МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ КЫРГЫЗСКОЙ РЕСПУБЛИКИ

КЫРГЫЗСКИЙ ГОСУДАРСТВЕННЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
им. И. РАЗЗАКОВА

КРАТКИЙ КУРС ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ. КОНТРОЛЬНЫЕ ЗАДАНИЯ И МЕТОДИЧЕСКИЕ УКАЗАНИЯ К РЕШЕНИЮ ЗАДАЧ

БИШКЕК 2011

Рекомендованы к публикации на заседании кафедры прикладной математики, протокол № 8 от 27 апреля 2011г., Ученым советом Кыргызского Государственного технического университета им. И. Раззакова, протокол № 9 от 30 мая 2011г.

СОСТАВИТЕЛЬ: ДЖАМАНБАЕВ М.Дж. – д.ф.-.м.н., профессор, кафедры прикладной математики КГТУ им.И. Раззакова.

РЕЦЕНЗЕНТЫ:

Уметалиев М.У. – к.ф.-м.н., доцент, кафедры прикладной математики КГТУ им. И. Раззакова

Батырканов М. - к.ф.- м.н.,профессор, заф. кафедрой математической статистики Кыргызского Национального Аграрного Университета им. К.И. Скрябина

Токтакунов Т. – к.ф.- м.н., доцент, зав. кафедрой прикладной информатики Кыргызского Экономического Университета

ПРЕДИСЛОВИЕ

Учебное пособие написана согласно ГОС для инженерно – технических специальностей по дисциплине теория вероятностей и математическая статистика и содержит краткое теоретическое изложение разделов дисциплины в виде определений и теорем с приведением многих примеров и указаний к решению задач. Материалы закрепляются вопросами для самопроверки. В конце пособии приводится сборник контрольных заданий, перечень модульных вопросов по теории вероятностей и математической статистике и правила выполнения контрольных заданий. Такое построение пособии удобно в подготовке лекции и при проведении практических занятий. При написании пособии использованы учебники, приведенные в списке литературы, а также материалы автора при чтении курса лекции по дисциплине теории вероятностей и математической статистике. Полезность пособии заключается в возможности использования студентами как очной, так и заочной формы обучения.

I.ОСНОВНЫЕ ПОНЯТИЯ И НЕПОСРЕДСТВЕННЫЙ ПОДСЧЕТ ВЕРОЯТНОСТЕЙ.

В природе и в любом процессе существуют три вида событии *достоверные, недостоверные и случайные. Достоверные* события, которые обязательно произойдет. Например 1 января будет новый год. *Недостоверные* события, которые никогда не произойдут. Например летом вода превратится в лед. *Случайные* события, которые могут или не могут произойти при одинаковых обстоятельствах. Среди трех видов событий интерес представляет *случайные* события. Установлением закономерностей не единичных, а массовых *случайных* событий занимается наука теория вероятностей. В результате испытания или эксперимента или опыта произойдет событие.

Испытанием мы будем называть тип опыта (эксперимента). Например, извлечение наудачу карты из колоды – испытание. Бросание наудачу игральной кости (монеты) – испытание.

Пример. Пусть испытание – извлечение карты из колоды. Тогда событиями являются: A– извлечена карты красной масти, B – извлечена “ картинка“, C – извлечен туз и т.п. Если в результате конкретного испытания из колоды достали, например, семерку бубен, то событие A наступило, события B и C – нет.

Пример. Пусть испытание – бросание игральной кости. Тогда событиями являются: A – число выпавших очков – четно, B – число выпавших очков – больше 4, C– на верхней грани игральной кости выпала “5”. *Случайные* события обозначаются большими буквами A, B, C, ….. Всякую *случайную событию* можно заранее численно охарактеризовать возможность появления. ***Эта численная характеристика называется вероятностью случайной событии.*** Вероятности события A обозначается P(A). Имеются различные виды случайных событий: *Два события называются **равновозможными,** если вероятности их наступления равны*

Пример. Пусть испытание – бросание монеты. Тогда события A – выпадение “орла” и B – выпадение “решки” являются равновозможными.

*Два события называются **несовместными (несовместимыми)**, если они не могут наступить одновременно.*

Пример. Испытание – извлечение карты из колоды. Если событие A – извлечена карта красной масти, событие B – извлечена карта черной масти, то A и B – несовместны.

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

Говорят, что события $A_1, A_2, \ldots, A_k$ образуют ***полную систему (группу),*** *если в результате испытания обязательно одно из них произойдет*

Пример. Пусть испытание – бросание игральной кости. Тогда события $A_1 = \{1\}, A_2 = \{2\}, \ldots, A_6 = \{6\}$ образуют полную группу сбытий.

Пример. Пусть по мишени производится 3 выстрела и m – число попаданий в мишень. Тогда события, например, $(m = 0), (1 \le m \le 2), (m = 3)$ образуют полную группу событий.

Если два события образуют полную систему, то они называются парой ***взаимно противоположных*** *событий.*
Если одно из событий такой пары обозначено, скажем, через A, другое будет обозначено $\overline{A}$.

Пример. Пусть испытание – бросание монеты. Тогда события A – выпадение “орла” и B – выпадение “решки” являются взаимно противоположными ($B = \overline{A}$).

Пример. Пусть по мишени производится 3 выстрела, и m – число попаданий в мишень. Тогда события, например, $(m < 2) = (m = 0 \text{ или } m = 1)$ и $(m \ge 2) = (m = 2 \text{ или } m = 3)$ – взаимно противоположны.

2. Классическое определение вероятности

Определение. *Пусть некоторое испытание имеет n исходов, причем эти исходы а) попарно несовместимы; б) единственно возможны; в) равновозможны и наступлению события A благоприятствует m исходов из n. Тогда вероятность $P(A)$ наступления события A (в одном испытании) определяется по формуле*

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

Пример. В коробке имеется 10 хороших деталей и 5 бракованных. Наудачу из коробки извлекается одна деталь. Найти вероятность наступления события A – извлеченная деталь – хорошая.

Решение. Общее число исходов $n = 15$ равно полному числу деталей в коробке. Извлечению хорошей детали благоприятствует $m = 10$ исходов из общего числа (число хороших деталей). Тогда

$$P(A) = \frac{10}{15} = \frac{2}{3}.$$

3. Статистическое определение вероятности события.

Пусть проведено N испытаний, в которых некоторое событие A наступает N_A раз. Тогда отношение $\frac{N_A}{N}$ называется *частостью (долей)* наступления события A в N испытаниях.

Определение. *Пусть условия проведения некоторого испытания можно в точности воспроизвести неограниченное число раз. Тогда* ***вероятностью*** *$P(A)$ наступления события A (в одном испытании) называется такое число, около которого группируются значения частости W(A) = $\frac{N_A}{N}$ при неограниченном*

увеличении числа испытаний N.
Символически это определение можно записать в виде

$$P(A) = \lim_{N \to \infty} \frac{N_A}{N}.$$

Отметим практическое следствие данного определения: если нас интересует значение вероятности наступления некоторого события A, то производят достаточно большое число испытаний *N,* по их результатам определяют значение частости W(A) = $\frac{N_A}{N}$ и затем полагают

$$P(A) \cong \frac{N_A}{N}. \qquad \mathrm{P(A)} \cong W(A).$$

(см. Закон больших чисел, теорему Бернулли.)
Пример: Статистическая вероятность: Брошено монета 10 раз. Герб появился 8 раз т.е. n = 10; m = 8

$$p = w = \frac{m}{n} = \frac{8}{10} = 0,8$$

Замечание:
1.Статистическая вероятность может быть найдена только после проведения опытов, а для классической вероятности опыты не нужны.
2.Статистическая вероятность получается различной для разных серий опытов, однако при достаточно большом количестве опытов практически достоверно, что статистическая вероятность будет сколь угодно мало отличатся от классической вероятности (устойчивость статистической вероятности).

4.Геометрическое определение вероятности события.

Если геометрическая мера всей области равна S, а геометрическая мера части этой области, попадание в которую благоприятствует данному событию A, есть S_A, то вероятность события

$$P(A) = S_A / S$$

5. Вопросы для самопроверки.

