ДОНСКОЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ

УПРАВЛЕНИЕ ДИСТАНЦИОННОГО ОБУЧЕНИЯ И ПОВЫШЕНИЯ КВАЛИФИКАЦИИ

Кафедра «Математика»

УЧЕБНОЕ ПОСОБИЕ

«Математическая статистика с элементами теории вероятностей»

(часть 1)

Авторы:
Смирнова И.Ю.
Коровина К.С.

Ростов-на-Дону, 2015

Аннотация

Методическое пособие предназначено для преподавателей, аспирантов, студентов всех специальностей и форм обучения. Основная задача пособия состоит в демонстрации практического применения методов теории вероятностей и математической статистики, методики типовых задач.

Авторы

Ст.преп. Смирнова И.Ю.

Ст.преп. Коровина К.С.

Оглавление

ГЛАВА 1. ВАРИАЦИОННЫЕ РЯДЫ И ИХ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ

1.1 Понятие вариационного ряда и графические методы его изображения

В реальных условиях обычно бывает трудно или экономически нецелесообразно, а иногда и невозможно, исследовать всю совокупность, характеризующую изучаемый признак (*генеральную совокупность*). Поэтому на практике широко применяется выборочное наблюдение, когда обрабатывается часть генеральной совокупности (*выборочная совокупность* или *выборка*).

Свойства генеральной совокупности неизвестны, поэтому возникает задача их оценки по выборке. Для получения хороших оценок характеристик генеральной совокупности необходимо, чтобы выборка была *репрезентативной* (представительной). Метод, основанный на изучении генеральной совокупности по ее выборочным представителям, называется *выборочным методом*.

Различают следующие типы выборок:

- *собственно случайная* выборка, образованная случайным выбором элементов без расчленения на части или группы;
- *механическая* выборка, в которую элементы из генеральной совокупности отбираются через определенный интервал, например, если объем выборки должен составлять 10% (10% выборка), то отбирается каждый 10-й ее элемент и т.д.;
- *типическая* (стратифицированная) выборка, в которую случайным образом отбираются элементы из типических групп, на которые по некоторому признаку разбивается генеральная совокупность;
- *серийная* (гнездовая) выборка, в которую случайным образом отбираются не элементы, а целые группы совокупности (серии), а сами серии подвергаются сплошному наблюдению;
- *комбинированная* выборка, в которой используются различные комбинации указанных методов, например, типическая выборка сочетается с механической и собственно случайной.

Для образования выборки используют два способа: *повторный отбор*, когда каждый элемент, случайно отобранный и обследованный, возвращается в общую совокупность и может быть повторно отобран и *бесповторный отбор*, когда отобранный элемент не возвращается в общую совокупность. Если объем генеральной совокупности велик, то различие между выборками с возвратом или без возврата (когда вновь встретившееся число пропускается) незначительно и практически не сказывается на окончательных результатах.

Расхождения между величиной какого-либо показателя, найденного посредством статистического наблюдения, и действительными его размерами называются ошибками наблюдения. Ошибки органически присущи выборочному наблюдению и возникают в силу того, что выборочная совокупность не полностью воспроизводит генеральную. Избежать ошибок нельзя, однако можно свести их к минимальным значениям, границы которых устанавливаются с заданной точностью.

В генеральной совокупности исследуется некоторый количественный признак X. Из нее случайным образом извлекается выборка объема n, то есть число элементов выборки равно n. Значения признака при переходе от одного элемента к другому изменяются (варьируют), поэтому в статистике различные значения признака называют *вариантами*, а совокупность значений признака $x_1, x_2, \ldots, x_n$, расположенных в порядке возрастания или убывания (ранжированные), -*вариационным рядом.*

Абсолютные числа, показывающие, сколько раз встречаются те или иные варианты в ряду, называют частотами, обозначают m_i .

Отношение частоты m_i того или иного варианта к объему выборки n называют *относительной частотой (частостью)* и обозначают w_i: $w_i = \dfrac{m_i}{n}$.

При изучении вариационного ряда также используют понятия накопленной частоты и накопленной частости. Пусть x некоторое число. Тогда количество вариантов, значения которых меньше x, называется *накопленной частотой*.

Отношение накопленной частоты к общему числу наблюдений n называется *накопленной частостью* $w_i^{нак}$ или v_i

Признак называется *дискретно варьируемым*, если его отдельные значения (варианты) отличаются друг от друга на некоторую конечную величину (обычно целое число). Вариационный ряд такого признака называется *дискретным вариационным рядом*.

Таблица 1. Общий вид дискретного вариационного ряда частот

Значения признака X	x_i	x_1	x_2	...	x_n
Частоты	m_i	m_1	m_2	...	m_n

Признак называется *непрерывно варьирующим*, если его значения отличаются друг от друга на сколь угодно малую величину, т.е. признак может принимать любые значения в некотором интервале. *Непрерывный вариационный ряд* для такого признака называется *интервальным*.

Таблица 2. Общий вид интервального вариационного ряда частот

Интервалы	$a_i - a_{i+1}$	$a_1 - a_2$	$a_2 - a_3$	...	$a_k - a_{k+1}$
Частоты	m_i	m_1	m_2	...	m_n

Разности $a_2 - a_1$, $a_3 - a_2$,.., $a_{i+1} - a_i$ называют интервальными разностями, или величинами (длинами) интервалов.

Если интервал имеет обе границы, то его называют закрытым. Первый и последний интервалы могут быть открытыми, т. е. иметь только одну границу. Например, 1-й интервал может быть задан как «до 100», 2-й — «100-110», ... предпоследний — «190-200», последний — «200 и более». Очевидно, что 1-й интервал не имеет нижней границы, а последний — верхней, оба они — открытые. Часто открытые интервалы приходится условно закрывать. Обычно для этого величину 1-го интервала принимают равной величине 2-го, а величину последнего — величине предпоследнего. В нашем примере величина 2-го интервала равна 110 - 100 = 10, следовательно, нижняя граница 1-го условно составит 100 - 10 = 90; величина предпоследнего равна 200 - 190 = 10, значит, верхняя граница последнего условно составит 200 + 10 = 210.

Кроме этого в интервальном вариационном ряде могут встречаться интервалы разной длины. Если интервалы в вариационном ряде имеют одинаковую длину (интервальную разность), их называют равновеликими, в противном случае — неравновеликими.

Дискретный ряд легко изобразить графически в виде полигона распределения частот, представляющего собой ломаную, отрезки которой соединяют точки с координатами (x_1, m_1),

(x_2, m_2),..., (x_k, m_k), где x_i откладываются на оси абсцисс, а m_i – на оси ординат. Если на оси ординат откладывать не абсолютные (m_i), а относительные (w_i) частоты, то получим *полигон относительных частот.*

Интервальные вариационные ряды графически можно представить с помощью *гистограммы*. При ее построении по оси абсцисс откладываются значения изучаемого признака (границы интервалов). В том случае, если интервалы одинаковой величины, по оси ординат можно откладывать частоты или частости. Если же интервалы имеют разную величину, по оси ординат необходимо откладывать значения абсолютной или относительной плотности распределения.

Абсолютная плотность — отношение частости интервала к его величине:

$$f(a)_i = \frac{m_i}{k_i},$$

где $f(a)_i$ — абсолютная плотность i-го интервала; m_i — его частота;

k_i— величина (интервальная разность).

Абсолютная плотность показывает, сколько единиц совокупности приходится на единицу интервала.

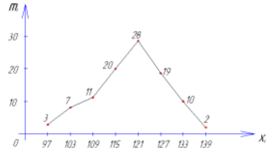
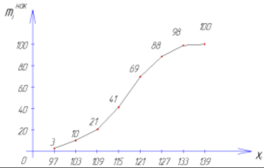
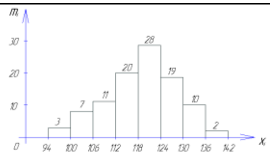
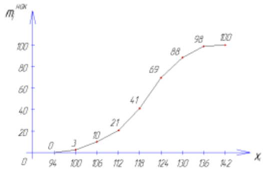
Относительная плотность — отношение частости интервала к его величине:

$$f(a)_i = \frac{w_i}{k_i},$$

где $f(a)_i$ — относительная плотность i-го интервала; w_i — его частость.

Дискретные и интервальные вариационные ряды графически можно представить в виде *кумуляты, или кумулятивной кривой.* При построении по оси ординат откладывается величина накопленных частот, а по оси абсцисс – возрастающие значения количественного признака. Накопленные частоты и кумулята – это интегральные показатели плотности распределения в вариационном ряду.

Таблица 3. Графические изображения вариационного ряда

Ряд	Полигон или гистограмма	Кумулята
Дискретный		
Интервальный		

1.2 Эмпирическая функция распределения.

Эмпирической функцией распределения совокупности по признаку X называют функцию $F^*(x)$ (или F_n), определяющую для каждого значения x относительную частоту события $X < x$. Таким образом,

$$F^*(x) = \frac{n_x}{n}$$

где n_x – число вариант, меньших x, n – объем выборки.

Эмпирическая функция распределения обладает свойствами:

1) $0 \le F^*(x) \le 1$.

2) $F^*(x)$ – неубывающая функция.

3) Если x_1 – наименьшее значение признака, то $F^*(x) = 0$ при $x \le x_1$; если $x_к$ – наибольшее значение признака, то $F^*(x) = 1$ при $x > x_к$.

Эмпирическую функцию $F^*(x)$ можно записать в виде:

$$F^*(x) = \begin{cases} 0, \ x \le x_1; \\ \dfrac{n_x}{n}, x_i < x \le x_{i+1} \ (i < m) \\ 1, \ x > x_m \end{cases}$$

Данная функция непрерывна, кусочно-постоянна и изменяется в каждой точке x_i, где x_i — варианта рассматриваемого статистического распределения.

1.3 Числовые характеристики вариационного ряда.

Числовые характеристики вариационных рядов вычисляют по данным, полученным в результате наблюдений (статистическим данным), поэтому их называют также статистическими характеристиками или оценками. На практике часто оказывается достаточным знание сводных характеристик вариационных рядов: средних или характеристик положения (центральной тенденции); характеристик рассеяния или вариации (изменчивости); характеристик формы (асимметрии и крутости распределения).

Самой известной и наиболее употребляемой характеристикой любого вариационного ряда является его средняя арифметическая, называемая также *выборочным средним*. Средняя арифметическая характеризует значения признака, вокруг которого концентрируются наблюдения, т.е. центральную тенденцию распределения, и вычисляется по формуле:

$$\overline{x_в} = \frac{\sum_{i=1}^{n} m_i x_i}{n}$$

где x_i – варианты, m_i - частоты.

Свойства средней арифметической:

1. Если находят среднюю арифметическую для интервального вариационного ряда, то в качестве значения признака для каждого интервала условно принимают его середину.
2. Средняя арифметическая постоянной величины равна этой постоянной.
3. Если все варианты ряда уменьшить (увеличить) на одно и то же число, то средняя арифметическая уменьшится (увеличится) на то же число.
4. Если все варианты ряда уменьшить (увеличить) в одно и то же число раз, то средняя арифметическая уменьшится (увеличится) во столько же раз.
5. Если частоты (частости) средней взвешенной разделить или умножить на постоянное число, то средняя арифметическая не изменится.
6. Сумма отклонений вариантов ряда от средней арифметической равна нулю.

В статистическом анализе кроме средней арифметической, называемой аналитической средней, широко применяют структурные, или порядковые, средние, к которым относятся медиана и мода.

Медиана – значение признака ряда, относительно которого вариационный ряд делится на две равные по числу вариантов части:

$M_e = x_{k+1}$, если $n = 2k+1$;

$M_e = (x_k + x_{k+1})/2$, если $n = 2k$.

Достоинство медианы как меры центральной тенденции заключается в том, что на нее не влияет изменение крайних членов вариационного ряда, если любой из них, меньший медианы, остается меньше ее, а любой, больший медианы, продолжает быть большее ее. Медиана предпочтительнее средней арифметической для ряда, у которого крайние варианты по сравнению с остальными оказались чрезмерно большими или малыми.

Мода - это значение признака, наиболее часто встречающееся в вариационном ряду:

$$Mo = x_j, \text{ если } m_j = m_{max}$$

Особенность *моды* как меры центральной тенденции заключается в том, что она также не изменяется при изменении крайних членов ряда, т.е. обладает определенной устойчивостью к вариации признака.

Для получения полного представления о вариационном ряде (определив центральную тенденцию распределения с помощью характеристик положения) далее оценивают рассеяние (вариацию, изменчивость) исследуемого признака вокруг этих величин. Простейшим и, весьма приближенным показателем вариации (изменчивости), является вариационный *размах* - разность между наибольшим и наименьшим значениями признака: $R = x_{max} - x_{min}$. Размах вариации наиболее полезен, если нужен быстрый и общий взгляд на изменчивость при сравнении большого количества выборок.

