)$ состоит в том, чтобы решить согласуются ли с ней значения x_i выборки (4.2.1.1). Для решения этой задачи поступим следующим образом. Разделим точками $-\infty = z_0 < z_1 < z_2 < \dots < z_{r-1} < z_r = +\infty$ всю прямую на r интервалов $[z_{k-1}, z_k)$.

Обозначим $P_k = F_0(z_k) - F_0(z_{k-1}) = P\{z_{k-1} \leq x_i < z_k\}$ – вероятность попадания x_i в интервал $[z_{k-1}, z_k)$ в случае, когда наша гипотеза справедлива. По выборке (4.2.1.1) определим числа v_k , $k=1, 2, \dots, r$, где v_k – число элементов x выборки (4.2.1.1), попавших в интервал $[z_{k-1}, z_k)$. Таким образом, мы свели задачу к более простой. Имеется n независимых испытаний с r исходами. Вероятность k -го исхода равна p_k . Набор вероятностей исходов $p_1, p_2, \dots, p_r$, $\sum_{k=1}^r p_k = 1$, определяется первоначальной статистической гипотезой. Случайные величины $v_1, v_2, \dots, v_r$, $\sum_{k=1}^r v_k = n$, определяются по выборке (4.2.1.1).

Если значения v_k соответствует вероятностям p_k , то разности $\frac{v_k}{n} - p_k$ должны быть малы. Таким образом, в качестве критерия проверки нулевой гипотезы примем случайную величину:

$$\chi^2 = \sum_{k=1}^r \frac{(v_k - np_k)^2}{np_k}, \quad (4.2.1.2)$$

которую будем называть χ^2 – статистикой Пирсона.

Справедливо следующее утверждение: при любом $x > 0$ $\lim_{n \rightarrow \infty} P\{\chi^2 < x\} = K_{r-1}(x)$, где $K_{r-1}(x)$ – функция распределения χ^2 с $(r-1)$ степенью свободы.

Этот результат используется следующим образом. Зададимся каким-либо малым значением вероятности α , которое будем называть **уровнем значимости критерия**. Заменим при больших n предельное соотношение приближенным равенством $P\{\chi^2 < x\} \approx K_{r-1}(x)$. Выбирая $x = \chi_{kp}^2$ таким, чтобы $K_{r-1}(\chi_{kp}^2) = 1 - \alpha$, получаем, что в случае, когда проверяемая гипотеза справедлива, событие $\{\chi^2 \geq \chi_{kp}^2\}$ может произойти лишь с малой вероятностью, которая приблизительно равна α . Обычно полагают $\alpha = 0,05$ или $\alpha = 0,01$. Если гипотеза верна, то маловероятное событие $\{\chi^2 \geq \chi_{kp}^2\}$ практически невозможно. Если оно произошло, то

будем считать, что гипотеза неверна. Если же $\chi^2 < \chi_{кр}^2$, то будем говорить, что выборочные данные не противоречат гипотезе $F_{\xi}(x) = F_0(x)$.

Замечание 1 Если функция распределения $F_x(x, \theta_1, \dots, \theta_m)$ элементов выборки зависит от неизвестных параметров $\theta_1, \dots, \theta_m$, то эти параметры следует оценить по выборке с помощью методов минимума χ^2 или максимума правдоподобности и в этом случае предельное распределение величины χ^2 является χ^2 – распределением, но уже с $(r-m-1)$ степенями свободы.

Замечание 2 Если $F_x(x, \theta_1, \dots, \theta_m)$ имеет нормальное распределение, то параметры можно оценивать с помощью метода моментов.

Каждый интервал должен содержать не менее 5-6 выборочных значений; малочисленные интервалы следует объединять, суммируя частоты.

Пример 42

Получены следующие данные о размере мужской обуви, проданной магазином в течение дня

Размер обуви, x_i	37	38	39	40	41	42	43	44
Количество проданных пар, n_i	1	4	14	37	35	20	8	3

Проверить гипотезу о том, что случайная величина ξ – размер обуви мужчины, имеет нормальное распределение, предварительно оценив по выборке математическое ожидание и дисперсию.