1. Образуют ли полную группу следующие события:
а) A_1 - появление герба, A_2 - появление цифры при одном бросании монеты? (Да);
б) B_1 - появление двух гербов. B - появлений двух цифр при бросании двух монет? (Нет);
в) C_1 - хотя бы одно попадание, C_2 - хотя бы один промах при двух выстрелах по мишени?
2. Являются ли несовместными следующие события:
а) B_1, - появление герба на первой монете, B_2 - появление цифры на второй монете при бросании двух монет? (Нет),
б) C_a- ни одного попадания, C_1 - одно попадание, C_2 - два попадания при двух выстрелах по мишени?

3. Являются ли равновозможными следующие события:

а) D_1 - появление двух гербов, D_2 - появление двух цифр, D_3 - появление одного герба и одной цифры при бросании двух симметричных монет? (Нет);

б) F_1 - появление не менее трех очков, F_2 - не более четырех очков? (Да).

4. Являются ли случаями следующие группы событий:

а) A_1, - появление герба, A_2 .- появление цифры при бросании симметричной монеты? (Да);

б) D_1 - попадание, D_2 - промах при одном выстреле по мишени? (Нет);

в) C_1 - появление не более двух очков, D_2 -появление трех или четырех очков, C_{2x}- появление не менее пяти очков при бросании игральной кости? (Да).

5. Привести примеры:

а) трех событий, образующих группу случаев;

б) трех событий, равновозможных и несовместных, но не образующих полной группы;

в) группу двух событий, несовместных и образующих полную группу, но не равновозможных;

г) двух событий, равновозможных и образующих полную группу, но совместных, (см. 3. 6).

6. Вычисление вероятностей событий

При непосредственном подсчете вероятностей событий необходимо знание вопросов из элементарной алгебры, связанных с понятиями комбинаторики. Введем эти понятия.

Пусть M - некоторое конечное множество, состоящее из элементов

$$(a,b,,c,..l) \qquad (1)$$

Сочетанием из *n* элементов, взятых по k , называется всякая часть множества (1), содержащая k элементов. Два различных сочетания из данных n элементов, взятых по k. , отличаются друг от друга только составом. а не порядком размещения, входящих в них элементов, т.е. если два сочетания различны, то в одном из них содержится хотя бы один элемент, не содержащийся в другом.

Число сочетаний из n элементов по k обозначается символом $C_n^k = (o \le k \le n)$и вычисляется по формуле

$$C_n^k = \frac{n!}{k!(n-k)!} = \frac{n(n-1)(n-2)...(n-k+1)}{1.2.3...k} \qquad (2)$$

Всякое упорядоченное множество (конечное) называется перестановкой, образованной из его элементов.

Число всевозможных перестановок, которые могут быть образованы из n элементов, равно

$$n! = 1.2.3...n$$

и обозначается символом,

$$P_n = n! \approx 1.2.3...n \qquad (3)$$

Размещением из *n* элементов по *k*. называется всякая упорядоченная часть множества (1), содержащая k элементов.

Два различных размещения из данных n элементов, взятых по k , различаются либо составом входящих в них элементов, либо при одном и том же составе элементов, порядком их расположения. Число различных размещений, взятых из n элементов по k, обозначается символом A_n^k и выражается формулой

$$A_n^k = n(n-1)(n-2)...(n-k+1) = \frac{n!}{(n-k)!} \qquad (4)$$

ABBYY PDF Transformer 2.0
Click here to buy
www.ABBYY.com

7. Примеры решения задач.

Пример. Одновременно бросаются три монеты. Найти вероятность того, что на двух из них выпадет "орел".

Решение. Для удобства будем предполагать, что монеты некоторым образом занумерованы. Единичным исходом здесь является совокупный результат по трем монетам (другими словами, для того, чтобы задать единичный исход, надо сказать, *что* выпало на первой монете, на второй и на третьей). Перечислим возможные исходы (см. Таблицу 1, в которой выпадение "орла" на соответствующей монете обозначено буквой "О", "решки" – "Р"). Видно, что общее число n исходов равно 8. Число m благоприятствующих исходов равно 3 – это исходы с номерами 2, 3, 5 Таблицы 1. Тогда

$$P(A) = \frac{m}{n} = \frac{3}{8}.$$

Таблица 1.

Номер исхода	Номер монеты		
	1	2	3
1	О	О	О
2	О	Р	О
3	О	О	Р
4	О	Р	Р
5	Р	О	О
6	Р	Р	О
7	Р	О	Р
8	Р	Р	Р

Пример. В коробке 6 белых шаров и 8 красных. Наудачу одновременно извлекаются 3 шара. Найти вероятность, того, что среди них будут:
а) два белых шара; б) не менее одного белого.

Решение. а) Для удобства будем предполагать, что имеющиеся шары некоторым образом перенумерованы. Пусть, например, белые шары имеют номера 1, 2, … ,6 красные – 7, 8 , … ,14. Тогда единичным исходом является произвольная тройка номеров: $\{123\}$, $\{124\}$, …, $\{12,13,14\}$. (Оставляем читателю в качестве упражнения проверку того, что данные исходы удовлетворяют всем условиям классического определения вероятностей.) Тогда общее число n исходов равно числу способов, которыми можно выбрать 3 номера из имеющихся 14-ти номеров. Напомним, что такое число равно соответствующему числу сочетаний: $n = C_{14}^{3}$.

(В общем случае,

$$C_k^s = \frac{k!}{s!(k-s)!}$$

равно числу способов, которыми можно выбрать s объектов из k *имеющихся объектов.*)
Таким образом,

$$n = C_{14}^{3} = \frac{14!}{3!\,11!} = \frac{1\cdot 2\cdot\ldots\cdot 10\cdot 11\cdot 12\cdot 13\cdot 14}{1\cdot 2\cdot 3\cdot 1\cdot 2\cdot\ldots\cdot 10\cdot 11} = 2\cdot 13\cdot 14 = 364.$$

Найдем теперь число m исходов, благоприятствующих появлению двух белых шаров среди трех извлеченных. Число способов, которыми можно выбрать 2 шара из имеющихся 6-ти белых шаров, равно C_6^2. Но число благоприятствующих исходов с фиксированной парой белых шаров равно числу способов, которыми можно выбрать оставшийся красный шар в тройку, т.е. равно C_8^1. Поэтому

$$m = C_6^2 \cdot C_8^1 = \frac{6!}{2!\cdot 4!}\cdot\frac{8!}{1!\cdot 7!} = 5\cdot 3\cdot 8 = 120.$$

Окончательно имеем

$$P(A) = \frac{m}{n} = \frac{120}{364} = \frac{30}{91},$$

где A – событие состоящее в том, что среди трех отобранных шаров ровно 2 белых шара.

б) Полное число n исходов найдено в п. а). Число троек, в которых не менее 2 белых шаров, равно сумме троек с двумя белыми шарами и троек с тремя белыми шарами:

$$m = C_6^2 \cdot C_8^1 + C_6^3 = 120 + 20 = 140.$$

Окончательно имеем

$$P(B) = \frac{m}{n} = \frac{140}{364} = \frac{35}{91},$$

где B – событие состоящее в том, что среди трех отобранных шаров не менее 2-х белых шаров.

8. Сложения и умножения событий

Сумой двух событий А и В называется событие С, состоящее в появлении хотя бы одного из событий А или В. Обозначение суммы событий : А+В, иначе А или В, иначе $A \cup B$.

Произведением двух событий А и В называется событие С, состоящее в совместном появлении события А и события В. Обозначение произведение событий : АВ, иначе А и В, иначе $A \cap B$.

Произведение нескольких событий называется событие, состоящее в совместном появлении всех этих событий.

9. Теоремы сложения вероятностей

Вероятность появление какого – либо одного из двух несовместных событий равна сумме вероятностей этих событий, т.е.

$$P(A+B) = P(A) + P(B) \qquad (5)$$

Если события А и В совместны, то вероятность появления хотя бы одного из них определяется формулой

$$P(A+B) = P(A) + P(B) - P(AB), \qquad (8)$$

Следствие I. Сумма вероятностей событий образующих полную группу, равна единице.

$$P(A_1) + P(A_2) + \ldots + P(A_k) = 1.$$

Следствие 2. Вероятность события противоположному данному, равна разности между единицей и вероятностью данного события, т.е. $P(\overline{A}) = 1 - P(A)$.