Но наибольший интерес представляют меры вариации (рассеяния) наблюдений вокруг средних величин, в частности, вокруг средней арифметической. К таким оценкам относятся выборочная дисперсия и среднее квадратичное отклонение.

Дисперсией вариационного ряда называется средняя арифметическая квадрата отклонения значений признака от их средней арифметической:

$$D_в = \frac{\sum_{i=1}^{n} m_i (x_i - \overline{x_в})^2}{n}$$

Свойства дисперсии:

1. Дисперсия постоянной величины равна нулю.
2. Если все варианты ряда уменьшить на одно и то же число, то дисперсия не изменится.
3. Если все варианты ряда уменьшить(увеличить) в r раз, то дисперсия уменьшится(увеличится) в r^2 раз.

Выборочная дисперсия обладает одним существенным недостатком: если среднее арифметическое выражается в тех же единицах, что и значения случайной величины, то, согласно определению, дисперсия выражается уже в квадратных единицах. Этого недостатка можно избежать, если использовать в качестве меры вариации признака *среднее квадратичное отклонение*:

$$\sigma_в = \sqrt{D_в}$$

При малых объемах выборки дисперсия является смещенной оценкой, поэтому при объемах $n \le 30$ используют *исправленную дисперсию*

$$S^2 = \frac{n}{n-1} D_в$$

и *исправленное среднее квадратичное отклонение*

$$S = \sqrt{S^2} \ .$$

Другой часто используемой характеристикой меры рассеяния признака является *коэффициент вариации*:

$$V = \frac{\sigma_в}{\overline{\overline{x}}_в} \cdot 100\%$$

Достоинством коэффициента вариации является то, что это безразмерная характеристика, позволяющая сравнивать варьирование несоизмеримых вариационных рядов. Кроме того, чем меньше значение коэффициента вариации, тем однороднее совокупность по изучаемому признаку и типичнее средняя. Совокупности с коэффициентом вариации *V* > 30-35% принято считать неоднородными.

Наряду с дисперсией используют и *среднее абсолютное от-*

клонение:

$$\theta = \frac{\sum_{i=1}^{n} m_i \left| x_i - \overline{x_в} \right|}{n}$$

Достоинством среднего линейного отклонения является его размерность, т.к. выражается в тех же единицах, что и значения случайной величины.

Дополнительным и простым показателем рассеяния значений признака является *квартильный размах:*

$$H = Q_В - Q_Н$$

Квартильный размах включает в себя медиану и 50% наблюдений, отражающих центральную тенденцию признака, исключая наименьшие и наибольшие значения.

К характеристикам формы относят коэффициент асимметрии и эксцесс.

Коэффициентом асимметрии вариационного ряда называется число

$$As = \frac{\sum_{i=1}^{n} m_i (x_i - \overline{x_в})^2}{n\sigma_в^2}$$

Если коэффициент асимметрии равен нулю, то распределение имеет симметричную форму. Если распределение асимметрично, одна из ветвей полигона частот имеет более пологий спуск, чем другая. Если асимметрия правосторонняя, то справедливо неравенство: $\overline{x_в} > M_e > M_o$, что означает преимущественное появление в распределении более высоких значений признака. Если асимметрия левосторонняя, то выполняется неравенство: $\overline{x_в} < Me < Mo$, означающее, что в распределении чаще встречаются более низкие значения. Чем больше значение коэффициента асимметрии, тем более асимметрично распределение (до 0,25 асимметрия незначительная; от 0,25 до 0,5 умеренная; свыше 0,5 – существенная).

Коэффициентом эксцесса называют величину

$$Ek = \frac{\sum_{i=1}^{n} m_i (x_i - x_в)^4}{n\sigma_в^4} - 3.$$

Эксцесс является показателем крутости (островершинности) вариационного ряда. Если эксцесс положителен, то полигон вариационного ряда имеет более крутую вершину. Это говорит о скоплении значений признака в центральной зоне ряда распределения, т.е. о преимущественном появлении в данных значений, близких к средней величине. Если эксцесс отрицателен - то полигон имеет более пологую вершину, это означает, что значения признака не концентрируются в центральной части ряда, а достаточно равномерно рассеяны по всему диапазону от минимального до максимального значения.

Значения коэффициента эксцесса лежат на полусегменте [-3;+∞).

1.4 Решение типовых задач.

Пример 1.

При проведении 20 серий из 10 бросков игральной кости число выпадений шести очков оказалось равным 1,1,4,0,1,2,1,2,2,0,5,3,3,1,0,2,2,3,4,1.

Составить вариационный ряд, найти моду и медиану.

Решение: Статистический ряд для абсолютных и относительных частот имеет вид:

x_i	0	1	2	3	4	5
n_i	3	6	5	3	2	1
w_i	0,15	0,3	0,25	0,15	0,1	0,05

Мода M_0 – варианта, имеющая наибольшую частоту, значит $M_0 =$ 5.

Медиана M_e - варианта, которая делит вариационный ряд на две части, равные по числу вариант. Если число вариант нечетно $n = 2k + 1$, то $M_e = x_{k+1}$, а при четном $n =2k$

$$M_e = \frac{x_k + x_{k+1}}{2}.$$

В нашей задаче число вариант четное, следовательно

$$M_e = \frac{2+3}{2} = 2{,}5$$

Пример 2.

Дана выборка: 1,7,8,3,3,7,3,1,1,3. Найти выборочное среднее.

Решение: Составим для данной выборки статистический ряд:

x_i	1	3	7	8
n_i	3	4	2	1

Объем выборки, то есть количество входящих в нее чисел, равен 10. Тогда

$$\overline{x}_B = \frac{\sum_{i=1}^{k} n_i x_i}{n} = \frac{1\cdot 3 + 3\cdot 4 + 7\cdot 2 + 8\cdot 1}{10} = 3{,}7.$$

Пример3.

Дана выборка, вариационный ряд которой имеет вид:

10,8; 11,1; 11,7; 12,2; 13,1; 13,4; 13,9; 14,3; 14,3; 14,4; 14,8; 16,5; 17,7; 18,2; 19,9; 20,0; 20,3; 20,8; 23,1; 24,2; 25,1; 25,1; 25,7; 28,4; 28,5; 29,3; 29,8; 29,9; 30,2; 30,4.

Составить статистический ряд распределения абсолютных и относительных частот, состоящий из пяти интервалов, и найти выборочное среднее.

Решение: Объем выборки n = 30. Выберем в качестве границ интервала a = 10,5 и b = 30,5. Тогда $h = \frac{30{,}5 - 10{,}5}{5} = 4$, и ($a$, b) разбивается на части (10,5; 14,5), (14,5; 18,5), (18,5; 22,5), (22,5; 26,5) и (26,5; 30,5). Статистический ряд при этом имеет вид:

Номер интервала	Границы интервала		Середина интервала	Абсолютные частоты
1	10,5	14,5	12,5	10
2	14,5	18,5	16,5	4
3	18,5	22,5	20,5	4
4	22,5	26,5	24,5	5
5	26,5	30,5	28,5	7

Найдем выборочное среднее, принимая в качестве вариант середины полученных интервалов:

$$\overline{x}_B = \frac{12,5\cdot 10+16,5\cdot 4+20,5\cdot 4+24,5\cdot 5+28,5\cdot 7}{30} =$$

$$= \frac{119}{6} \approx 19,833.$$

Пример 4.

Дана выборка, состоящая из чисел: 3,2; 4,1; 8,1; 8,1; 6,7; 4,4; 4,4; 3,2; 5,0; 6,7; 6,7; 7,5; 3,2; 4,4; 6,7; 6,7; 5,0; 5,0; 4,4; 8,1. Найти выборочную дисперсию.

Решение: Составим статистический ряд:

x_i	3,2	4,4	5,0	6,7	7,5	8,1
n_i	3	5	3	5	1	3

Найдем выборочную среднюю и выборочную дисперсию по формулам:

$$\overline{x_B} = \frac{\sum_{i=1}^{n} m_i x_i}{n} \qquad D_B = \frac{\sum_{i=1}^{n} m_i (x_i - \overline{x_B})^2}{n}.$$

Получим

$$\overline{x}_B = \frac{3,2\cdot 3+4,4\cdot 5+5\cdot 3+6,7\cdot 5+7,5\cdot 1+8,1\cdot 3}{20} = 5,595;$$

$$D_B = \frac{(3,2-5,595)^2\cdot 3+\ldots+(8-5,595)^2\cdot 3}{20} = 2,698.$$

Пример 5.

Дана выборка значений случайной величины: 2, 3, 3, 4, 2, 5, 5, 5, 6, 3, 6, 3, 4, 4, 4, 6, 5, 7, 3, 5. Найти исправленное выборочное среднее квадратическое отклонение.

Решение: Составим статистический ряд:

x_i	2	3	4	5	6	7
n_i	2	5	4	5	3	1

Воспользуемся формулами для исправленной выборочной дисперсии и исправленного выборочного среднего квадратического отклонения:

$$s^2 = \frac{n}{n-1} D_B, \quad s = \sqrt{s^2}.$$

Тогда

$$n = 20, \quad \bar{x}_B = \frac{2\cdot 2 + 3\cdot 5 + 4\cdot 4 + 5\cdot 5 + 6\cdot 3 + 7\cdot 1}{20} = 4{,}25;$$

$$s^2 = \frac{(2-4{,}25)^2\cdot 2 + (3-4{,}25)^2\cdot 5 + \ldots + (7-4{,}25)^2\cdot 1}{19} = 1{,}987;$$

$$s = \sqrt{1{,}987} = 1{,}41.$$

Пример 6. При обследовании 50 членов семей рабочих и служащих установлено следующее количество членов семьи: 5;3;2;1;4;6;3;7;9;1;3;2;5; 6; 8; 2; 5; 2; 3; 6; 8; 3; 4; 4; 5; 6; 5; 4; 7; 5; 6; 4; 8; 7; 4; 5; 7; 8; 6; 5; 7; 5; 6; 6; 7; 3; 4; 6; 5; 4.

1) Составьте вариационный ряд распределения частот.
2) Постройте полигон распределения частот, кумуляту.
3) Постройте эмпирическую функцию распределения.
4) Определите средний размер (среднее число членов) семьи.
5) Охарактеризуйте колеблемость размера семьи с помощью показателей вариации (дисперсии, среднего квадратического отклонения, коэффициента вариации).

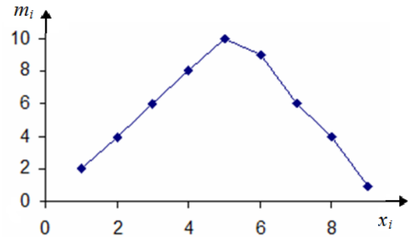
Объясните полученные результаты, сделайте выводы.

Решение 1) В данной задаче изучаемый признак является дискретно варьирующим, так как размер семей не может отличаться друг от друга менее чем на одного человека. Следовательно, необходимо построить дискретный вариационный ряд. Чтобы сделать это, необходимо подсчитать, сколько раз встречаются те или иные значения признака, и расположить их в порядке возрастания или убывания. Значения изучаемого признака — размер семьи — обозначим x_i, частоты — m_i.

Произведем упомянутые расчеты и запишем их результаты в таблицу:

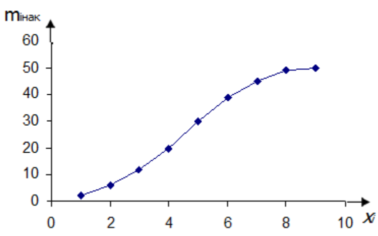
X_i	1	2	3	4	5	6	7	8	9
m_i	2	4	6	8	10	9	6	4	1

2) Дискретный вариационный ряд можно представить графически, построив полигон распределения частот или частостей.



Для того чтобы построить кумуляту, необходимо рассчитать накопленные частоты или частости. Накопленная частота 1-го варианта (x_1 = 1) равна самой частоте этого варианта, т. е. v_1 = 2. Накопленная частота 2-го варианта (x_2 = 2) равна сумме частот 1-го и 2-го вариантов, т.е. v_2 = 2 + 4 = 6. Далее, аналогично v_3 = 12; v_4 = 20; v_5 = 30; v_6 = 39; v_7= 45; v_8 = 49; v_9 =50.

Построим кумуляту.



3) Построим эмпирическую функцию распределения.

$$F_n(x)=\begin{cases} 0, & x\le 1\\ 0{,}04, & 1<x\le 2\\ 0{,}12, & 2<x\le 3\\ 0{,}24, & 3<x\le 4\\ 0{,}4, & 4<x\le 5\\ 0{,}6, & 5<x\le 6\\ 0{,}78, & 6<x\le 7\\ 0{,}9, & 7<x\le 8\\ 0{,}98, & 8<x\le 9\\ 1, & x>9 \end{cases}$$

Рассчитаем средний размер (среднее число членов) семьи. Так как частоты отличны друг от друга, расчет средней арифметической произведем по формуле

$$\overline{x_B}=\frac{\sum_{i=1}^{n} m_i x_i}{n}$$

$$\overline{x_B}=\frac{1\cdot 2+2\cdot 4+3\cdot 6+4\cdot 8+5\cdot 10+6\cdot 9+7\cdot 6+8\cdot 4+9\cdot 1}{50}=4{,}94$$

Средний размер семьи — 4,94 чел.