Решение. Оценим сначала параметры распределения, используя метод моментов. Для этого найдем выборочные среднее и дисперсию.

$$\bar{x} = \frac{1}{122} \sum_{i=1}^8 n_i x_i = 40,705; (\sigma^2)^* = 1,749, \quad \sigma^* = \sqrt{(\sigma^2)^*} = 1,322$$

Принимаем эти величины соответственно за математическое ожидание и дисперсию случайной величины ξ . Таким образом, нам нужно проверить гипотезу $H_0: F_{\xi}(x) = F_0(x)$, где $F_0(x)$ – функция нормального распределения с параметрами $(40,705; 1,749)$. Для этого вычислим величину χ^2 . Разобьем множество значений случайной величины ξ на 5 интервалов: $(-\infty; 39,5)$, $[39,5; 40,5)$, $[40,5; 41,5)$, $[41,5; 42,5)$, $[42,5; +\infty)$ и подсчитаем число выборочных значений, попадающих в каждый интервал: $v_1 = 19$, $v_2 = 37$, $v_3 = 35$, $v_4 = 20$, $v_5 = 11$.

Далее вычисляем вероятности p_i – попадания с.в. ξ в i -й интервал. Принимая во внимание, что при справедливости гипотезы с.в. ξ распределенная нормально с параметрами (40,705; 1,749), для вычисления вероятностей p_i , получим следующую формулу:

$$p_i = P \alpha_i \leq \xi < \beta_i = \Phi\left(\frac{\beta_i - 40,705}{1,322}\right) - \Phi\left(\frac{\alpha_i - 40,705}{1,322}\right), \text{ где } \Phi(x) - \text{функция Лапласа.}$$

Используя приведенную выше формулу и таблицу значений функции Лапласа, получим:

$$p_1 = \Phi\left(\frac{39,5 - 40,705}{1,322}\right) - \Phi(-\infty) = \Phi(-\infty) - \Phi(0,9115) = 0,5 - 0,3186 = 0,1814;$$

$$p_2 = \Phi\left(\frac{40,5 - 40,705}{1,322}\right) - \Phi\left(\frac{39,5 - 40,705}{1,322}\right) = \Phi(-0,155) - \Phi(-0,9115) = \Phi(0,9115) - \Phi(0,155) = 0,3186 - 0,0636 = 0,255;$$

$$p_3 = \Phi\left(\frac{41,5 - 40,705}{1,322}\right) - \Phi\left(\frac{40,5 - 40,705}{1,322}\right) = \Phi(0,6) + \Phi(0,155) = 0,2257 + 0,0636 = 0,2893;$$

$$p_4 = \Phi\left(\frac{42,5 - 40,705}{1,322}\right) - \Phi\left(\frac{41,5 - 40,705}{1,322}\right) = \Phi(1,36) - \Phi(0,6) = 0,4131 - 0,2257 = 0,1874;$$

$$p_5 = \Phi(+\infty) - \Phi\left(\frac{42,5 - 40,705}{1,322}\right) = 0,5 - \Phi(1,36) = 0,5 - 0,4131 = 0,0869$$

По формуле (4.2.1.2) вычисляем величину χ^2 :

$$\begin{aligned} \chi^2 &= \frac{(19 - 122 \cdot 0,1814)^2}{122 \cdot 0,1814} + \frac{(1 - 122 \cdot 0,255)^2}{122 \cdot 0,255} + \frac{(35 - 122 \cdot 0,2893)^2}{122 \cdot 0,2893} + \\ &+ \frac{(20 - 122 \cdot 0,1874)^2}{122 \cdot 0,1874} + \frac{(11 - 122 \cdot 0,0869)^2}{122 \cdot 0,0869} = 0,4429 + 1,1151 + \\ &+ 0,0025 + 0,3585 + 0,0150 = 1,934 \end{aligned}$$

Зададим уровень значимости $\alpha = 0,01$. Принимая во внимание замечание 1, найдем критическую точку χ^2 распределения, отвечающую уровню значимости $\alpha = 0,01$ и числу степеней свободы

$$k = 5 - 1 - 2 = 2, \text{ то есть } \chi_{kp}^2 = \chi^2(0,01; 2) = 9,2.$$

Поскольку $\chi^2 < \chi_{kp}^2$, то можно считать, что выборочные данные не противоречат нашей гипотезе о нормальности распределения размера обуви мужчин.