Событие $\overline{A}$ называется противоположным событию А, если оно состоит в не появлении события А.

10.Теорема умножения вероятностей для зависимых событий

Условной вероятностью события А при наличии В называется вероятность события А, вычисленная при условии, что событие В произошло. Это вероятность обозначается $P(A/B)$ иначе $P_B(A)$

Событие А и В называются независимыми, если появление одного из них не меняет вероятности появление другого.

$$\boxed{\begin{gathered} P(AB) = P(A)P_A(B), \\ P(ABC) = P(A)P_A(B)P_{AB}(C), \\ P(ABCD) = P(A)P_A(B)P_{AB}(C)P_{ABC}(D), \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.. \end{gathered}}$$

Теорема умножения вероятностей для независимых событий. *Вероятность произведения двух независимых событий равна произведению их вероятностей, т.е.*

$$\boxed{P(AB) = P(A)P(B).}$$

Аналогичное утверждение справедливо для любого числа независимых событий.

11. Примеры решения задач

Пример. Два стрелка одновременно выстреливают в мишень. Вероятность попадания для первого стрелка равна 0,6, для второго – 0,8. Найти вероятность того, что в мишени будет: а) одна пробоина; б) хотя бы одна пробоина.

Решение. а) Прежде всего, укажем, когда может наступать интересующее нас событие, перебирая все возможные варианты. В мишени будет одна пробоина тогда и только тогда, когда первый стрелок попал и второй стрелок промахнулся или первый стрелок промахнулся и второй стрелок попал. Пусть событие A – в мишени будет одна пробоина, событие B_1 – первый стрелок попал, событие B_2 – второй стрелок попал. Тогда $\overline{B}_1$ – первый стрелок промахнулся, $\overline{B}_2$ – второй стрелок промахнулся. “Тогда и только тогда, когда” соответствует отношению равенства событий. Соединительный союз “или” соответствует операции сложения событий. Соединительный союз “и” соответствует умножению событий. Тогда для наступления события A, равносильна следующему символическому равенству

$$A = B_1\overline{B}_2 + \overline{B}_1 B_2.$$

Откуда следует равенство вероятностей

$$P(A) = P(B_1\overline{B}_2 + \overline{B}_1 B_2).$$

Так как события $B_1\overline{B}_2$ и $\overline{B}_1 B_2$ несовместны, то, применяя теорему сложения вероятностей для несовместных событий, приходим к равенству

$$P(A) = P(B_1\overline{B}_2) + P(\overline{B}_1 B_2).$$

События $B_1, \overline{B}_2$ и $\overline{B}_1, B_2$ попарно независимы, поэтому, применяя теорему умножения вероятностей для независимых событий, получаем

$$P(A) = P(B_1)P(\overline{B}_2) + P(\overline{B}_1)P(B_2).$$

По условию, $P(B_1) = 0{,}6$ и $P(B_2) = 0{,}8.$ Тогда, по свойству взаимно противоположных событий (см. следствие из теоремы сложения вероятностей для несовместных событий, $k = 2$), $P(\overline{B}_1) = 1 - P(B_1) = 1 - 0{,}6 = 0{,}4$ и $P(\overline{B}_2) = 1 - P(B_2) = 1 - 0{,}8 = 0{,}2.$ Окончательно имеем

$$P(A) = 0{,}6 \cdot 0{,}2 + 0{,}4 \cdot 0{,}8 = 0{,}44.$$

б) Пусть m – число попаданий в мишень, тогда искомой является вероятность $P(m \geq 1)$ (заметим, что слова “хотя бы один”, “не менее чем один”, “по-крайней мере один” являются синонимами). Событие $(m \geq 1)$ равносильно тому, что число попаданий в мишень будет равно 1 или 2, т.е.

$$(m \geq 1) = (m = 1) + (m = 2).$$

Тогда, учитывая несовместность событий $(m = 1)$ и $(m = 2)$, получаем

$$P(m \geq 1) = P(m = 1) + P(m = 2).$$

$P(m = 1) = P(A) = 0{,}44$ (см. п. а) данного примера). Событие $(m = 2)$ (два попадания в мишень) наступает тогда и только тогда, когда первый стрелок попадет в мишень и второй стрелок попадет, т.е.

$$(m = 2) = B_1 B_2.$$

Поэтому

$$P(m=2)=P(B_1)P(B_2)=0{,}6\cdot 0{,}8=0{,}48$$

(см. теорему умножения вероятностей для независимых событий). Окончательно имеем

$$P(m\ge 1)=P(m=1)+P(m=2)=0{,}44+0{,}48=0{,}92.$$

Отметим, что эта задача допускает и другое решение. Так как события $(m\ge 1)$ и $(m=0)$ взаимно противоположны, то

$$P(m\ge 1)=1-P(m=0).$$

Но $P(m=0)=P(\overline{B}_1)P(\overline{B}_2)=0{,}4\cdot 0{,}2=0{,}08.$ Следовательно

$$P(m\ge 1)=1-P(m=0)=1-0{,}08=0{,}92.$$

Пример. В коробке лежат 4 белых шара и 6 красных. Наудачу, один за другим из коробки извлекается 2 шара. Найти вероятность того, что среди них будет:
а) один красный шар; б) менее 2-х красных шаров.

Решение. а) Пусть событие A – среди двух извлеченных шаров – ровно один красный. Это событие наступает тогда и только тогда, когда первый из извлеченных шаров – красный, а второй – белый или первый шар – белый, а второй – красный. Напомним, что соединительный союз “или” соответствует сложению событий, союзы “и”, “а” соответствуют умножению событий. Тогда описание всех возможностей наступления события A равносильно следующему формальному равенству

$$A=К_1Б_2+Б_1К_2,$$

где $К_1$ $(К_2)$ – первый (второй) шар – красный, $Б_1$ $(Б_2)$ – первый (второй) шар – белый. События $К_1Б_2$ и $Б_1К_2$ – несовместны, поэтому, используя теорему сложения вероятностей для несовместных событий, получаем

$$P(A)=P(К_1Б_2)+P(Б_1К_2).$$

Применяя теперь теорему умножения вероятностей, приходим к равенству

$$P(A)=P(К_1)P_{К_1}(Б_2)+P(Б_1)P_{Б_1}(К_2).$$

Для вычисления вероятностей из правой части последнего равенства используем классическое определение вероятности. Тогда

$$P(A)=\frac{6}{10}\cdot\frac{4}{9}+\frac{4}{10}\cdot\frac{6}{9}=\frac{8}{15}.$$

б) Пусть m – число красных шаров среди двух извлеченных. Тогда искомой является вероятность $P(m<2)$. Очевидно, что $(m<2)=(m=0)+(m=1)$, и $P(m=1)=P(A)$ (см. п. а) данного примера). Вместе с тем, событие $(m=0)$ – среди извлеченных шаров нет красных – равносильно тому, что первый шар окажется белым и второй – также белым, т.е. $(m=0)=Б_1Б_2$, поэтому

$$P(m=0)=P(Б_1Б_2)=P(Б_1)P_{Б_1}(Б_2)=\frac{4}{10}\cdot\frac{3}{9}=\frac{2}{15}.$$

Окончательно имеем

$$P(m<2)=P(m=0)+P(m=1)=\frac{2}{15}+\frac{8}{15}=\frac{10}{15}=\frac{2}{3}.$$

Заметим, что вероятность $P(m<2)$ может быть также найдена по-другому. События $(m<2)$ и $(m=2)$ взаимно противоположны, поэтому

$$P(m<2)=1-P(m=2).$$

Но

$$P(m=2)=P(К_1К_2)=P(К_1)P_{К_1}(К_2)=\frac{6}{10}\cdot\frac{5}{9}=\frac{1}{3}.$$

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

Тогда

$$P(m<2)=1-\frac{1}{3}=\frac{2}{3}.$$

Пример. На предприятии брак составляет в среднем 1,5% от общего выпуска изделий. Среди годных изделии первый сорт составляет 80%. Какова вероятность того, что наудачу взятое изделие окажется изделием первого сорта, если оно взято из общей массы изготовленной продукции'[7]