5) Так как частоты неодинаковы, для расчета дисперсии размера семьи используем формулу

$$D_B=\frac{\sum_{i=1}^{n} m_i (x_i-\overline{x_B})^2}{n}.$$

$$D_B=\frac{2(1-4{,}94)^2+4(2-4{,}94)^2+6(3-4{,}94)^2+8(4-4{,}94)^2+10(5-4{,}94)^2}{50}+$$

$$+\frac{9(6-4{,}94)^2+6(7-4{,}94)^2+4(8-4{,}94)^2+1(9-4{,}94)^2}{50}=3{,}6964$$

Дисперсия размера семьи $D_в$ =3,6964 чел^2.

Найдем среднее квадратическое отклонение размера семьи по формуле

$\sigma_B = \sqrt{D_B}$, $\sigma_в$ =1,9226 чел.

Найдем коэффициент вариации размера семьи по формуле

$$V = \frac{\sigma_B}{\overline{\overline{x}}_B} \cdot 100\%$$

Коэффициент вариации составляет 38%. Так как коэффициент вариации больше 35%, можно сделать вывод о том, что изучаемая совокупность семей является неоднородной, чем и объясняется высокая колеблемость размера семьи в данной совокупности.

Ввиду неоднородности семей, попавших в выборку, использование средней арифметической для характеристики наиболее типичного уровня размера семьи не вполне оправданно — средняя арифметическая нетипична для изучаемой совокупности, в качестве характеристики наиболее типичного уровня размера семьи в данной совокупности лучше использовать моду или медиану.

1.5 Задачи к главе 1

1. В отделе обуви универмага в течение дня были проданы туфли следующих размеров: 37,36,35,37,38,37,36,37,39,38,38,37, 36,37,37,37,36. Составьте вариационный ряд. Постройте полигон и кумуляту. Найдите медиану, моду, среднюю арифметическую и коэффициент вариации.

2. Двадцати подросткам, отобранным случайным образом, показали блок коммерческой рекламы о новых сортах жевательной резинки и попросили оценить рекламу в баллах о 0 до 100. Результаты оценки дали следующие баллы: 89,75,59,96,88, 71,43,62,80,92,76,72,67,60,79,85,77,83,87,53. Найдите среднюю арифметическую, медиану, моду, дисперсию и стандартное отклонение выборочного рейтинга. Постройте интервальный ряд без корректировки границ первого и последнего интервалов.

3. Постройте гистограмму частот, найдите среднюю арифметическую и среднее квадратическое отклонение для данных о дневной выручке в магазине электроники (тыс. руб.):

X_i	100-200	200-300	300-400	400-500	500-600	600-700
m_i	3	5	9	14	8	3

4. Инженер по контролю качества продукции обнаружил в 10 партиях электро ламп, произведенных заводом, следующее число бракованных изделий: 5,3,7,1,0,6,3,4,5,2. Найдите среднее число и стандартное отклонение бракованных ламп. Начертите график эмпирической функции распределения. Вычислите коэффициент вариации, моду и медиану.

ГЛАВА 2. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

2.1 Основные формулы комбинаторики

Группы, составленные из каких-либо предметов (безразлично каких, например, букв, кубиков, цветных шаров, чисел) называются комбинациями.

Комбинаторика – наука, изучающая комбинации, которые можно составить по определенным правилам из элементов некоторого конечного множества. Определим основные типы комбинаций.

Перестановки – это комбинации, составленные из всех n элементов данного множества и отличающиеся только порядком их расположения. Число всех возможных перестановок $P_n = n!$

Пример.

Сколько различных списков (отличающихся порядком фамилий) можно составить из 7 различных фамилий?

Решение.

$P_7 = 7! = 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 = 5040.$

Размещения – комбинации из m элементов множества, содержащего n различных элементов, отличающиеся либо составом элементов, либо их порядком. Число всех возможных размещений

$$A_n^m = n(n-1)(n-2)...(n-m+1).$$

Пример.

Сколько возможно различных вариантов пьедестала почета (первое, второе, третье места), если в соревнованиях принимают участие 10 человек?

Решение.

$$A_{10}^3 = 10 \cdot 9 \cdot 8 = 720.$$

Сочетания – неупорядоченные наборы из m элементов множества, содержащего n различных элементов (то есть наборы, отличающиеся только составом элементов). Число сочетаний

$$C_n^m = \frac{n!}{m!(n-m)!} \qquad (5)$$

Пример.

В отборочных соревнованиях принимают участие 10 человек, из которых в финал выходят трое. Сколько может быть различных троек финалистов?

Решение.

В отличие от предыдущего примера, здесь не важен порядок финалистов, следовательно, ищем число сочетаний из 10 по 3:

$$C_{10}^{3}=\frac{10!}{3!(10-3)!}=\frac{8\cdot 9\cdot 10}{1\cdot 2\cdot 3}=120$$

Перестановки с повторениями - это комбинации, составленные из всех *n* элементов данного множества, среди которых один элемент повторяется *m* раз, другой *p* раз, и т.д., например, последний элемент повторяется s раз. Число всех возможных перестановок

$$P(m,p,...,s)=\frac{n!}{m!\,p!...s!}$$

Пример.

Сколькими способами можно разделить 15 предметов на 3 группы, чтобы в одной группе было 2, другой 6, в третьей 7 предметов?

Решение.

$$P(2,6,...,7)=\frac{15!}{2!6!7!}=180180$$

Размещение с повторениями - размещения из *n* элементов по *m*, в которых некоторые из элементов (и даже все) могут оказаться одинаковыми. Число всех возможных размещений

$$\left(A_n^m\right)_{повт.}=n^m$$

Пример

Замок сейфа представляет собой систему из 3 цифровых дисков по 30 позиций в каждой. Для того чтобы открыть сейф, каждый из трех дисков замка должен быть установлен в определенной позиции. Сколько существует различных цифровых комбинаций в этом замке?

Решение.

$$\left(A_{30}^{3}\right)_{повт.} = 30^{3} = 27000.$$

Сочетания с повторениями - сочетания из n элементов по m, в которых некоторые из элементов (и даже все) могут оказаться одинаковыми. Число всех возможных сочетаний

$$\left(C_{n}^{m}\right)_{повт.} = C_{n+m-1}^{m} = \frac{(n+m-1)!}{m!(n-1)!}$$

Пример.

На конкурс представлены 10 научных студенческих работ. Денежные премии будут присуждаться по следующим номинациям: оригинальная научная идея; использование современного экономико-математического аппарата; применение компьютерного обеспечения; презентация на научных конференциях. Сколько существует вариантов распределения премий, если по каждой комбинации установлены одинаковые премии.

Решение.

Поскольку по каждой номинации установлены одинаковые премии, то порядок следования работ в комбинации 4 премиальных работ значения не имеет. Тогда используем формулу сочетания с повторениями

$$\left(C_{10}^{4}\right)_{повт} = \frac{(10+4-1)!}{4!9!} = \frac{13!}{4!9!} = 715.$$

2.2 Основные правила комбинаторики

При решении статистических задач часто приходится рассматривать различные комбинации из некоторой совокупности элементов, например, составлять различные выборки.

Пример. В магазине продаются 3 марки телевизоров и два вида видеомагнитофонов. У покупателя есть возможность приобрести либо телевизор, либо видеомагнитофон. Сколькими способами он может совершить одну покупку? Сколько различных комплектов, содержащих телевизор и магнитофон, можно приобрести в этом магазине, если покупатель собирается приобрести в паре телевизор и видеомагнитофон?

Решение. 1) Один телевизор можно выбрать тремя способами, а магнитофон - другими двумя способами. Тогда телевизор или магнитофон можно купить N=3+2=5 способами.

2) Обозначим марки телевизоров a,b,c, а марки магнитофонов x,y. Если выбран телевизор a, то можно составить комплекты ax, ay. Если же выбран магнитофон b, то получим комплекты bx, by. Если выбран телевизор c, то присоединяя к нему магнитофоны x,y, получим комплекты cx,cy. Таким образом, после каждого выбора телевизора магнитофон можно выбрать двумя способами. Следовательно, всего можно выбрать N=$3\cdot 2$=6 различных комплектов.

Сформулируем *основные правила комбинаторики.*

Если объект *a* может быть выбран *m* способами, объект *b* может быть выбран другими *n* способами (не такими как *a)*, то выбор одного элемента *a* или *b* из объединенной совокупности может быть осуществлен *m+n* способами.

Если объект *a* может быть выбран *m* способами и после каждого такого выбора объект *b* может быть выбран *n* способами, то выбор пары объектов *a* и *b* в указанном порядке может быть осуществлен $m\cdot n$ способами.

Эти правила могут быть распространены на случай трех и более выборок.

Пример. Сколькими способами можно из 40 человек, поступающих в вуз, создать 4 группы разных специальностей по 10 человек в каждой?

Решение. Первую группу можно создать C_{40}^{9} способами. Вторую группу можно создать из оставшихся 30 человек C_{30}^{9} способами. Третью группу можно создать из оставшихся 20 человек C_{20}^{9} способами. Оставшиеся 10 человек составят четвертую группу. Таким образом, число всех различных способов составления четырех групп из 40 человек равно

$$C_{40}^{10}\cdot C_{30}^{10}\cdot C_{20}^{10}=\frac{40!}{10!30!}\cdot\frac{30!}{10!20!}\cdot\frac{20!}{10!10!}=\frac{40!}{(10!)^4}$$

2.3 Решение типовых задач

Задача 1.

Сколькими способами можно расставить в одну шеренгу 6 человек?

Решение.

Первого человека в шеренге можно выбрать шестью способами, второго – пятью и так далее, то есть число способов

$N = P_6 = 6! = 1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6 = 720.$

ОТВЕТ: 720.

Задача 2.

Сколько различных трехзначных чисел можно составить, используя по одному разу цифры 1,2,3,4,5,6,7,8?

Решение.

Первую цифру можно выбрать 8 способами, вторую – 7, третью – 6, то есть

$$N = A_8^3 = \frac{8!}{(8-3)!} = \frac{8!}{5!} = \frac{1\cdot 2\cdot 3\cdot 4\cdot 5\cdot 6\cdot 7\cdot 8}{1\cdot 2\cdot 3\cdot 4\cdot 5} =$$

$$= 6\cdot 7\cdot 8 = 336.$$

ОТВЕТ: 336.

Задача 3.

Сколько различных букетов, состоящих из трех цветов, можно составить, если имеется 10 цветов разных сортов?

Решение.

Букеты различаются только составом входящих в них цветов, следовательно, это неупорядоченные наборы, поэтому

$$N = C_{10}^3 = \frac{10!}{3!(10-3)!} = \frac{10!}{3!\cdot 7!} = \frac{8\cdot 9\cdot 10}{1\cdot 2\cdot 3} = 120.$$

ОТВЕТ: 120.

Задача 4.

В классе 14 девочек и 12 мальчиков. Сколькими способами можно выбрать ученика из этого класса?

Решение.

Девочку из класса можно выбрать 14 способами, а мальчика – 12 способами. Так как выбирают одного ученика из класса, то выбор можно сделать 14+12=26 способами.

Ответ: 26 способов.

Задача 5.

Из города А в город В ведут две дороги, из города В в город С — три дороги, из города С до пристани — две дороги. Туристы хотят проехать из города А через города В и С к пристани. Сколькими способами они могут выбрать маршрут?

Решение.

Путь из А в В туристы могут выбрать двумя способами. Далее в каждом случае они могут проехать из В в С тремя способами. Значит, имеются $2 \cdot 3$ вариантов маршрута из А в С. Так как из города С на пристань можно попасть двумя способами, то всего существует $2 \cdot 3 \cdot 2$, т.е. 12 способов выбора туристами маршрута из города А к пристани.

Ответ: 12 способами.

Задача 6.

В семье 6 человек, и за столом в кухне стоят 6 стульев. Семья решила каждый вечер, ужиная, рассаживаться на эти 6 стульев по-новому. Сколько дней члены семьи смогут осуществлять задуманное?

Решение.

Для удобства рассуждений будем считать, что семья (бабушка, дедушка, мама, папа, дочь, сын) будет рассаживаться на стулья поочередно. Нас интересует сколько всего существует различных способов их размещения на стульях. Предположим, что первой усаживается бабушка. У нее имеется 6 вариантов выбора стула. Вторым садится дедушка и независимо выбирает стул из 5 оставшихся. Мама делает свой выбор третьей и выбор у нее будет из 4 стульев. У папы будет уже 3 варианта, у дочки — 2, ну а сын сядет на единственный незанятый стул. По правилу умножения получаем, что всего имеется $6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$ различных способов размещения. Таким образом, в «игру с рассаживаниями» семья может играть 720 дней, т. е. почти 2 года.