4.3 Элементы теории корреляции; линейная корреляция

При изучении влияния одних признаков явлений на другие из ряда признаков, характеризующих данное явление, выделяют два признака – факториальный и результативный. Необходимо установить, какой из признаков является факториальным и какой результативным.

Пример. Себестоимость промышленной продукции отдельного предприятия зависит от многих факторов, в том числе от объема продукции на данном предприятии. Себестоимость продукции в этом случае выступает как результативный признак, а объем продукции как факториальный.

Одной из основных задач теории корреляции является выявление на основе экспериментальных данных того, как изменяется результативный признак в связи с изменением данного фактора. Эта задача решается нахождением уравнения связи. Под уравнением связи будем понимать функциональную зависимость между результативным и факториальными признаками. Применение той или иной функции в качестве уравнения связи разграничивает корреляцию на линейную, параболическую и др.

Рассмотрим уравнение связи для линейной зависимости от одного признака. Такое уравнение называют уравнением линейной регрессии. Выборочное уравнение линейной регрессии Y на X имеет вид:

$$y - \bar{y} = r_B \frac{S_y}{S_x} (x - \bar{x}), \text{ где}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i; \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i; \quad S_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2};$$

$$S_y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2};$$

а $r_B = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) / (S_x S_y)$ есть выборочный коэффициент корреляции.

При проведении корреляционного анализа прежде всего возникает вопрос о реальности связи, т.е. о том, является ли полученный из наблюдений коэффициент корреляции значимым и не объясняется ли получение его случайностями выборки. Таким образом, требуется проверить гипотезу $H_0: g = 0$, где g – коэффициент корреляции признаков X и Y . В качестве критерия служит следующая случайная величина

на $D = r_B \sqrt{n} / (1 - r_B^2)$. Если гипотеза H_0 верна, то случайная величина D распределена нормально с параметрами $(0, 1)$. Следовательно, критическую область выбирают так, чтобы $P\{|D| \leq t_\alpha\} = \alpha$, где α - уровень значимости, а t_α находят из уравнения $\Phi(t_\alpha) = \frac{1-\alpha}{2}$ по таблицам функции Лапласа.

Если вычисленная по данным выборки величина $|D| \leq t_\alpha$, то гипотезу $H_0 : g = 0$ отвергают. В противном случае гипотезу принимают.

Пример 43

Рассмотрим зависимость между размером предприятия по стоимости основных средств и себестоимостью единицы продукции. Факторальным признаком у нас является стоимость основных средств, а результативным себестоимость единицы продукции.

X (млн. руб.)	0,5	1,5	2,5	3,5	4,5	5,5
Y	15	11	12	12	9	10

Решение. Найдем числовые характеристики случайных величин X , Y , найденные по выборке:

$$\bar{x} = \frac{1}{6}(0,5 + 1,5 + 2,5 + 3,5 + 4,5 + 5,5) = \frac{18}{6} = 3;$$

$$\bar{y} = \frac{1}{6}(15 + 11 + 12 + 12 + 9 + 10) = \frac{69}{6} = 11,5;$$

$$S_x \sqrt{\frac{1}{5} \sum_{i=1}^6 (x_i - \bar{x})^2} = 1,707 \sqrt{\frac{6}{5}}; S_y = \sqrt{\frac{1}{5} \sum_{i=1}^6 (y_i - \bar{y})^2} = 1,892 \sqrt{\frac{6}{5}}$$

Для вычисления выборочного коэффициента корреляции вычислим $\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$, предварительно преобразовав ее:

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} n \sum_{i=1}^n x_i y_i - \frac{n}{n-1} \bar{x} \bar{y}.$$

Таким образом,

$$\frac{1}{5} \sum_{i=1}^6 (x_i - 3)(y_i - 11,5) = \frac{1}{5} (0,5 \cdot 15 + 1,5 \cdot 11 + 2,5 \cdot 12 + 3,5 \cdot 12 +$$

$$+ 4,5 \cdot 9 + 5,5 \cdot 10) - \frac{6}{5} \cdot 3 \cdot 11,5 = -2,583 \cdot \frac{6}{5}.$$

Таким образом, выборочный коэффициент корреляции равен:

$$r_B = \frac{-2,583}{1,707 \cdot 1,892} \approx -0,79.$$

Следовательно, выборочное уравнение линейной регрессии Y на X будет иметь вид:

$$y - 11,5 = -0,79 \frac{1,892}{1,707} (x - 3).$$

$$y = -0,869 \cdot 10 + 14,107 = 5,417.$$

Т.е. прогнозируемая стоимость будет равна 5,417.