Решение: Введем обозначения: Событие A - взятое изделие первого сорта, событие B -взятое изделие годное, событие $\overline{B}$ - взятое изделие бракованное. Событие А может произойти только совместно с событием B

$$A=BA=B\cap A$$

Применяя теорему умножения, найдем вероятность события A

$$P(A)=P(BA)=P(B)P(A/B) \qquad (9)$$

Из условия задачи следует, что

$$P(\overline{B})={}^{1{,}5\%}\!/_{100\%}=0{,}015; \qquad \mathrm{P(A/B)}=\frac{80\%}{100\%}=0{,}8$$

Тогда $P(B)=1-P(\overline{B})=1-0{,}015=0{,}985$

Подставляя значения вероятностей $P(B)$ и $P(A/B)$ в формуле (9), получим

$$P(A)=0.985*0{,}8=0{,}788$$

Пример. Студент ищет нужную формулу в 3-х справочниках. Вероятность того, что формула имеется в первом справочнике равна 0,9, во втором - 0,7, в третьем - 0,8. Найти вероятности следующих событий: A - формула имеется только в одном справочнике, В - формула имеется только в двух справочниках, С - формула имеется во всех трех справочниках, D - формула имеется хотя бы в одном справочнике.

Решение: События A_1, A_2, A_3 - независимые и обозначают соответствующие справочники, т.к. появление любого из них влияет на вероятность появление двух других. Противоположные события $\overline{A}_1, \overline{A}_2, \overline{A}_3$ также независимые. Введем обозначения: событие A_i - формула имеется i-ом справочнике (i=1,2,3), $\overline{A}_i$ - противоположное событие. По условии задачи имеем

$$P(A_1)=p_1=0{,}9, \qquad P(\overline{A}_1)=1-P(A_1)=1-P_1=1-0{,}9=0{,}1; q_1=0{,}1$$

$$P(A_2)=p_2=0{,}7, \qquad P(\overline{A}_2)=1-P(A_2)=1-p_2=1-0{,}7=0{,}3; q_2=0{,}3$$

$$P(A_3)=p_3=0{,}8, \qquad P(\overline{A}_3)=1-P(A_3)=1-P_3=1-0{,}8=0{,}2; q_3=0{,}2$$

I. Событие А можно представить в виде суммы произведений событий

$$A=A_1\overline{A}_2\overline{A}_3+\overline{A}_1A_2\overline{A}_3+\overline{A}_1\overline{A}_2A_3.$$

Применяя теоремы сложения и умножения вероятностей, получим

$$P(A)=P\left(A_1\overline{A}_2\overline{A}_3\right)+P\left(\overline{A}_1A_2\overline{A}_3\right)+P(\overline{A}_1\overline{A}_2A_3)=P\left(A_1\right)P(\overline{A}_2)P\left(\overline{A}_3\right)+P\left(\overline{A}_1\right)P\left(A_2\right)P\left(\overline{A}_3\right)+P\left(\overline{A}_1\right)P\left(\overline{A}_2\right)P\left(A_3\right)=$$

$$=P\left(A_2\right)P\left(\overline{A}_2\right)P\left(\overline{A}_3\right)+P\left(\overline{A}_1\right)P\left(A_2\right)P\left(\overline{A}_3\right)+P(\overline{A}_1)P\left(\overline{A}_2\right)P\left(A_3\right)=p_1q_2q_3+q_1p_2q_3+q_1q_2p_3=$$

$$=0{,}9*0{,}3*0{,}2+0{,}1*0{,}7*0{,}2+0{,}1*0{,}3*0{,}8=0{,}092$$

Вероятность, что формула имеется только в одном справочнике равна 0,092.

2. Событие В можно предоставить в виде суммы произведений событий

$$B=A_1A_2\overline{A}_3+A_1\overline{A}_2A_3+\overline{A}_1A_2A_3$$

тогда

$$P(B)=P(A_1A_2\overline{A}_3+P(A_1\overline{A}_2A_3)+P(\overline{A}_1A_2A_3)=P(A_1A_2\overline{A}_3)+P(A_1\overline{A}_2A_3)+P(\overline{A}_1A_2A_3)=$$
$$p_1p_2q_3+p_1q_2p_3+q_1p_2p_3=0{,}9*0{,}7*0{,}2+0{,}9*0{,}3*0{,}8+0{,}1*0{,}7*0{,}8=0{,}398$$

Вероятность того, что формула имеется только в двух справочниках равна 0,398.

3. Событие С можно представить в виде произведения событий $C=A_1A_2A_3$

тогда $P(C)=P(A_1A_2A3)=P(A_1)P(A_2)P(A_3)=p_1p_2p_3=0{,}9*0{,}7*0{,}8=0{,}504$

Вероятность, того что формула имеется во всех трех справочниках равна 0,504

4. Событие D можно представить в виде суммы событий A, B, C. Событие D=A+B+C, тогда

$$P(D)=P(A+B+C)=P(A)+P(B)+P(C)=0{,}092+0{,}398+0{,}504=0{,}994$$

Вероятность события D можно найти с помощью противоположного события $\overline{D}$, $P(D)=1-P(\overline{D})$.

Событие $\overline{D}$, заключается в том, что формулы нет ни в одном справочнике $\overline{D}=\overline{A}_1\overline{A}_2\overline{A}_3$,

$$P(\overline{D})=P(\overline{A}_1\overline{A}_2\overline{A}_3)=P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3)=q_1q_2q_3P(D)=1-q_1q_2q_3=1-0{,}1*0{,}3*0{,}2=1-0{,}006=0{,}994$$

Вероятность, что формула имеется хотя бы в одном справочнике равна 0,994.

12. Вопросы для самопроверки

I. Может ли сумма двух событий A и B совпадать с их произведением?

Решение: Да, может, если из события A вытекает B и, наоборот, из B вытекает A . Приведите конкретный пример.

2. Событие B есть частный случай события A. Следует ли из B , что A произошло ?

3. Зависимы или независимы: а) несовместные события;
б) события, образующие полную группу;
в) равновозможные события?

4. Событие B является частным случаем события A , т.е.из появления события B с достоверностью вытекает появление события A . Чему равны: а) их сумма, б) их произведение?

5. Назовите противоположные события для следующих событий:
A - появление двух гербов при бросании двух монет, B - три попадания при трех выстрелах.

6. Представьте в виде сумм, произведений или сумм произведений следующие события:
A - все три попадания, B - все три промаха, C - хотя бы одно попадание, D - хотя бы один промах, E - не меньше двух попаданий, F - не больше одного попадания,
G - попадание в мишень не раньше, чем при третьем выстреле, если производится три выстрела по мишени. (Обозначьте попадание в мишень через Ac , = 1,2,3).

13. ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ И ФОРМУЛА БЕЙЕСА.

Формула полной вероятности объединяет теоремы сложения и умножения.
Если событие A может появиться с одним и только из событий (гипотез) $H_1,H_2,H_3,...,H_n$

Образующих полную группу попарно несовместных событий $\left(\sum P(H_i)=1\right)$, вероятность события А вычисляется по формуле полной вероятности:

$$P(A)=P(H_1)P(A/H_1)+P(H_2)P(A/H_2)+P(H_3)P(A/H_3)+\ldots+P(H_n)P(A/H_n)$$

или $P(A)=\sum\limits_{i=1}^{n} P(H_i)P(A/H_i)$, где $P(H_i)$ - вероятность гипотезы H_i,

$P(A/H_i)$ - условная вероятность события А при гипотезе H_i.

Формула Бейеса является следствием формулы полной вероятности.

Пусть в результате опыта появилось событие A, тогда с учетом этого условные вероятности гипотез Hi вычисляются по формуле Бейеса:

$$P(H_i/A)=\frac{P(H_i)P(A/H_i)}{\sum\limits_{i=1}^{n} P(H_i)P(A/H_i)}$$

14. Примеры решения задач.