Ответ: 720.

Задача 7.

Сколько различных четырехзначных чисел, в которых цифры не повторяются, можно составить из цифр 0, 2, 4, 6?

Решение.

Из цифр 0, 2, 4, 6 можно получить P_4 перестановок. Из этого числа надо исключить те перестановки, которые начинаются с 0, так как натуральное число не может начинаться с цифры 0. Число таких перестановок равно P_3. Значит, искомое число четырехзначных чисел (без повторения цифр), которые можно составить из цифр 0, 2, 4, 6, равно P_4-P_3 . Получаем: $P_4-P_3 = 4!-3!=24-6=18$.

Ответ: 18 чисел.

Задача 8.

Имеется девять различных книг, четыре из которых — учебники. Сколькими способами можно расставить эти книги на полке так, чтобы все учебники стояли рядом?

Решение.

Сначала будем рассматривать учебники как одну книгу. Тогда на полке надо расставить не девять, а шесть книг. Это можно сделать P_6 способами. В каждой из полученных комбинаций можно выполнить P_4 перестановок учебников. Значит, искомое число способов расположения книг на полке равно произведению

$$P_6 \cdot P_4 = 6! \cdot 4! = 720 \cdot 24 = 17280$$

Ответ: 17280 способов.

Задача 9.

В кондитерском отделе продаются пирожные четырех сортов: наполеоны, эклеры, песочные и слоеные. Сколькими способами можно купить семь пирожных?

Решение.

При покупке пирожных важен состав. Количество покупаемых пирожных больше, чем видов. Значит, купить пирожные можно

$$\left(C_4^7\right)_{повт.} = C_{4+7-1}^7 = \frac{(4+7-1)!}{7!(4-1)!} = 120$$

Ответ: 120 способами.

Задача 10.

Для запирания автоматической камеры применяется секретный замок, который открывается лишь тогда, когда набрано «тайное слово». Это слово набирают с помощью пяти дисков, на каждом из которых изображено 12 букв. Сколько неудачных попыток может быть сделано человеком, не знающим секретного слова и подбирающим его наудачу?

Решение.

Из условия задачи видно, что порядок выбираемых букв существенен. Мы имеем пять дисков, на каждом из которых изображено 12 букв. Поэтому число комбинаций равно

$$\left(A_{12}^5\right)_{повт.} = 12^5 = 248832$$

Следовательно неудачных попыток может быть 248831.

Ответ. 248831 попытка.

Задача 11.

Сколькими способами можно переставить буквы в слове «математика»? Решение.

В этом слове буква «м» повторяется 2 раза, «а» - 3 раза, «т» - 2, «е», «к», «и» - по одному разу, по формуле перестановки с повторениями, получим

$P_{повт.}(1,1,1,2,2,3) = 1!\cdot 1!\cdot 1!\cdot 2!\cdot 2!\cdot 3!\cdot 10! = 151200.$

Ответ: 151200 способами.

Задачи к главе 2

Задача 1.

У Ирины пять подруг: Вера, Зоя, Марина, Полина и Светлана. Она решила пригласить двух из них в кино. Укажите все возможные варианты выбора подруг. Сколько таких вариантов?

Задача 2.

Стадион имеет четыре входа: A, B, C и D. Укажите все возможные способы, какими посетитель может войти через один вход, а выйти через другой. Сколько таких способов?

Задача 3.

В палатке имеется три сорта мороженого: рожок, брикет и эскимо. Наташа и Данил решили купить по одной порции. Сколько вариантов такой покупки? Решите задачу двумя способами.

Задача 4.

В цветочный магазин привезли 25 роз, 45 гвоздик и 30 лилий. Сколькими способами он может выбрать розу или гвоздику?

Задача 5.

На вершину горы ведут пять дорог. Сколькими способами турист может подняться и спуститься с нее, при условии, что спуск и подъем происходят по разным дорогам?

Задача 6.

Номер машины состоит из трех букв русского алфавита и трех цифр. Сколько можно составить различных номеров автомашин?

Задача 7.

У мамы два яблока и три груши. Каждый день в течение пяти дней она дает сыну по одному фрукту. Сколькими способами это может быть сделано?

Задача 8.

Сколькими способами 6 студентов, сдающих экзамен, могут занять места в аудитории, в которой стоит 20 одноместных столов?

Задача 9.

Сколькими способами можно составить расписание на один день, чтобы в нем было 4 различных предмета?

Задача 10.

Алфавит племени Мумбо-Юмбо состоит из букв А, У и С. Словом является любая последовательность, состоящая из 4 букв. Сколько слов в языке племени Мумбо-Юмбо?

Задача 11.

В классе 7 человек успешно занимаются математикой. Сколькими способами можно выбрать из них двоих для участия в математической олимпиаде? мальчиков и трех девочек. Сколькими способами это можно сделать?

Задача 12.

У профессора есть три любимых каверзных вопроса. В группе 20 студентов. а) Профессор решил задавать каждому из студентов по одному из каверзных вопросов; б) Профессор решил наудачу по списку группы выбрать студента, чтобы задать ему первый вопрос, потом из всего списка выбрать второго студента, чтобы задать ему второй вопрос, потом также выбрать третьего студента. в) Профессор решил спрашивать только троих студентов, каждому по одному вопросу (так, чтобы вопросы не повторялись). Сколько есть возможностей провести опрос в группе в каждом из случаев?

Задача 13.

Сколько слов получится при перестановке букв в слове: а) «бабушка», б) «комбинаторика» в) «солнце» г) «хоровод» д) «Миссисипи» е) «арбуз»?

Задача 14.

Сколькими способами можно вписать в колонку фамилии 30 учеников?

Задача 15.

В турнире по шахматам каждый участник сыграл с каждым по одной партии, всего было сыграно 36 партий. Определите число участников турнира?

Задача 16.

Сколькими способами 5 мальчиков и 5 девочек могут занять в театре в одном ряду места с 1 по 10? Сколькими способами они могут это сделать, если мальчики будут сидеть на нечетных местах, а девочки — на четных?

Задача 17.

Сколько различных трехзначных чисел (без повторения цифр) можно составить из цифр 1, 2, 3, 4, 5, таких, которые являются: а) четными; б) кратными 5?

Задача 18.

На референдуме предложены четыре вопроса, на которые можно ответить «да» или «нет». Сколько есть возможностей заполнения бюллетеня (на все вопросы надо дать ответ)?

Задача 19.

Сколькими способами можно составить команду из четырех человек для соревнований по бегу, если имеется семь бегунов?

Задача 20.

В классе 30 человек. Сколькими способами можно: а) назначить двух дежурных; б) выбрать 28 человек для осеннего кросса?

Задача 21.

В шахматном кружке занимаются 16 человек. Сколькими способами тренер может выбрать из них для предстоящего турнира: а) команду из четырех человек; б) команду из четырех человек, указав при этом, кто из членов команды будет играть на первой, второй, третьей и четвертой досках?

Задача 22.

В состав сборной включены два вратаря, пять защитников, шесть полузащитников, шесть нападающих. Сколькими способами тренер может выставить на поле команду из одного вратаря, трех защитников, четырех полузащитников и трех нападающих?

ГЛАВА 3. ЭЛЕМЕНТЫ ТЕОРИИ ВЕРОЯТНОСТЕЙ

3.1. Случайные события. Алгебра событий

В различных разделах науки и техники нередко возникают ситуации, когда результат каждого из многих проводимых опытов заранее предугадать невозможно, однако можно исследовать закономерности, возникающие при проведении серии опытов. Нельзя, например, точно сказать, какая сторона монеты окажется сверху при данном броске: герб или цифра – но при большом количестве бросков число выпадений герба приближается к половине количества бросков; нельзя заранее предсказать результат одного выстрела из данного орудия по данной цели, но при большом числе выстрелов частота попадания приближается к некоторому постоянному числу. Исследование вероятностных закономерностей массовых однородных явлений составляет предмет **теории вероятностей**.

Основным интуитивным понятием классической теории вероятностей является **случайное событие**. События, которые могут произойти в результате опыта, можно подразделить на три вида:

а) **достоверное событие** – событие, которое всегда происходит при проведении опыта;

б) **невозможное событие** – событие, которое в результате опыта произойти не может;

в) **случайное событие** – событие, которое может либо произойти, либо не произойти. Например, при броске игральной кости достоверным событием является выпадение числа очков, не превышающего 6, невозможным – выпадение 10 очков, а случайным – выпадение 3 очков.

Алгебра событий

Определение. **Суммой** $\boldsymbol{A+B}$ двух событий A и B называют событие, состоящее в том, что произошло хотя бы одно из событий A и B. **Суммой нескольких событий**, соответственно, называется событие, заключающееся в том, что произошло хотя бы одно из этих событий.

Пример 1.

Два стрелка делают по одному выстрелу по мишени. Если событие A – попадание первого стрелка, а событие B – второго, то сумма $A+B$ – это хотя бы од- но попадание при двух выстрелах.

Пример 2.

Если при броске игральной кости событием A_i назвать выпадение i очков, то выпадение нечетного числа очков является суммой событий $A_1+A_3+A_5$.

Назовем все возможные результаты данного опыта его *исходами* и предположим, что множество этих исходов, при которых происходит событие *A* (исходов, *благоприятных* событию *A*), можно представить в виде некоторой области на плоскости. Тогда множество исходов, при которых произойдет событие *A+B*, является объединением множеств исходов, благоприятных событиям *A* или *B* (рис. 1).

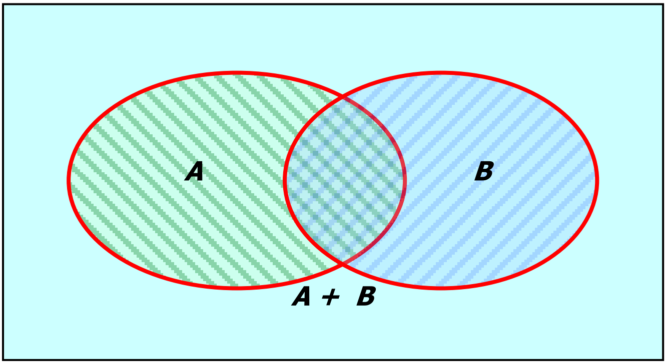


Рис.1.

Определение. **Произведением *AB*** событий *A* и *B* называется событие, состоящее в том, что произошло и событие *A*, и событие *B*. Аналогично **произведением нескольких событий** называется событие, заключающееся в том, что произошли все эти события.

Пример 3.

В примере 1 (два выстрела по мишени) событием *AB* будет попадание обоих стрелков.

Пример 4.

Если событие *A* состоит в том, что из колоды карт извлечена карта пиковой масти, а событие *B* – в том, что из колоды вынута дама, то событием *AB* будет извлечение из колоды дамы пик.

Геометрической иллюстрацией множества исходов опыта, благоприятных появлению произведения событий *A* и *B*, является пересечение областей, соответствующих исходам, благоприятным *A* и *B*.

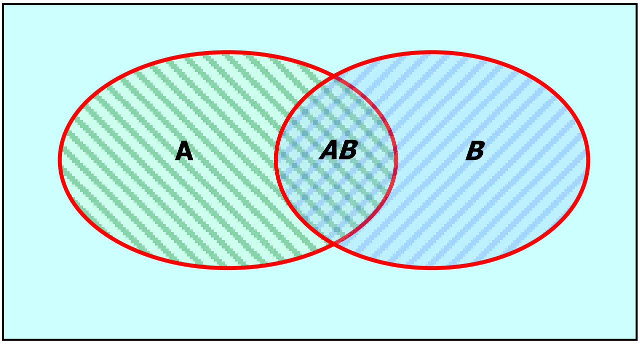


Рис.2.

Определение. **Разностью A\B** событий *A* и *B* называется событие, состоящее в том, что *A* произошло, а *B* – нет.

Пример 5.

Вернемся к примеру 1, где A\B – попадание первого стрелка при промахе второго.

Пример 6.

В примере 4 событие A\B – извлечение из колоды любой карты пиковой масти, кроме дамы. Наоборот, B\A – извлечение дамы любой масти, кроме пик.

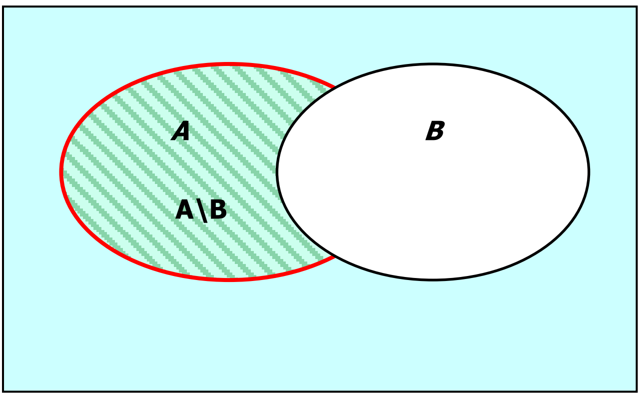


Рис.3.