Мы нашли выборочный коэффициент корреляции и на основании его строили уравнение линейной регрессии, т.е. уравнение связи.

Или окончательно

$$y = -0,869x + 14,107$$

Полученной зависимостью можно воспользоваться для определения ожидаемой себестоимости единицы изделия на предприятии со стоимостью основных фондов в 10 млн.руб.

Выясним вопрос о реальности связи, т.е. является ли полученный по наблюдениям коэффициент корреляции значимым? Следовательно, нам предстоит проверить гипотезу $H_0: g = 0$.

$$\text{Вычислим величину } D = \frac{-0,79\sqrt{6}}{1 - 0,79^2} \approx -5,1.$$

Для уровня значимости $\alpha = 0,01$ по таблице функции Лапласа находим $t_{0,01} = 2,58$. Поскольку $|D| = 5,1 > 2,58$, то гипотезу $H_0: g = 0$ следует отвергнуть. Таким образом, на основании экспериментальных данных коэффициент корреляции X и Y следует признать значимым.

Литература

- 1 **Андронов, А. М.** Теория вероятностей и математическая статистика : учеб. для вузов / А. М. Андронов, Е. А. Копытов, Л. Я. Гринглаз. – СПб. : Питер, 2004. – 461 с.
- 2 **Бородин, А. Н.** Элементарный курс теории вероятностей и математической статистики / А. Н. Бородин. – СПб. : Лань, 1998. – 224 с.
- 3 **Вентцель, Е. С.** Теория вероятностей : учеб. для вузов / Е. С. Вентцель. – 5-е изд., стер. – М. : Высш. шк., 1998. – 576 с.
- 4 **Герасимович, А. И.** Математическая статистика / А. И. Герасимович. – Минск : Выш. шк., 1983. – 275 с.
- 5 **Гмурман, В. Е.** Теория вероятностей и математическая статистика : учеб. пособие для вузов / В. Е. Гмурман. – М. : Высш. шк., 1998. – 479 с.
- 6 **Гмурман, В. Е.** Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. – М. : Высш. шк., 1998. – 400 с.
7. **Евдокимович, В.Е.** Основы теории вероятностей : учеб.-метод. пособие / В.Е. Евдокимович. – Гомель : УО «БелГУТ», 2007. – 122 с.
8. **Курносенко, Н.М., Евдокимович, В.Е.** Теория вероятностей и математическая статистика : практическое пособие / Н.М. Курносенко, В.Е. Евдокимович. – Гомель : УО «ГГУ им.Ф.Скорины», 2006. – 86 с.
9. **Малинковский, Ю. В.** Теория вероятностей и математическая статистика (часть 1. Теория вероятностей) : учеб. пособие / Ю. В. Малинковский. – Гомель : УО «ГГУ им. Ф.Скорины», 2004. – 355 с.
10. **Прищепова, Т.В.** Основы теории вероятностей : учеб.-метод. пособие / Т. В. Прищепова. – Гомель : УО «БелГУТ», 2008. – 140 с.
11. **Пугачёв, В. С.** Теория вероятностей и математическая статистика / В. С. Пугачёв. – М. : Наука, 1979. – 496 с.

12. **Лагойкин, А. Н.** Теория вероятностей (сборник заданий и методические указания по РГР) / А. Н. Лагойкин, В. С. Серёгина, А. Ю. Сокольский. – Гомель : БелГУТ, 1994. – 52 с.

13. **Сазонова, Е. Л.** Теория вероятностей : пособие для студентов ФБО. Ч.1. Теория вероятностей / Е. Л. Сазонова; под ред. В. С. Серёгиной. – Гомель : БелГУТ, 2000. – 95 с.

ГТУ ИМЕНИ Ф. СКОРИНЫ