Пример. В группе из 10 студентов, пришедших на экзамен, 3 - подготовленных отлично, 4 - хорошо, 2 - посредственно и I -плохо. В экзаменационных билетах имеется 20 вопросов. Отлично подготовленный студент может ответить на все 20 вопросов, хорошо подготовленный - на 16 вопросов, посредственно - на 10, плохо - на 5. Вызванный наугад студент ответил на три произвольно заданных вопроса. Найти вероятность того, что этот студент подготовлен: *а)* отлично, б) плохо.

Событие A - студент ответил на 3 вопроса.

Решение. Гипотезы: H_1 - студент подготовлен отлично,

H_2 - студент подготовлен хорошо,

H_3 - студент подготовлен посредственно,

H_4 - студент подготовлен плохо.

До опыта. Вероятности гипотез.

$P(H_1)=3/10=0{,}3;$ $\quad P(H_2)=4/10=0{,}4$

$P(H_3)=2/10=0{,}2;$ $\quad P(H_4)=1/10=0{,}1$

События A/H_i - студент ответил на три произвольно заданных вопроса, если он:

H_1 - подготовлен отлично,

H_2 - хорошо,

H_3 - посредственно,

H_4 - плохо.

Условно вероятности событий A/H_i равны:

$$P(A/H_1)=1; \qquad P(A/H_2)=\frac{16}{20}*\frac{15}{19}*\frac{14}{18}=0{,}491$$

$$P(A/H_3)=\frac{10}{20}*\frac{9}{19}*\frac{8}{18}=0{,}105; \qquad P(A/H_4)=\frac{5}{20}*\frac{4}{19}*\frac{3}{18}=0{,}009$$

После опыта.

Вероятность того, что студент подготовлен отлично, если он ответил на 3 вопроса .

$$P(H_i / A) = \frac{0{,}3 * 0{,}1}{0{,}3 * 01 + 0{,}4 * 0{,}491 + 0{,}2 * 0{,}105 + 0{,}1 * 0{,}009} = 0{,}58$$

Вероятность того, что студент подготовлен плохо, если он ответил на 3 вопроса
$$P(H_4 / A) = \frac{0{,}1 * 0{,}009}{0{,}518} = 0{,}002$$

Пример. Два стрелка независимо один от другого стреляют по одной мишени, делая каждый по одному выстрелу. Вероятность попадания в мишень для первого стрелка 0,8, для второго - 0,4. После стрельбы в мишени обнаружена одна пробоина. Найти вероятность того, что эта пробоина принадлежит первому стрелку.

Решение: до опыта возможны следующие гипотезы:
Событие A - в мишени обнаружена одна пробоина. Возможны гипотезы- H_1, H_2, H_3, H_4. H_1 - ни первый, ни второй стрелок не попадает, H_2 - оба стрелка попадут, H_3 - первый стрелок попадает, а второй не попадет, H_4 - первый стрелок не попадает, второй попадает.

Вероятности этих гипотез:

$P(H_1) = q_1 q_2 = 0{,}2 * 0{,}6 = 0{,}12$ $\qquad p_1 = 0{,}8; p_2 = 0{,}4$

$P(H_2) = p_1 p_2 = 0{,}8 * 0{,}4 = 0{,}32,$ $\qquad q_1 = 1 - p_1 = 0{,}2$

$P(H_3) = p_1 q_2 = 0{,}8 * 0{,}6 = 0{,}48,$ $\qquad q_2 = 1 - p_2 = 0{,}6$

$P(H_4) = q_1 p_2 = 0{,}2 * 0{,}4 = 0{,}08$

Условные вероятности наблюдаемого события А при этих гипотезах равны:

$P(A / H_1) = 0,$ $\qquad P(A / H_2) = 0,$ $\qquad P(A / H_3) = 1,$ $\qquad P(A / H_4) = 1.$

После опыта гипотезы H_1 и H_2 становятся невозможными, а вероятности гипотез H_3 и H_4 будут равны

$$P(H_3 / A) = \frac{0{,}48 * 1}{0{,}48 * 1 + 0{,}08 * 1} = \frac{6}{7}; \qquad P(H_4 / A) = \frac{0{,}08 * 1}{0{,}48 * 1 + 0{,}08 * 1} = \frac{1}{7}$$

Вероятность того, что пробоина принадлежит первому стрелку равна $6/7$.

15. Повторные события. Формула Бернули. Примеры решения задач.

Опыты называются независимыми, если вероятность того или иного исхода каждого опыта не зависит от того, какие исходы имели другие опыты.

Пример. Предполагается произвести 4 выстрела по мишени. Вероятность попадания при каждом выстреле считается известной и равной 0,7. Найти вероятность того, что число попаданий в мишень будет: а) равно 2; б) не менее 2-х; в) менее 4-х.

Решение. Введем обозначения, которые ниже будем использовать в подобных случаях. Число выстрелов по мишени обозначим через n (здесь $n = 4$), $p = 0{,}7$ – вероятность попадания в мишень при каждом выстреле, $q = 1 - p = 1 - 0{,}7 = 0{,}3$ – вероятность промаха при каждом выстреле, m – число попаданий. Требуется найти $P(m = 2)$, эту же вероятность обозначим через $P_{2,4}$. Перебирая все случаи, в которых число попаданий в мишень будет равно 2, получаем

$$P_{2,4} = ppqq + pqpq + pqqp + qppq + qpqp + qqpp =$$
$$= 6p^2q^2 = 6 \cdot 0{,}7^2 \cdot 0{,}3^2 = 0{,}2646.$$

В общем случае справедлива следующая теорема:

Теорема. *Пусть произведено n повторных независимых испытаний, в каждом из которых некоторое событие A наступает с вероятностью p. Тогда вероятность* $P_{m,n}$ *того, что в этих n испытаниях событие A наступит* m *раз, вычисляется по формуле*

$$\boxed{P_{m,n} = C_n^m p^m q^{n-m},}$$

где C_n^m *– число сочетаний из n по* m , $q = 1 - p$, где
$m = 0,1,2,\dots,n : q = 1 - p;$

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

Полученная формула носит название *формула Бернулли.*
Завершим рассмотрение нашего примера.

б) Так как $(m \ge 2) = (m = 2) + (m = 3) + (m = 4)$, то, применяя теорему сложения вероятностей для несовместных событий, получаем

$$P(m \ge 2) = P(m = 2) + P(m = 3) + P(m = 4) = P_{2,4} + P_{3,4} + P_{4,4}.$$

Первое слагаемое последней суммы найдено в п. а) данного примера. Аналогично для остальных:

$$P_{3,4} = C_4^3 p^3 q^1 = \frac{4!}{3! \cdot 1!} \cdot 0{,}7^3 \cdot 0{,}3 = 4 \cdot 0{,}7^3 \cdot 0{,}3 = 0{,}4116,$$

$$P_{4,4} = C_4^4 p^4 q^0 = \frac{4!}{4! \cdot 0!} \cdot 0{,}7^4 \cdot 0{,}3^0 = 1 \cdot 0{,}7^4 \cdot 1 = 0{,}2401.$$

Окончательно имеем

$$P(m \ge 2) = 0{,}2646 + 0{,}4116 + 0{,}2401 = 0{,}9163.$$

в) По аналогии с предыдущим пунктом задания,

$$P(m < 4) = P_{0,4} + P_{1,4} + P_{2,4} + P_{3,4},$$

т.е. решение требует, вообще говоря, четырех применений формулы Бернулли. Однако возможно и более короткое решение. Действительно, события $(m < 4)$ и $(m = 4)$ – взаимно противоположны, следовательно

$$P(m < 4) = 1 - P(m = 4).$$

Вероятность $P(m = 4) = P_{4,4}$ найдена в п. б) примера. Таким образом, получаем

$$P(m < 4) = 1 - P_{4,4} = 1 - 0{,}2401 = 0{,}7599.$$

Вероятность хотя бы одного появления события A при n независимых опытах в различных условиях равна

$$R_{1,n} = 1 - q_1 q_2 q_3 *** q_n = 1 - \prod_{i=1}^{n} q_i \qquad (10)$$

Для любых условий опыта (как одинаковых, так и различных)

$$\sum_{m=0}^{n} P_{m,n} = 1$$

Вероятность $R_{k,n}$ того, что при n опытах событие A появится не менее R раз, выражается формулой

$$R_{k,n}=\sum_{m=k}^{n} P_{m,n}, \qquad R_{k,n}=1-\sum_{m=0}^{k-1} P_{m,n}$$

Пример. Имеется 5 станций, с которыми поддерживается связь. Время от времени связь прерывается из-за атмосферных помех. В следствии удаленности станций перерыв друг от друга связи с каждой из них происходит независимо от остальных с вероятностью *p=0,2* . Найти вероятность того, что в данный момент времени будет иметься связь не менее чем с тремя станциями.