Если все исходы опыта, при которых происходит событие B, содержатся в множестве исходов, при которых происходит событие A, то говорят, что $B \subseteq A$ (рис. 4).

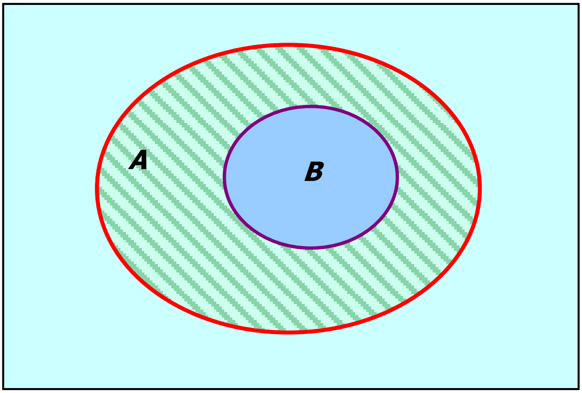


Рис. 4

Введем еще несколько категорий событий.

Определение. События *A* и *B* называются **совместными**, если они могут произойти оба в результате одного опыта. В противном случае (то есть если они не могут произойти одновременно) события называются **несовместными.**

Примеры: совместными событиями являются попадания двух стрелков в примере 1 и появление карты пиковой масти и дамы в примере 4; несовместными – события A_1 – A_6 в примере 2.

Замечание 1. Если изобразить графически области исходов опыта, благоприятных несовместным событиям, то они не будут иметь общих точек.

Замечание 2. Из определения несовместных событий следует, что их произведение является невозможным событием.

Определение. Говорят, что события $A_1, A_2, \ldots, A_n$ образуют **полную группу**, если в результате опыта обязательно произойдет хотя бы одно из событий этой группы.

Замечание. В частности, если события, образующие полную группу, попарно несовместны, то в результате опыта произойдет *одно и только одно* из них. Такие события называют **элементарными событиями**.

Пример 7.

В примере 2 события A_1 – A_6 (выпадение одного, двух,..., шести очков при одном броске игральной кости) образуют полную группу несовместных событий.

Определение. События называются **равновозможными**, если нет оснований считать, что одно из них является более возможным, чем другое.

Примеры: выпадение любого числа очков при броске игральной кости, появление любой карты при случайном извлечении из колоды, выпадение герба или цифры при броске монеты и т.п.

3.2 Классическое определение вероятности

При изучении случайных событий возникает необходимость количественно сравнивать возможность их появления в результате опыта. Например, при последовательном извлечении из колоды пяти карт более возможна ситуация, когда появились карты разных мастей, чем появление пяти карт одной масти; при десяти бросках монеты более возможно чередование гербов и цифр, нежели выпадение подряд десяти гербов, и т.д. Поэтому с каждым таким событием связывают по определенному правилу некоторое число, которое тем больше, чем более возможно событие. Это число называется **вероятностью события** и является вторым основным понятием теории вероятностей.

Отметим, что само понятие вероятности, как и понятие случайного события, является аксиоматическим и поэтому не поддается строгому определению. То, что в дальнейшем будет называться различными определениями вероятности, представляет собой способы вычисления этой величины.

Если все события, которые могут произойти в результате данного опыта,

а) попарно несовместны;

б) равновозможны;

в) образуют полную группу,

то говорят, что имеет место **схема случаев.**

Можно считать, что случаи представляют собой все множество исходов опыта. Пусть их число равно n (число возможных исходов), а при m из них происходит некоторое событие A (число благоприятных исходов).

Определение. **Вероятностью события *A*** называется отношение числа исходов опыта, благоприятных этому событию, к числу возможных исходов:

$$p(A)=\frac{m}{n} \qquad (1)$$

- *классическое определение вероятности*.

Свойства вероятности

Из определения вытекают следующие свойства вероятности:

Свойство 1. Вероятность достоверного события равна единице.

Доказательство. Так как достоверное событие всегда происходит в результате опыта, то все исходы этого опыта являются для него благоприятными, то есть $m = n$, следовательно,

$P(A) = 1$.

Свойство 2. Вероятность невозможного события равна нулю.

Доказательство. Для невозможного события ни один исход опыта не является благоприятным, поэтому $m = 0$ и $p(A) = 0$.

Свойство 3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей.

Доказательство. Случайное событие происходит при некоторых исходах опыта, но не при всех, следовательно, $0 < m < n$, и из (1) следует, что $0 < p(A) < 1$.

Пример 8.

Из урны, содержащей 6 белых и 4 черных шара, наудачу вынут шар. Найти вероятность того, что он белый.

Решение.

Будем считать элементарными событиями, или исходами опыта, извлечение из урны каждого из имеющихся в ней шаров. Очевидно, что эти события удовлетворяют всем условиям, позволяющим считать их схемой случаев. Следовательно, число возможных исходов равно 10, а число исходов, благоприятных событию А (появлению белого шара) – 6 (таково количество белых шаров в урне). Значит,

$$p(A)=\frac{m}{n}=\frac{6}{10}=0,6.$$

Классическое определение вероятности применимо только для очень узкого класса задач, где все возможные исходы опыта можно свести к схеме случаев. В большинстве реальных задач эта схема неприменима. В таких ситуациях требуется определять вероятность события иным образом. Для этого введем вначале понятие **относительной частоты *W(A)*** события *A* как отношения числа опытов, в которых наблюдалось событие *A*, к общему количеству проведенных испытаний:

$$W(A) = \frac{M}{N}, \qquad (2)$$

где *N* – общее число опытов, *M* – число появлений события *A*.

Большое количество экспериментов показало, что если опыты проводятся в одинаковых условиях, то для большого количества испытаний относительная частота изменяется мало, колеблясь около некоторого постоянного числа. Это число можно считать вероятностью рассматриваемого события.

Определение. **Статистической вероятностью события** считают его относительную частоту или число, близкое к ней.

Замечание 1. Из формулы (2) следует, что свойства вероятности, доказанные для ее классического определения, справедливы и для статистического определения вероятности.

Замечание 2. Для существования статистической вероятности события *A* требуется:

возможность производить неограниченное число испытаний;

устойчивость относительных частот появления *A* в различных сериях достаточно большого числа опытов.

Замечание 3. Недостатком статистического определения является неоднозначность статистической вероятности.

Пример 9.

Если в задаче задается вероятность попадания в мишень для данного стрелка (скажем, $p = 0{,}7$), то эта величина получена в результате изучения статистики большого количества серий выстрелов, в которых этот стрелок попадал в мишень около семидесяти раз из каждой сотни выстрелов.

3.3 Геометрическая вероятность.

Одним из недостатков классического определения вероятности является то, что оно неприменимо к испытаниям с бесконечным количеством исходов. В таких случаях можно воспользоваться понятием **геометрической вероятности**.

Пусть на отрезок *L* наудачу брошена точка. Это означает, что точка обязательно попадет на отрезок *L* и с равной возможностью может совпасть с любой точкой этого отрезка. При этом вероятность попадания точки на любую часть отрезка *L* не зависит от расположения этой части на отрезке и пропорциональна его длине. Тогда вероятность того, что брошенная точка попадет на отрезок *l*, являющийся частью отрезка *L*, вычисляется по формуле:

$$p = \frac{l}{L},$$

где *l* – длина отрезка *l*, а *L* – длина отрезка *L*.

Можно дать аналогичную постановку задачи для точки, брошенной на плоскую область *S* и вероятности того, что она попадет на часть этой области *s*:

$$p = \frac{s}{S},$$

где *s* – площадь части области, а *S* – площадь всей области.

В трехмерном случае вероятность того, что точка, случайным образом расположенная в теле *V*, попадет в его часть *v*, задается формулой:

$$p = \frac{v}{V},$$

где v – объем части тела, а *V* – объем всего тела.

Пример 1.

Найти вероятность того, что точка, наудачу брошенная в круг, не попадет в правильный шестиугольник, вписанный в него.

Решение.

Пусть радиус круга равен R, тогда сторона шестиугольника тоже равна *R*. При этом площадь круга $S = \pi R^2$, а площадь шестиугольника

$$s = \frac{3\sqrt{3}}{2} R^2.$$

Следовательно,

$$p = \frac{S - s}{S} = \frac{\pi R^2 - \frac{3\sqrt{3}}{2} R^2}{\pi R^2} = \frac{\pi - 3\sqrt{3}}{2\pi} \approx 0,174.$$

Пример 2.

На отрезок *AB* случайным образом брошены три точки: *C, D* и *M*. Найти вероятность того, что из отрезков *AC, AD* и *AM* можно построить треугольник.

Решение.

Обозначим длины отрезков *AC, AD* и *AM* через x, y и z и рассмотрим в качестве возможных исходов множество точек трехмерного пространства с координатами (x, y, z). Если принять длину отрезка равной 1, то эти множество возможных исходов представляет собой куб с ребром, равным 1. Тогда множество благоприятных исходов состоит из точек, для координат которых выполнены неравенства треугольника: $x + y > z$, $x + z > y$, $y + z > x$. Это часть куба, отрезанная от него плоскостями $x + y = z$, $x + z = y$, $y + z = x$

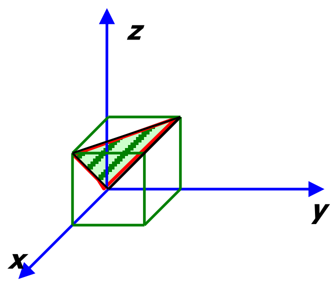


Рис. 4

(одна из них, плоскость $x + y = z$, проведена на рис.4). Каждая такая плоскость отделяет от куба пирамиду, объем которой равен

$$\frac{1}{3}\cdot\frac{1}{2}\cdot 1=\frac{1}{6}.$$

Следовательно, объем оставшейся части

$$v=1-3\cdot\frac{1}{6}=\frac{1}{2}.$$

Тогда

$$p=\frac{v}{V}=\frac{1}{2}:1=\frac{1}{2}.$$

3.4 Теоремы сложения и умножения вероятностей

Теорема сложения. Вероятность $p(A+B)$ суммы событий A и B равна

$$p(A+B)=p(A)+p(B)-p(AB). \quad (1)$$

Доказательство.

Докажем теорему сложения для схемы случаев. Пусть n – число возможных исходов опыта, m_A – число исходов, благоприятных событию A, m_B – число исходов, благоприятных событию B, а m_{AB} – число исходов опыта, при которых происходят оба события (то есть исходов, благоприятных произведению AB). Тогда число исходов, при которых имеет место событие $A+B$, равно $m_A+m_B-m_{AB}$ (так как в сумме (m_A+m_B) m_{AB} учтено дважды: как исходы, благоприятные A, и исходы, благоприятные B). Следовательно, вероятность суммы можно определить по формуле (1):

$$p(A+B)=\frac{m_A+m_B-m_{AB}}{n}=\frac{m_A}{n}+\frac{m_B}{n}-\frac{m_{AB}}{n}=p(A)+p(B)-p(AB),$$

что и требовалось доказать.

Следствие 1. Теорему сложения можно распространить на случай суммы любого числа событий. Например, для суммы трех событий A, B и C

$P(A+B+C)=p(A)+p(B)+p(C)-p(AB)-p(AC)-p(BC)+p(ABC)$

и т.д.

Следствие 2. Если события A и B несовместны, то $m_{AB}=0$, и, следовательно, вероятность суммы несовместных событий равна сумме их вероятностей:

$$p(A+B)=p(A)+p(B). \qquad (2)$$

Определение. **Противоположными событиями** называют два несовместных события, образующих полную группу. Если одно из них назвать A, то второе принято обозначать $\bar{A}$.

Замечание. Таким образом, $\bar{A}$ заключается в том, что событие A не произошло.

Теорема. Сумма вероятностей противоположных событий равна 1:

$$p(A)+p(\bar{A})=1. \qquad (3)$$

Доказательство.

Так как A и $\overline{A}$ образуют полную группу, то одно из них обязательно произойдет в результате опыта, то есть событие $A + \overline{A}$ является достоверным. Следовательно, $P(A + \overline{A}) = 1$. Но, так как A и $\overline{A}$ несовместны, из (7) следует, что

$P(A + \overline{A}) = p(A) + p(\overline{A})$.

Значит, $p(A) + p(\overline{A}) = 1$, что и требовалось доказать.

Замечание. В ряде задач проще искать не вероятность заданного события, а вероятность события, противоположного ему, а затем найти требуемую вероятность по формуле (3).

Пример 3.

Из урны, содержащей 2 белых и 6 черных шаров, случайным образом извлекаются 5 шаров. Найти вероятность того, что вынуты шары разных цветов.

Решение.

Событие $\overline{A}$, противоположное заданному, заключается в том, что из урны вынуто 5 шаров одного цвета, а так как белых шаров в ней всего два, то этот цвет может быть только черным. Множество возможных исходов опыта найдем по формуле для числа сочетаний:

$$n = C_8^5 = \frac{8!}{5! \cdot 3!} = \frac{6 \cdot 7 \cdot 8}{6} = 56,$$

а множество исходов, благоприятных событию $\overline{A}$ - это число возможных наборов по 5 шаров только из шести черных:

$$m_{\overline{A}} = C_6^5 = 6.$$

Тогда $p(\overline{A}) = \frac{6}{56} = \frac{3}{28}$, а $p(A) = 1 - \frac{3}{28} = \frac{25}{28}$.

Определение. Назовем **условной вероятностью** $\boldsymbol{p_A(B)}$ **события** $\boldsymbol{B}$ вероятность события B при условии, что событие A произошло.

Замечание.

Понятие условной вероятности используется в основном в случаях, когда осуществление события A изменяет вероятность события B.

Пример 4.

Пусть событие *A* – извлечение из колоды в 32 карты туза, а событие *B* – то, что и вторая вынутая из колоды карта окажется тузом. Тогда, если после первого раза карта была возвращена в колоду, то вероятность вынуть вторично туз не меняется: $p(B)=p(A)=\frac{4}{32}=\frac{1}{8}=0,125.$ Если же первая карта в колоду не возвращается, то осуществление события *A* приводит к тому, что в колоде осталась 31 карта, из которых только 3 туза. Поэтому

$$p_A(B)=\frac{3}{31}\approx 0,097.$$

Пример 5.

Если событие *A* – попадание в самолет противника при первом выстреле из орудия, а *B* – при втором, то первое попадание уменьшает маневренность самолета, поэтому $p_A(B)$ увеличится по сравнению с $p(A)$.

Теорема умножения. Вероятность произведения двух событий равна произведению вероятности одного из них на условную вероятность другого при условии, что первое событие произошло:

$$p(AB)=p(A)\cdot p_A(B). \qquad (4)$$

Доказательство.

Воспользуемся обозначениями теоремы сложения. Тогда для вычисления $p_A(B)$ множеством возможных исходов нужно считать m_A (так как *A* произошло), а множеством благоприятных исходов – те, при которых произошли и *A*, и *B* (m_{AB}). Следовательно,

$$p_A(B)=\frac{m_{AB}}{m_A}=\frac{m_{AB}}{n}\cdot\frac{n}{m_A}=p(AB):p(A),$$

откуда следует утверждение теоремы.

Пример 6.

Для поражения цели необходимо попасть в нее дважды. Вероятность первого попадания равна 0,2, затем она не меняется при промахах, но после первого попадания увеличивается вдвое. Найти вероятность того, что цель будет поражена первыми двумя выстрелами.

Решение.

Пусть событие A – попадание при первом выстреле, а событие B – попадание при втором. Тогда

$p(A) = 0{,}2$, $p_A(B) = 0{,}4$, $p(AB) = 0{,}2 \cdot 0{,}4 = 0{,}08$.

Следствие. Если подобным образом вычислить вероятность события BA, совпадающего с событием AB, то получим, что

$$p(BA) = p(B) \cdot p_B(A).$$

Следовательно,

$$p(A) \cdot p_A(B) = p(B) \cdot p_B(A). \quad (5)$$

Определение. Событие B называется **независимым** от события A, если появление события A не изменяет вероятности B, то есть $p_A(B) = p(B)$.

Замечание. Если событие B не зависит от A, то и A не зависит от B. Действительно, из (10) следует при этом, что

$p(A) \cdot p(B) = p(B) \cdot p_B(A)$,

откуда $p_B(A) = p(A)$. Значит, **свойство независимости событий взаимно**.

Теорема умножения для независимых событий имеет вид:

$$p(AB) = p(A) \cdot p(B), \qquad (6)$$

то есть вероятность произведения независимых событий равна произведению их вероятностей.

При решении задач теоремы сложения и умножения обычно применяются вместе.

Пример 7.

Два стрелка делают по одному выстрелу по мишени. Вероятности их попадания при одном выстреле равны соответственно 0,6 и 0,7. Найти вероятности следующих событий:

A – хотя бы одно попадание при двух выстрелах;

B – ровно одно попадание при двух выстрелах;

C – два попадания;

D – ни одного попадания.

Решение.

Пусть событие H_1 – попадание первого стрелка, H_2 – попадание второго. Тогда

$$A = H_1 + H_2,\ B = H_1 \cdot \bar{H}_2 + \bar{H}_1 \cdot H_2,\ C = H_1 \cdot H_2,\ D = \bar{H}_1 \cdot \bar{H}_2.$$

События H_1 и H_2 совместны и независимы, поэтому теорема сложения применяется в общем виде, а теорема умножения – в виде (6). Следовательно, $p(C) = 0{,}6 \cdot 0{,}7 = 0{,}42$, $p(A) = 0{,}6 + 0{,}7 – 0{,}42 = 0{,}88$, $p(B) = 0{,}6 \cdot 0{,}3 + 0{,}7 \cdot 0{,}4 = 0{,}46$ (так как события $H_1 \cdot \bar{H}_2$ и $\bar{H}_1 \cdot H_2$ несовместны), $p(D) = 0{,}4 \cdot 0{,}3 = 0{,}12$. Заметим, что события A и D являются противоположными, поэтому

$$p(A) = 1 - p(D).$$

Теорема. Вероятность появления хотя бы одного из попарно независимых событий $A_1, A_2, \ldots, A_n$ равна

$$p(A) = 1 - q_1 q_2 \ldots q_n, \qquad (7)$$

где q_i – вероятность события $\bar{A}_i$, противоположного событию A_i.

Доказательство.

Если событие A заключается в появлении хотя бы одного события из $A_1, A_2, \ldots, A_n$, то события A и $\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n$ противоположны, поэтому сумма их вероятностей равна 1. Кроме того, поскольку $A_1, A_2, \ldots, A_n$ независимы, то независимы и $\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n$, следовательно,

$$p\left(\bar{A}_1 \bar{A}_2 \ldots \bar{A}_n\right) = p(\bar{A}_1) p(\bar{A}_2) \ldots p(\bar{A}_n) = q_1 q_2 \ldots q_n.$$

Отсюда следует справедливость формулы (7).

Пример 8.

Сколько нужно произвести бросков монеты, чтобы с вероятностью не менее 0,9 выпал хотя бы один герб?

Решение.

Вероятность выпадения герба при одном броске равна вероятности противоположного события (выпадения цифры) и равна 0,5. Тогда вероятность выпадения хотя бы одного герба при n выстрелах равна $1 - (0{,}5)^n$. Тогда из решения неравенства $1 - (0{,}5)^n > 0{,}9$ следует, что $n > \log_2 10 \geq 4$.

3.5 Формула полной вероятности и формула Байеса.

Пусть событие A может произойти только совместно с одним из событий $H_1, H_2, \ldots, H_n$, образующих полную группу несовместных событий. Тогда события $H_1, H_2, \ldots, H_n$ называются **гипотезами*.***

Теорема. Вероятность события A, наступающего совместно с гипотезами H_1, H_2,..., H_n, равна:

$$p(A)=\sum_{i=1}^{n}p(H_i)p(A/H_i), \qquad (1)$$

где $p(H_i)$ – вероятность i- й гипотезы, а $p(A/H_i)$ – вероятность события A при условии реализации этой гипотезы. Формула (1) носит название **формулы полной вероятности**.

Доказательство.

Можно считать событие A суммой попарно несовместных событий AH_1, AH_2,..., AH_n. Тогда из теорем сложения и умножения следует, что

$$p(A)=p(AH_1+AH_2+...+AH_n)=$$

$$=p(AH_1)+p(AH_2)+...+p(AH_n)=\sum_{i=1}^{n}p(H_i)p(A/H_i),$$

что и требовалось доказать.

Пример 1.

Имеются три одинаковые урны с шарами. В первой из них 3 белых и 4 черных шара, во второй – 2 белых и 5 черных, в третьей – 10 черных шаров. Из случайно выбранной урны наудачу вынут шар. Найти вероятность того, что он белый.

Решение.

Будем считать гипотезами H_1, H_2 и H_3 выбор урны с соответствующим номером. Так как по условию задачи все гипотезы равновозможны, то

$$p(H_1)=p(H_2)=p(H_3)=\frac{1}{3}.$$

Найдем условную вероятность A при реализации каждой гипотезы:

$$p(A/H_1)=\frac{3}{7},\ p(A/H_2)=\frac{2}{7},\ \ p(A/H_3)=0.$$

Тогда

$$p(A)=\frac{1}{3}\cdot\frac{3}{7}+\frac{1}{3}\cdot\frac{2}{7}+\frac{1}{3}\cdot 0=\frac{5}{21}\approx 0{,}238.$$

Формула Байеса (теорема гипотез)

Пусть известен результат опыта, а именно то, что произошло событие A. Этот факт может изменить априорные (то есть известные до опыта) вероятности гипотез. Например, в предыдущем примере извлечение из урны белого шара говорит о том, что этой урной не могла быть третья, в которой нет белых шаров, то есть $p(H_3/A) = 0$. Для переоценки вероятностей гипотез при известном результате опыта используется **формула Байеса**:

$$p(H_i / A) = \frac{p(H_i)p(A / H_i)}{p(A)}. \qquad (2)$$

Действительно, из теоремы умножения получим, что

$$p(A)p(H_i / A) = p(H_i)p(A / H_i),$$

откуда следует справедливость формулы (2).

Пример 2.

После двух выстрелов двух стрелков, вероятности попаданий которых равны 0,6 и 0,7, в мишени оказалась одна пробоина. Найти вероятность того, что попал первый стрелок.

Решение.

Пусть событие A – одно попадание при двух выстрелах, а гипотезы: H_1 – первый попал, а второй промахнулся, H_2 – первый промахнулся, а второй попал, H_3 – оба попали, H_4 – оба промахнулись. Вероятности гипотез: $p(H_1) = 0{,}6 \cdot 0{,}3 = 0{,}18$, $p(H_2) = 0{,}4 \cdot 0{,}7 = 0{,}28$, $p(H_3) = 0{,}6 \cdot 0{,}7 = 0{,}42$, $p(H_4) = 0{,}4 \cdot 0{,}3 = 0{,}12$. Тогда

$p(A/H_1) = p(A/H_2) = 1$, $p(A/H_3) = p(A/H_4) = 0$.

Следовательно, полная вероятность

$p(A) = 0{,}18 \cdot 1 + 0{,}28 \cdot 1 + 0{,}42 \cdot 0 + 0{,}12 \cdot 0 = 0{,}46$.

Применяя формулу Байеса, получим:

$$p(H_1 / A) = \frac{0{,}18 \cdot 1}{0{,}46} = \frac{9}{23} \approx 0{,}391.$$

3.6 Схема повторения испытаний. Формула Бернулли

Рассмотрим серию из n испытаний, в каждом из которых событие A появляется с одной и той же вероятностью p, причем результат каждого испытания не зависит от результатов остальных. Подобная постановка задачи называется **схемой повторения испытаний**. Найдем вероятность того, что в такой серии событие A произойдет ровно k раз (неважно, в какой последовательности). Интересующее нас событие представляет собой сумму равновероятных несовместных событий, заключающихся в том, что A произошло в некоторых k испытаниях и не произошло в остальных $n - k$ испытаниях. Число таких событий равно числу сочетаний из n по k, то есть C_n^k, а вероятность каждого из них: $p^k q^{n-k}$, где $q = 1 - p$ – вероятность того, что в данном опыте A не произошло. Применяя теорему сложения для несовместных событий, получим **формулу Бернулли**:

$$p_n(k) = C_n^k \cdot p^k \cdot q^{n-k}. \qquad (3)$$

Пример 3.

Для получения приза нужно собрать 5 изделий с особым знаком на этикетке. Найти вероятность того, что придется купить 10 изделий, если этикетки с этим знаком имеют 5% изделий.

Решение.

Из постановки задачи следует, что последнее купленное изделие имеет особый знак. Следовательно, из предыдущих девяти эти знаки имели 4 изделия. Найдем вероятность этого по формуле Бернулли:

$$p_9(4) = C_9^4 \cdot (0,05)^4 \cdot (0,95)^5 = 0,0006092.$$

Тогда p = 0,0006092·0,05 = 0,0000304.

Приближение Пуассона для схемы Бернулли

Формула Бернулли требует громоздких расчетов при большом количестве испытаний. Можно получить более удобную для расчетов приближенную формулу, если при большом числе испытаний вероятность появления A в одном опыте мала, а произведение $np = \lambda$ сохраняет постоянное значение для разных серий опытов (то есть среднее число появлений события A в разных сериях испытаний остается неизменным). Применим формулу Бернулли:

$$p_n(k) = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!} p^k (1-p)^{n-k} =$$

$$= \frac{n(n-1)\ldots(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

Найдем предел полученного выражения при $n \to \infty$:

$$p_n(k) \approx \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\ldots\left(1 - \frac{k-1}{n}\right)\left(1 - \frac{\lambda}{n}\right)^{n-k}\right) =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1.$$

Таким образом, **формула Пуассона**

$$p_n(k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad (4)$$

позволяет найти вероятность *k* появлений события *A* для массовых (*n* велико) и редких (*p* мало) событий.