Решение: По формуле

$$R_{m,n}=\sum_{i=m}^{n} P_{i,n}$$

получим $R_{3,5}=\sum_{i=3}^{5} P_{i,5}=P_{3,5}+P_{4,5}+P_{5,5}=C_5^3p^3q^2+C_5^4p^4q+C_5^5p^5q^0=$

$$=\frac{5!}{3!2!}0{,}2^3*0{,}8^2+\frac{5!}{4!1!}0{,}2^4*0{,}8+\frac{5!}{5!0!}0{,}2^5*0{,}8^0=0{,}0512+0{,}0064+0{,}0003=0{,}0579$$

Пример. Прибор состоит из 4 узлов. Надежность (вероятность безотказной работы в течение времени *t*) для каждого узла равна 0.95. Узлы выходят из строя независимо один от другого. Найти вероятность того, что за время t

а) откажет хотя бы один узел; *б)* откажет ровно один узел; в) откажут ровно два узла; г) откажет не менее двух узлов.

Решение:

а) Найдем вероятность противоположного события $P_{0,4}$ - не откажет ни один узел

$$P_{0,4}=p^4=0{,}95^4=0{,}8145$$

Вероятность отказа хотя бы одного узла равна единице минус вероятность противоположного события

$R_{1,4}=1-P_{0,4}=1-0{,}8145=0{,}1855$

б) По формуле Бернулли $P_{m,n}=C_n^m p^m q^{n-m}$ найдем вероятность отказа ровно одного узла

$P_{1,4}=C_4^1qp^3=4*0{,}05*0{,}95^3$, где q=1-p=0,05 – вероятность отказала узлу.

в) По формуле Бернулли найдем вероятность отказа ровно двух узлов

$$P_{2,4}=C_4^2q^2p^2=\frac{4!}{2!2!}0{,}05^2*0{,}95^2=0{,}2256$$

г) Применяя противоположное событие, найдем вероятность отказа не менее двух узлов

$R_{2,4}=1-(P_{0,4}+P_{1,4})=1-(0{,}8145+0{,}1715)=0{,}0140$

16. Формула Пуассона (редких событий)

Теорема. *Пусть произведено n повторных независимых испытаний, в каждом из которых некоторое событие A наступает с вероятностью* p *, причем*
а) число испытаний достаточно велико ($n\geq 100$*) ; б)* $\lambda=np\leq 10$.

Тогда вероятность $P_{m,n}$ *того, что в этих n испытаниях событие A наступит m раз, вычисляется по следующей приближенной формуле*

$$P_{m,n} = \frac{\lambda^m}{m!} e^{-\lambda}.$$

Эта формула и называется *формулой Пуассона (редких событий).*

Пример. По каналу связи передано 1000 сигналов. Вероятность ошибки при передаче каждого из сигналов равна 0,005. Найти вероятность того, что неверно передано: а) 7 сигналов; б) не менее 4-х сигналов.

Решение. а) Воспользуемся формулой Пуассона, т.к. условия ее применимости в данном случае выполнены: число испытаний достаточно велико $(n = 1000 \ge 100)$ и $\lambda = np = 1000 \cdot 0{,}005 = 5 \le 10$. Искомое значение $P_{7,1000}$ найдем по таблице функции Пуассона при $m = 7$ и $\lambda = 5$ (см. учебник В.Е. Гмурман): $P_{7,1000} = 0{,}1045$.

б) Требуется найти $P(m \ge 4)$, где m – число неверно принятых сигналов. Так как $(m \ge 4) = (m = 4) + (m = 5) + \ldots + (m = 1000)$, то $P(m \ge 4) = P_{4,1000} + P_{5,1000} + \ldots + P_{1000,1000}$.

Искать каждое из слагаемых этой суммы и затем выполнять суммирование – такое решение не представляется рациональным из-за большого числа слагаемых и потому, что таблица функции Пуассона не дает искомых значений с требуемой в данном случае точностью. Воспользуемся переходом к противоположному событию: $P(m \ge 4) = 1 - P(m < 4) = 1 - (P_{0,1000} + P_{1,1000} + P_{2,1000} + P_{3,1000})$.

Находя вероятности из правой части последнего равенства по таблице функции Пуассона, окончательно получаем

$$P(m \ge 4) = 1 - (0{,}0067 + 0{,}0337 + 0{,}0842 + 0{,}1404) = 0{,}735.$$

17.Локальная теорема Муавра-Лапласа

Теорема. *Пусть произведено n повторных независимых испытаний, в каждом из которых некоторое событие A наступает с вероятностью* p *, причем число испытаний достаточно велико* $(n \ge 100)$*.Тогда вероятность* $P_{m,n}$ *того, что в этих n испытаниях событие A наступит* m *раз, вычисляется по следующей приближенной формуле*

$$P_{m,n} = \frac{f(x)}{\sqrt{npq}},$$

где $f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-x^2/2}$ *– функция Гаусса,* $x = \frac{m - np}{\sqrt{npq}}$, $q = 1 - p$.

Пример. Имеется партия деталей, состоящая из 1000 штук. В среднем среди деталей такого вида стандартные детали составляют 90%. Найти вероятность того, что число стандартных деталей в данной партии окажется равным 890.

Решение. Число испытаний в данном случае достаточно велико $(n = 1000 \ge 10)$, поэтому локальная теорема Муавра-Лапласа применима. Из условия следует, что вероятность быть стандартной для произвольной детали данной партии равна

$p = \frac{90}{100} = 0{,}9$, $q = 1 - p = 1 - 0{,}9 = 0{,}1$, $m = 890$. Тогда

$$x = \frac{m - np}{\sqrt{npq}} = \frac{890 - 1000 \cdot 0{,}9}{\sqrt{1000 \cdot 0{,}9 \cdot 0{,}1}} = -1{,}05.$$

По локальной теореме Муавра-Лапласа,

$$P_{890,1000} = \frac{f(-1{,}05)}{\sqrt{1000 \cdot 0{,}9 \cdot 0{,}1}}.$$

Учитывая, что функция Гаусса четная, используя таблицу этой функции (см. учебник В.Е. Гмурман), находим $f(-1{,}05) = f(1{,}05) = 0{,}2299$. Окончательно, получаем

$$P_{890,1000} = \frac{0{,}2299}{\sqrt{1000 \cdot 0{,}9 \cdot 0{,}1}} = 0{,}0242.$$

Свойства функции Гаусса.

1) Функция Гаусса четна: $f(-x) = f(x)$, поэтому ее график симметричен относительно оси O_y;

2) $f(x) > 0$ при всех x, т.е. график $y = f(x)$ расположен строго выше оси O_x;

3) $\lim\limits_{n \to +\infty} f(x) = \lim\limits_{n \to -\infty} f(x) = 0$, т.е. ось O_x является горизонтальной асимптотой графика этой функции; <u>на практике полагаем $f(x) \approx 0$ при $x > 5$</u>.

Схематично график функции Гаусса изображен на рис. 1.