3.7 Решение типовых задач

Задача 1.

Среди студентов, собравшихся на лекцию по теории вероятностей, выбирают наудачу одного. Пусть событие *B* заключается в том, что он не курит, а событие *C* в том, что он живет в общежитии. Когда будет справедливо соотношение $\overline{C} \subseteq B$?

РЕШЕНИЕ. Событие, противоположное событию *C*, происходит только в случае, если произошло событие *B*. Событие $\overline{C}$ заключается в том, что выбранный студент НЕ живет в общежитии. Оно должно происходить всегда, если произошло событие *B*, следовательно, все студенты, не живущие в общежитии, не курят. Однако событие *B* может произойти и в случае, если $\overline{C}$ не произошло, то есть нельзя утверждать, что не курят только студенты, не живущие в общежитии. Следовательно, верный ответ: все студенты, не живущие в общежитии, не курят.

ОТВЕТ: все студенты, не живущие в общежитии, не курят.

Задача 2.

Какова вероятность того, что наудачу взятую кость домино можно приставить к данной: (2; 5)?

Решение. В наборе домино 28 костей, соответствующих неупорядоченным парам чисел (0,0), (0,1),…, (6,6). Поскольку одна кость извлечена из набора, число *n* возможных исходов опыта равно 27 – числу оставшихся костей. Каждое из чисел 2 и 5 присутствует на семи костях, одна из которых: (2,5) – общая и в следующем выборе не участвует. Следовательно, в наборе осталось 6 + 6 = 12 костей, каждую из которых можно приставить к данной, то есть число *m* благоприятных исходов равно 12. Поэтому

$$p = \frac{m}{n} = \frac{12}{27} = \frac{4}{9}.$$

ОТВЕТ: $\frac{4}{9}$.

Задача 3.

Найти вероятность того, что при трех бросках монеты герб выпадет 1 раз.

РЕШЕНИЕ Каждый бросок монеты может иметь 2 различных исхода, поэтому общее число n исходов при трех бросках равно 2·2·2 = 8. Очевидно, что при этом выпадение ровно одного герба возможно трижды: при первом, втором или третьем броске. Следовательно, m = 3, а вероятность

$$p = \frac{m}{n} = \frac{3}{8}.$$

ОТВЕТ: $\frac{3}{8}$.

Задача 4.

Из коробки, в которой лежат 6 красных и 5 синих карандашей, случайным образом извлечены 3 карандаша. Найти вероятность того, что все они красные.

РЕШЕНИЕ. Число возможных исходов опыта – число сочетаний из 11 (общее число карандашей в коробке) по 3, а число благоприятных исходов - число сочетаний из 6 (количество карандашей нужного цвета) по 3.

$$n = C_{11}^{3} = \frac{11!}{3!\cdot 8!} = \frac{9\cdot 10\cdot 11}{1\cdot 2\cdot 3} = 165,$$

$$m = C_{6}^{3} = \frac{6!}{3!\cdot 3!} = \frac{4\cdot 5\cdot 6}{1\cdot 2\cdot 3} = 20, \quad p = \frac{m}{n} = \frac{20}{165} = \frac{4}{33}.$$

ОТВЕТ: $\frac{4}{33}$

Задача 5.

Из 10 служащих туристического агентства 7 знают французский язык. Какова вероятность того, что среди трех человек, находящихся в данный момент в офисе, французским владеют двое?

РЕШЕНИЕ Число благоприятных исходов – это произведение количества способов, которыми можно выбрать двух человек, знающих французский язык, из семи, и способов, которыми можно выбрать одного человека, не знающего этого языка, из оставшихся трех.

$$n = C_{10}^{3} = \frac{10!}{3!\cdot 7!} = 120,$$

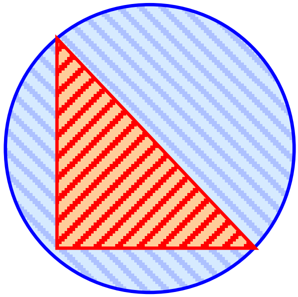
$$m = C_{7}^{2}\cdot C_{3}^{1} = \frac{7!}{2!\cdot 5!}\cdot 3 = 63, \quad p = \frac{63}{120} = \frac{21}{40}.$$

ОТВЕТ: $\frac{21}{40}$.

Задача 6.

Найти вероятность того, что точка, брошенная наудачу в круг, попадет в равнобедренный прямоугольный треугольник, вписанный в этот круг.

Решение. Искомая вероятность равна отношению площадей треугольника и круга.



Используем формулу геометрической вероятности:

$$p = \frac{s}{S},$$

где S – мера (в данном случае площадь) области возможных исходов, а s – площадь области благоприятных исходов.

Если радиус круга равен R, то гипотенуза прямоугольного треугольника является диаметром окружности и и равна $2R$, а высота, опущенная на гипотенузу, равна R. Следовательно,

$$s_\Delta = \frac{1}{2} \cdot 2R \cdot R = R^2, \quad S_{круга} = \pi R^2,$$

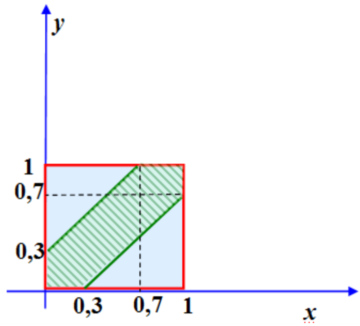
$$p = \frac{s_\Delta}{S_{круга}} = \frac{R^2}{\pi R^2} = \frac{1}{\pi}.$$

Ответ: $\frac{1}{\pi}$.

Задача 7.

Числа x и y могут равновозможно принимать любые значения из промежутка $[0;1]$. Найти вероятность того, что $|x - y| < 0{,}3$.

Решение. Будем считать, что данные числа являются координатами точки на плоскости. Тогда множеством возможных исходов является квадрат со стороной, равной единице, а множество благоприятных исходов – область, заданная неравенством $|x - y| < 0,3$.



Зададим множество возможных исходов опыта в виде квадрата $0 \leq x \leq 1$, $0 \leq y \leq 1$. Тогда множество благоприятных исходов представляет собой часть этого квадрата, расположенную между прямыми $x - y = 0,3$ и $x - y = -0,3$. Найдем площадь этой области как разность площади квадрата, равной единице, и суммы площадей прямоугольных треугольников, катеты которых равны 0,7:

$$s = 1 - 2\cdot\frac{1}{2}\cdot 0,7^2 = 1 - 0,49 = 0,51.$$

$$p = \frac{s}{S} = \frac{0,51}{1} = 0,51.$$

Ответ: 0,51.

Задача 8.

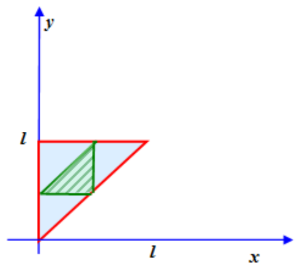
На отрезке длиной l наудачу ставятся две точки. Найти вероятность того, что из полученных трех отрезков можно построить треугольник.

Решение.

Пусть OC – данный отрезок, A и B – поставленные на нем точки. Рассмотрите в качестве множества возможных исходов прямоугольный треугольник с катетами, равными *l*, в котором координаты каждой точки равны длинам отрезков *OA* и *OB*. Множество возможных исходов задайте с помощью неравенства треугольника (сумма любых двух сторон треугольника больше третьей стороны). Пусть *OC* – данный отрезок, *A* и *B* – поставленные на нем точки. Обозначим длины отрезков *OA* и *OB* соответственно за x и y ($y > x$). Тогда множество возможных исходов опыта представляет собой треугольник $0 \le x \le l$, $0 \le y \le l$, $y > x$. Определим множество благоприятных исходов с помощью неравенства треугольника:

$$\begin{cases} OA + AB > BC \\ OA + BC > AB \\ AB + BC > OA \end{cases} \Rightarrow \begin{cases} y > l - y \\ x + l - y > y - x \\ l - x > x \end{cases} \Rightarrow$$

$$\Rightarrow \begin{cases} y > \dfrac{l}{2} \\ y - x < \dfrac{l}{2} \\ x < \dfrac{l}{2} \end{cases}.$$



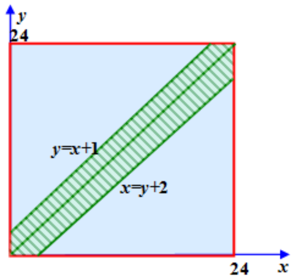
Из рисунка видно, что множество благоприятных исходов является треугольником, площадь которого в 4 раза меньше площади множества возможных исходов. Следовательно, искомая вероятность равна $\dfrac{1}{4}$.

Ответ: $\dfrac{1}{4}$.

Задача 9.

Два парохода должны подойти к одному и тому же причалу. Время прихода обоих пароходов независимо и равновозможно в течение данных суток. Найти вероятность того, что одному из пароходов придется ожидать освобождения причала, если время стоянки первого парохода – один час, а второго – два часа (ответ округлить до третьего знака после запятой).

Решение. Примите за x время прибытия первого парохода, за y – время прибытия второго, и рассмотрите в качестве множества возможных исходов опыта квадрат $0 \le x \le 24$, $0 \le y \le 24$.



Если принять за x время прибытия первого парохода, за y – время прибытия второго, то множество возможных исходов опыта является квадратом $0 \le x \le 24$, $0 \le y \le 24$. Множество благоприятных исходов задается условиями $x \le y \le x + 1$, $y \le x \le y + 2$.

Найдем площадь полученной области:

$$s = 24^2 - \frac{1}{2} \cdot 23^2 - \frac{1}{2} \cdot 22^2 = 69,5.$$

$$S = 24^2 = 576, \quad p = \frac{s}{S} = \frac{69,5}{576} \approx 0,121.$$

Ответ: 0,121.

Задача 10.

Из полного набора костей домино наудачу извлекаются три кости. Найти вероятность того, что все они – дубли.

Решение. Примените теорему умножения для зависимых событий.

Будем считать, что кости извлекаются по одной. Событием A назовем извлечение дубля в первый раз, событием B – извлечение второго дубля, событием C – третьего. Эти события являются зависимыми, так как после извлечения каждой очередной кости изменяется и число оставшихся костей (28, затем 27, затем 26), и количество дублей среди них (7, затем 6, затем 5).

Нам требуется найти вероятность произведения ABC.

$$p(A)=\frac{7}{28}=\frac{1}{4};\quad p_A(B)=\frac{6}{27}=\frac{2}{9};\quad p_{AB}(C)=\frac{5}{26};$$

$$p(ABC)=p(A)\cdot p_A(B)\cdot p_{AB}(C)=\frac{1}{4}\cdot\frac{2}{9}\cdot\frac{5}{26}=\frac{5}{468}.$$

Ответ: $\frac{5}{468}$.

Задача 11.

В первой урне находится 5 белых шаров, 11 черных и 8 красных, во второй соответственно 10, 8 и 6. Из каждой урны наудачу извлекается по одному шару. Найти вероятность того, что они одного цвета (ответ округлить до третьего знака после запятой).

Решение Событие, вероятность которого требуется определить, представляет собой сумму для трех несовместных событий: A – извлечены белые шары, B – черные, C – красные. Каждое их них, в свою очередь, является произведением двух независимых событий: извлечение шара нужного цвета из первой и второй урны. Применим теорему умножения для независимых событий и теорему сложения для несовместных событий:

$$p(A)=\frac{5}{24}\cdot\frac{10}{24}=\frac{25}{288};\quad p(B)=\frac{11}{24}\cdot\frac{8}{24}=\frac{11}{72};$$

$$p(C)=\frac{8}{24}\cdot\frac{6}{24}=\frac{1}{12};$$

$$p(A+B+C)=\frac{25}{288}+\frac{11}{72}+\frac{1}{12}=\frac{31}{96}\approx 0,323.$$

Ответ: 0,323.

Задача 12.

Найти вероятность выпадения хотя бы одной шестерки при четырех бросках игральной кости (ответ округлить до третьего знака после запятой).

Решение. Найдем вначале вероятность события, противоположного данному.

Событие, противоположное данному, заключается в том, что при четырех бросках шестерка не выпала ни разу. Найдем его вероятность, используя теорему умножения для независимых событий:

$$p(\bar{A})=\left(\frac{5}{6}\right)^4=\frac{625}{1296};\quad p(A)=1-p(\bar{A})=\frac{671}{1296}\approx 0,518.$$

Ответ: 0,518.

Задача 13.

В первой урне 4 белых шара и 6 черных, во второй 1 белый и 9 черных, в третьей 10 черных шаров. Из наудачу выбранной урны вынут шар. Найти вероятность того, что он белый.

Решение Используем формулу полной вероятности, считая гипотезами выбор урны (все три гипотезы равновероятны).

Пусть событие A – извлечение белого шара, гипотезы H_1, H_2, H_3 – выбор соответствующей урны. Тогда

$$p(H_1)=p(H_2)=p(H_3)=\frac{1}{3},$$

$$p_{H_1}(A)=\frac{4}{10},\quad p_{H_2}(A)=\frac{1}{10},\quad p_{H_3}(A)=\frac{0}{10}=0,$$

$$p(A)=\frac{1}{3}\cdot\frac{4}{10}+\frac{1}{3}\cdot\frac{1}{10}+\frac{1}{3}\cdot 0=\frac{1}{6}.$$

Ответ: $\frac{1}{6}$.

Задача 14.

В ящике 15 теннисных мячей, из которых 9 новых. Для первой игры наугад берутся три мяча, которые после игры возвращаются в ящик. Для второй игры также наугад берутся три мяча. Найти вероятность того, что все мячи, взятые для второй игры, новые (ответ округлить до третьего знака после запятой).

Решение Пусть событие A – выбор для второй игры трех новых мячей.

Зададим пространство гипотез следующим образом:

H_1 – все мячи, взятые для первой игры, старые (тогда после первой игры в ящике по-прежнему 9 новых мячей);

H_2 – среди мячей, взятых для первой игры, один новый (тогда после первой игры новых мячей осталось 8);

H_3 – среди мячей, взятых для первой игры, два новых (для второй игры остается 7 новых);

H_4 – все мячи, взятые для первой игры, новые (для второй игры остается 6 новых).

Найдем вероятности гипотез, используя классическое определение вероятности и формулу для числа сочетаний:

$$p(H_1)=\frac{C_6^3}{C_{15}^3}=\frac{20}{455}=\frac{4}{91};\quad p(H_2)=\frac{C_6^2\cdot C_9^1}{C_{15}^3}=\frac{15\cdot 9}{455}=\frac{27}{91};$$

$$p(H_3)=\frac{C_6^1\cdot C_9^2}{C_{15}^3}=\frac{6\cdot 36}{455}=\frac{216}{455};\quad p(H_4)=\frac{C_9^3}{C_{15}^3}=\frac{84}{455}=\frac{12}{65}.$$

Аналогичным образом, учитывая изменение количества новых мячей перед второй игрой в зависимости от выбора гипотезы, найдем условную вероятность события *A* при реализации каждой гипотезы:

$$p_{H_1}(A)=\frac{C_9^3}{C_{15}^3}=\frac{12}{65},\quad p_{H_2}(A)=\frac{C_8^3}{C_{15}^3}=\frac{56}{455}=\frac{8}{65},$$

$$p_{H_3}(A)=\frac{C_7^3}{C_{15}^3}=\frac{35}{455}=\frac{1}{13},\quad p_{H_1}(A)=\frac{C_6^3}{C_{15}^3}=\frac{4}{91}.$$

Теперь можно найти полную вероятность события *A*:

$$p(A)=\frac{4}{91}\cdot\frac{12}{65}+\frac{27}{91}\cdot\frac{8}{65}+\frac{216}{455}\cdot\frac{1}{13}+\frac{12}{65}\cdot\frac{4}{91}=$$

$$=\frac{528}{5915}\approx 0,089.$$

Ответ: 0,089.

Задача 15.

На двух станках изготовляются одинаковые детали. Вероятность брака для детали, изготовленной на первом станке, равна 0,02, на втором – 0,05. Из ящика, в котором находятся 20 деталей, изготовленных на первом станке, и 30, изготовленных на втором, случайно выбрана деталь, которая после проверки оказалась доброкачественной. Найти вероятность того, что она изготовлена на втором станке.

Решение. Примените формулу Байеса, выбрав в качестве гипотез изготовление детали на первом или втором станке. Тогда требуется, зная результат опыта, переоценить вероятность гипотезы H_2.

Пусть гипотеза H_1 – деталь изготовлена на первом станке, H_2 – на втором, событие *A* – выбранная деталь доброкачественная. Найдем полную вероятность события *A*:

$$p(H_1)=\frac{20}{50}=0,4, \quad p(H_2)=\frac{30}{50}=0,6,$$

$$p_{H_1}(A)=1-0,02=0,98, \quad p_{H_2}(A)=1-0,05=0,95,$$

$$p(A)=0,4\cdot 0,98+0,6\cdot 0,95=0,962.$$

Применяя формулу Байеса, найдем вероятность гипотезы H_2 при условии, что событие A произошло:

$$p_A(H_2)=\frac{0,6\cdot 0,95}{0,962}=\frac{285}{481}\approx 0,593.$$

Ответ: 0,593.

Задача16.

В некоторой местности 20% дней в году – дождливые и 80% - солнечные. Прогноз погоды дают два предсказателя. Первый ошибается в 30% случаев, второй – в 10%. Первый утверждает, что завтра будет солнце, второй – что дождь. Чей прогноз сбудется с большей вероятностью?

Решение Пусть гипотеза H_1 заключается в том, что в выбранный день будет солнечная погода, а гипотеза H_2 – дождливая. Тогда при реализации первой гипотезы прогноз первого предсказателя верен, а второго – ошибочен, а при реализации второй гипотезы – наоборот. Найдем полную вероятность события A (выдачи соответствующих прогнозов):

$$p(H_1)=0,8, \quad p(H_2)=0,2,$$

$$p_{H_1}(A)=0,7\cdot 0,1=0,07, \quad p_{H_2}(A)=0,3\cdot 0,9=0,27,$$

$$p(A)=0,8\cdot 0,07+0,2\cdot 0,27=0,11.$$

Итак, вероятность получения именно таких взаимоисключающих прогнозов равна 0,11. Переоценим вероятность первой гипотезы, зная, что событие A произошло. Если она окажется больше 0,5, то с большей вероятностью сбудется прогноз первого предсказателя, если меньше 0,5 – прогноз второго.

$$p_A(H_1)=\frac{0,8\cdot 0,07}{0,11}=\frac{56}{110}>\frac{1}{2} \quad -$$

более вероятно, что сбудется прогноз первого предсказателя.

Ответ: прогноз первого.

Задача 17.

Найти вероятность выпадения 2-х шестерок при 5 бросках игральной кости.

Решение Воспользуемся формулой Бернулли.

$$n=5, \quad k=2, \quad p=\frac{1}{6}, \quad q=1-\frac{1}{6}=\frac{5}{6},$$

$$p_5(2)=C_5^2\cdot\left(\frac{1}{6}\right)^2\cdot\left(\frac{5}{6}\right)^3=10\cdot\frac{1}{36}\cdot\frac{125}{216}=\frac{625}{3888}\approx 0{,}161.$$

Ответ: 0,161.

Задача 18.

Стрелку выдано 5 патронов для поражения трех мишеней. Вероятность попадания при каждом выстреле равна 0,4. Найти вероятность того, что он израсходует 5 патронов и поразит все цели.

Решение Событие, вероятность которого требуется определить, заключается в том, что последним, пятым выстрелом поражена последняя, третья мишень. Следовательно, при первых четырех выстрелах было ровно два попадания.

Представим событие, вероятность которого требуется определить, в виде произведения двух независимых событий: *A* – два попадания при четырех выстрелах, *B* – попадание при пятом выстреле. Тогда

$$p(A)=C_4^2\cdot 0{,}4^2\cdot 0{,}6^2=6\cdot 0{,}16\cdot 0{,}36=0{,}3456;$$

$$p(B)=0{,}4; \quad p(AB)=0{,}3456\cdot 0{,}4=0{,}13824\approx 0{,}138.$$

Ответ: 0,138.

Задача 19.

Для поражения цели произвели 7 выстрелов. Вероятность попадания при каждом выстреле равна 0,6. При одном попадании цель выходит из строя с вероятностью 0,1, при двух – с вероятностью 0,5, при трех – всегда выходит из строя. Найти вероятность выхода цели из строя после 7 выстрелов.

Решение. Пусть событие *A* – выход цели из строя. Вероятность этого события зависит от количества попаданий, поэтому рассмотрим гипотезы H_1 – одно попадание из семи выстрелов, H_2 – два попадания, H_3 – три или более попаданий. Вероятности гипотез можно найти по формуле Бернулли:

$$p(H_1) = C_7^1 \cdot 0,6 \cdot 0,4^6 \approx 0,0172;$$

$$p(H_2) = C_7^2 \cdot 0,6^2 \cdot 0,4^5 \approx 0,0774;$$

$$p(H_3) = 1 - p(H_1) - p(H_2) - p(0) \approx$$

$$\approx 1 - 0,0172 - 0,0774 - 0,4^7 \approx 0,9038;$$

$$p(A) = 0,0172 \cdot 0,1 + 0,0774 \cdot 0,5 + 0,9038 \cdot 1 \approx 0,944.$$

Ответ: 0,944.

3.8 Задачи к главе 3

Задача 1.

В ящике 15 шаров: 7 синих и 8 желтых. Наудачу из ящика вынули один шар, а затем второй (не возвращая их обратно). Найти вероятность того, что первый из взятых шаров синий, а второй желтый.

Задача 2.

На контроль поступают детали с двух станков. Производительность станков не одинакова. На первом станке изготовляют 60% всех деталей, на втором – 40% . Вероятность брака на первом станке 0,002, на втором – 0,004. Найти вероятность того, что поступившая на контроль деталь бракованная.

Задача 3.

В условиях примера 1, проверенная деталь оказалась бракованной. Определить вероятность того, что она была изготовлена на первом станке.

Задача 4.

Из колоды в 36 карт наудачу внимают три карты. Найти вероятность того, что они будут одной и той же масти.

Задача 5.

Из колоды в 36 карт извлекается одна карта, после чего она возвращается назад. Потом из колоды извлекаются две карты. Определить вероятность того, что три карты будут одной и той же масти.

Задача 6.

Пять экзаменаторов принимают экзамен. Известно, что вероятность сдать экзамен двум из них ("строгим") равна 0.6, а трем остальным ("нестрогим") 0.8. Известно, что студент сдал экзамен. Найти вероятность того, что он сдавал "нестрогому" экзаменатору.

Задача 7.
Партия деталей содержит 20% деталей, изготовленных заводом 1, 30%-заводом 2, 50%-заводом 3. Для завода 1 вероятность выпуска бракованной детали равна 0.05, для завода 2 - 0.01, для завода 3 - 0.06. Чему равна вероятность того, что наудачу взятая из партии деталь окажется бракованной?

Задача 8.
Среди n лотерейных билетов m выигрышных. Какова вероятность выиграть для лица, покупающего один билет, если перед этим было куплено только два билета?

Задача 9.
У ребенка в руках 8 карт с номерами 1,1,1,3,3,5,5 и 5. Он берет их по очереди и кладет обратно 6 карт. Найти вероятность, что до появления цифры 5 у него было 4 попытки. (Выбор до происхождения определенного события.)

Задача 10.
Берут два случайных (положительных) числа Х и У, которые оба меньше 4. Найти вероятность, что сумма этих чисел больше 1 (Х+У>1) , а разность меньше 3 (Х-У<3). (геометрическая теория вероятности)

Задача 11.
Из 20 сбербанков 10 расположены за чертой города. Для аудиторской проверки случайно выбрано 5 банков. Какова вероятность, что хотя бы 2 из них окажется в черте города.

Задача 12.
В связке 5 разных ключей, и один из них соответствующей двери. Делается попытка открыть наудачу взятым ключем, ключ неподходящий более не используется. Найти вероятность того, что

а) дверь будет открыта 1-ым ключом;

б) для открытия двери будет использовано не более двух ключей.

Задача 13.

Вероятность выигрыша по лотерейному билету p=1/7. Какова вероятность того, что обладатель 5 билетов выиграет:

а) по всем 5;

б) ни по одному;

в) хотя бы по одному билету?

Задача 14.

С 1-го станка на сборку поступает 40 %, со 2-го – 30 %, с 3-го – 20 %, с 4-го – 10 %. Вероятности брака для каждого из станков 0,1 %, 0,2 %, 0,25 %, 0,5 % соответственно. Найти вероятность того, что поступившая на сборку деталь – бракованная.

Задача 15.

Для участия в спорт. соревнованиях из 1-ой группы было выделено 4 студента; из 2-ой -6 ; из 3-й – 5 студентов. Вероятность того, что студент каждый из групп попадает в сборную института равны 0,5 ; 0,4; 0,3 соотв. для каждой из групп. Наудачу выбранный участник попал в сборную. К какой из 3-х групп он вероятнее всего принадлежит?

СПИСОК ЛИТЕРАТУРЫ

1. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика. Основы моделирования и первичная обработка данных.-М.:Финансы и статистика, 1983.

2.Вентцель Е.С. Теория вероятностей.-М.:Наука, 1964.

3.Гмурман В.Е. Введение в теорию вероятностей и математическую статистику.-М.:Наука,1970.

4. Гмурман В.Е. Теория вероятностей и математическая статистика.-М.:Высшая школа,1998.

5. Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике.-М.:Высшая школа,1979.

6.Павловский З.В. Введение в математическую статистику.-М.:Статистика,1967.