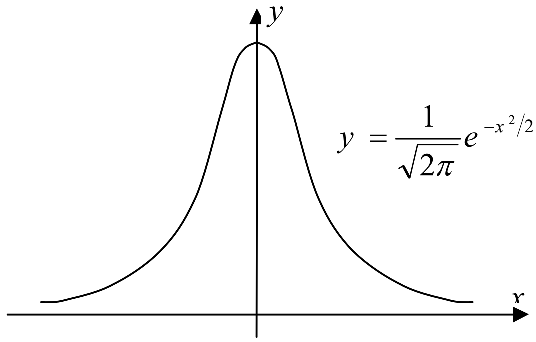


Рис.1

18.Интегральная теорема Муавра-Лапласа

Теорема. *Пусть произведено n повторных независимых испытаний, в каждом из которых некоторое событие A наступает с вероятностью p , причем число испытаний достаточно велико $(n \geq 100)$. Тогда вероятность того, что число m наступлений события A в этих n испытаниях будет заключено в границах от m_1 до m_2, вычисляется по следующей приближенной формуле*

$$P(m_1 \leq m \leq m_2) = \frac{1}{2}\left(\Phi\left(\frac{m_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{m_1 - np}{\sqrt{npq}}\right)\right),$$

где $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int\limits_0^{+\infty} e^{-x^2/2} dx$ *– функция Лапласа,* $q = 1 - p$ *. Функция Лапласа является первообразной функции Гаусса.*

Пример. Каждая из 1000 деталей партии стандартна с вероятностью 0,9. Найти вероятность того, что число стандартных деталей этой партии будет не меньше 880.

Решение. Число n повторных независимых испытаний в данном случае равно числу деталей в партии (каждая из деталей партии будет проверяться на предмет качества, а в этой проверке и состоит испытание). $n = 1000 \geq 100$, поэтому интегральная теорема Муавра-Лапласа применима; неравенство $(m \geq 880)$, где m – число стандартных деталей в партии, здесь равносильно $(880 \leq m \leq 1000)$, поэтому $m_1 = 880$, $m_2 = 1000$; $p = 0{,}9$, $q = 1 - p = 1 - 0{,}9 = 0{,}1$; $np = 1000 \cdot 0{,}9 = 900$; $npq = 1000 \cdot 0{,}9 \cdot 0{,}1 = 90$. Тогда

$$P(880 \le m \le 1000) = \frac{1}{2}\left(\Phi\left(\frac{1000-900}{\sqrt{90}}\right) - \Phi\left(\frac{880-900}{\sqrt{90}}\right)\right) =$$

$$= \frac{1}{2}\left(\Phi(10{,}5) - \Phi(-2{,}11)\right).$$

По свойствам функции Лапласа (см. ниже), $\Phi(10{,}5)=1$, $\Phi(-2{,}11) = -\Phi(2{,}11)$. По таблице функции Лапласа (см. учебник В.Е. Гмурман) находим $\Phi(2{,}11)=0{,}9651$. Тогда окончательно имеем

$$P(880 \le m \le 1000) = \frac{1}{2}(1+\Phi(2{,}11)) = \frac{1}{2}(1+0{,}9651) = 0{,}9826.$$

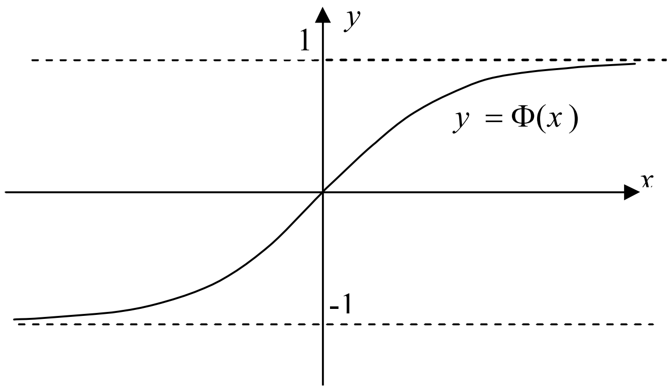


Рис. 2

Свойства функции Лапласа

1. Функция Лапласа нечетна: $\Phi(-x) = -\Phi(x)$.
2. Функция Лапласа – монотонно возрастающая;
3. $\lim\limits_{n\to+\infty}\Phi(x)=1,\ \lim\limits_{n\to-\infty}\Phi(x)=-1,$ т.е. прямые $y=1$ и $y=-1$ являются горизонтальными асимптотами (правой и левой соответственно) графика $y=\Phi(x)$; <u>на практике полагаем $\Phi(x)\approx 1$ при $x \ge 4$.</u>

График функции Лапласа схематично изображен на рис. 2.

Следствия из интегральной теоремы Муавра-Лапласа

Пусть выполнены условия применимости интегральной теоремы Муавра-Лапласа.

Следствие 1. *Вероятность того, что число m наступлений события A в n повторных независимых испытаниях будет отличаться от величины np не более чем на ε (по абсолютной величине), вычисляется по формуле*

$$P\left(|m-np| \le \varepsilon\right) = \Phi\left(\frac{\varepsilon}{\sqrt{npq}}\right).$$

Следствие 2. *Вероятность того, что доля(частость) m/n наступлений события A в n повторных независимых испытаниях будет отличаться от вероятности p наступления этого события в одном испытании не более чем на Δ (по абсолютной величине), вычисляется по формуле*

$$P\left(\left|\frac{m}{n}-p\right| \le \Delta\right) = \Phi\left(\frac{\Delta\sqrt{n}}{\sqrt{pq}}\right).$$

Пример. Подлежат исследованию 1000 проб руды. Вероятность промышленного содержания металла в каждой пробе равна 0,15. Найти границы, в которых с вероятностью 0,9973 будет заключено число проб руды с промышленным содержанием металла.

Решение. Искомые границы для числа m проб руды с промышленным содержанием металла (из данных 1000 проб) определяются величинами m_1 и m_2 (см. интегральную теорему Муавра-Лапласа). Будем предполагать, что искомые границы симметричны относительно величины np, где $n=1000$ и $p=0{,}15$. Тогда $m_1 = np-\varepsilon$,

$m_2 = np + \varepsilon$ для некоторого $\varepsilon \geq 0$, и, тем самым, единственной определяющей неизвестной данной задачи становится величина ε. Из следствия 1 и условия задачи следует, что

$$\Phi\left(\frac{\varepsilon}{\sqrt{npq}}\right) = 0{,}9973.$$

По таблице значений функции Лапласа найдем такое t, что $\Phi(t) = 0{,}9973$: $t = 3$.

Тогда $\varepsilon / \sqrt{npq} = 3$ и $\varepsilon = 3 \cdot \sqrt{npq} = 3 \cdot \sqrt{1000 \cdot 0{,}15 \cdot 0{,}85} \cong 33{,}8748 \cong 34$. Окончательно получаем искомые границы: $np - \varepsilon = 1000 \cdot 0{,}15 - 34 = 150 - 34 = 116$, $np + \varepsilon = 150 + 34 = 184$, т.е. с вероятностью 0,9973 число проб руды с промышленным содержанием металла (из данных 1000 проб) попадет в интервал $(116; 184)$.

Пример. В лесхозе приживается в среднем 80% саженцев. Сколько саженцев надо посадить, чтобы с вероятностью 0,9981 можно было утверждать, что доля прижившихся саженцев будет находиться в границах от 0,75 до 0,85.

Решение. $p = 80/100 = 0{,}8$ – вероятность прижиться для каждого из саженцев, $q = 1 - p = 1 - 0{,}8 = 0{,}2$. Пусть n – необходимое число саженцев (искомая величина данной задачи) и m – число прижившихся из них, тогда m/n – доля прижившихся саженцев. По условию,

$$P\left(0{,}75 \leq \frac{m}{n} \leq 0{,}85\right) = 0{,}9981.$$

Данные границы для доли m/n симметричны относительно величины $p = 0{,}8$, поэтому неравенство $0{,}75 \leq m/n \leq 0{,}85$ равносильно неравенству $|m/n - 0{,}8| \leq 0{,}05$.

Следовательно, вероятность 0,9981 – это та самая вероятность, которая вычисляется по следствию 2 из интегральной теоремы Муавра-Лапласа при $\Delta = 0{,}05$, $p = 0{,}8$, $q = 0{,}2$:

$$P\left(\left|\frac{m}{n} - 0{,}8\right| \leq 0{,}05\right) = \Phi\left(\frac{0{,}05\sqrt{n}}{\sqrt{0{,}8 \cdot 0{,}2}}\right) = 0{,}9981.$$

По таблице функции Лапласа найдем такое значение t, что $\Phi(t) = 0{,}9981$. Это значение: $t = 3{,}1$. Тогда

$$\frac{0{,}05\sqrt{n}}{\sqrt{0{,}8 \cdot 0{,}2}} = 3{,}1,$$

$\sqrt{n} = 3{,}1 \cdot \sqrt{0{,}8 \cdot 0{,}2} / 0{,}05$ и

$$n = \frac{3{,}1^2 \cdot 0{,}8 \cdot 0{,}2}{0{,}05^2} = 615{,}04 \cong 616.$$

Заметим, что значение n округлено до целых в большую сторону, чтобы обеспечить, как говорят, “запас по вероятности”. Кроме того, видно, что полученное значение n достаточно велико (более 100), поэтому применение интегральной теоремы Муавра-Лапласа для решения данной задачи было возможно.

II. Дискретная случайная величина

1. Закон распределения дискретной случайной величины

Определение. ***Случайной величиной*** *называется переменная, которая в результате испытания принимает то или иное числовое значение.*

Пример. Число попаданий в мишень при n выстрелах – случайная величина.

Пример. Рост наудачу взятого человека – случайная величина.

Определение. *Случайная величина называется **дискретной**, если число возможных значений конечно или счетно.*
(Напомним, что множество называется *счетным*, если его элементы можно перенумеровать натуральными числами.)
В этом смысле, число попаданий в мишень – пример дискретной случайной величины. Рост человека – непрерывная случайная величина (такие случайные величины будут рассмотрены ниже).

Случайные величины характеризуются следующими понятиями как закон распределения, числовые характеристики (математическое ожидание, дисперсия, среднеквадратическое отклонение). Для обозначения случайных величин будем использовать заглавные буквы латинского алфавита (возможно с индексами), например, $X, Y, Z, \ldots, X_1, Y_2, Z_3, \ldots$ и т.п.

Определение. ***Законом распределения*** *дискретной случайной величины называется такая таблица, в которой перечислены все возможные значения этой случайной величины (без повторений) с соответствующими им вероятностями.*

В общем виде закон распределения для случайной величины, например, X :

X :

x_i	x_1	x_2	$\cdots$	x_k
p_i	p_2	p_2	$\cdots$	p_k

где $p_i = P(X = x_i),\ i = 1, 2, \ldots, k$.

Из определения закона распределения следует, что события $(X = x_1)$, $(X = x_2)$, … , $(X = x_k)$ образуют полную систему (группу), поэтому

$$P(X = x_1) + P(X = x_2) + \ldots + P(X = x_k) = 1,$$

т.е.

$$\boxed{p_1 + p_2 + \ldots + p_k = 1.}$$

Данное равенство называется *основным свойством закона распределения.*

Пример. Два стрелка одновременно выстреливают в мишень. Вероятность попадания для первого равна 0,6, для второго – 0,8. Составить закон распределения случайной величины Z – общего числа попаданий в мишень.

Решение. Возможные значения данной случайной величины: 0, 1, 2. Обозначим через B_1 и B_2 события, состоящие в попадании в мишень первого и второго стрелков (соответственно). Тогда аналогично упомянутому примеру получаем

$$P(Z = 0) = P(\overline{B}_1)P(\overline{B}_2) = (1 - 0{,}6)\cdot(1 - 0{,}8) = 0{,}4\cdot 0{,}2 = 0{,}08,$$

$$P(Z = 1) = P(B_1)P(\overline{B}_2) + P(\overline{B}_1)P(B_2) = 0{,}6\cdot 0{,}2 + 0{,}4\cdot 0{,}8 = 0{,}44,$$

$$P(Z = 2) = P(B_1)P(B_2) = 0{,}6\cdot 0{,}8 = 0{,}48.$$

Окончательно, закон распределения случайной величины Z имеет вид:

Z :

z_i	0	1	2	Σ
p_i	0,08	0,44	0,48	1

Пример. В коробке – 3 белых шара и 2 красных. Шары извлекаются последовательно до появления белого шара. Составить закон распределения случайной величины X – числа извлеченных шаров.

Решение. Возможные значения данной случайной величины: 1, 2, 3. Событие

$(X=1)$ (из коробки будет извлечен один единственный шар) наступает тогда и только тогда, когда первый из шаров оказывается белым, т.к. появление именно белого шара является сигналом к прекращению последующих извлечений (см. условие). Поэтому

$$P(X=1)=P(Б_1)=\frac{3}{5},$$

где событие $Б_1$ – первый из извлеченных шаров – белый. Событие $(X=2)$ (из коробки будет извлечено ровно 2 шара) наступает тогда и только тогда, когда первый из извлеченных шаров оказывается красным, а второй – белым. Поэтому

$$P(X=2)=P(К_1Б_2)=P(К_1)P_{К_1}(Б_2)=\frac{2}{5}\cdot\frac{3}{4}=\frac{3}{10},$$

где событие $К_1$ – первый из извлеченных шаров – красный, $Б_2$ – второй шар – белый. Наконец событие $(X=3)$ (из коробки будет извлечено 3 шара) наступает тогда и только тогда, когда первый шар – красный, второй – красный и третий – белый. Поэтому

$$P(X=3)=P(К_1К_2Б_3)=P(К_1)P_{К_1}(К_2)P_{К_1К_2}(Б_3)=\frac{2}{5}\cdot\frac{1}{4}\cdot\frac{3}{3}=\frac{1}{10}.$$

Окончательно искомый закон распределения имеет вид:

$X:$

x_i	1	2	3	Σ
p_i	0,6	0,3	0,1	1

Пример. Стрелок стреляет в мишень 3 раза. Вероятность попадания при каждом выстреле равна 0,8. Составить закон распределения случайной величины X – числа попаданий в мишень.

Решение. Возможные значения для числа попаданий: 0, 1, 2, 3. Вероятности того, что случайная величина X примет эти значения вычисляются по формуле Бернулли при $n=3,\ p=0{,}8,\ q=0{,}2$:

$$P(X=0)=P_{0,3}=C_3^0p^0q^3=1\cdot1\cdot0{,}2^3=0{,}008,$$

$$P(X=1)=P_{1,3}=C_3^1p^1q^2=3\cdot0{,}8\cdot0{,}2^2=0{,}096,$$

$$P(X=2)=P_{2,3}=C_3^2p^2q^1=3\cdot0{,}8^2\cdot0{,}2^1=0{,}384,$$

$$P(X=3)=P_{3,3}=C_3^3p^3q^0=1\cdot0{,}8^3\cdot0{,}2^0=0{,}512.$$

Окончательно искомый закон распределения имеет вид:

$X:$

x_i	0	1	2	3	Σ
p_i	0,008	0,096	0,384	0,512	1

Полученный закон распределения является частным случаем так называемого *биномиального закона распределения* (при $n=3,\ p=0{,}8$).

Определение. *Случайная величина X имеет* ***биномиальный закон распределения*** *с параметрами n и p, если ее закон распределения имеет вид* :

$X:$

x_i	0	1	2	…	n
p_i	$P_{0,n}$	$P_{1,n}$	$P_{2,n}$	…	$P_{n,n}$

,

где вероятности $P_{m,n}$ вычисляются по формуле Бернулли:

$$P_{m,n}=C_n^mp^mq^{n-m},$$

n – *положительное целое число,* $m = 0,1,2,\ldots,n$, $0 < p < 1$.

В пределе при $n \to \infty$ и $\lambda = np = const$ биномиальное распределение переходит в так называемое распределение Пуассона.

Определение. *Говорят, что случайная величина X имеет* ***распределение Пуассона*** *с параметром* λ*, если ее закон распределения имеет вид:*

X:

x_i	0	1	2	…
p_i	P_0	P_1	P_2	...

,

где

$$P_m = \frac{\lambda^m}{m!}e^{-\lambda},$$

$m = 0,1,2,\ldots$, λ – *положительное число.*

2.Арифметические операции над случайными величинами

Определение. *Случайные величины X и Y называются* ***равными****, если их законы распределения точно совпадают, и для произвольного числа* α *справедливо равенство:* $(X = \alpha) = (Y = \alpha)$.

Пример. Пусть законы распределения случайных величин *X и Y* имеют вид:

X:

x_i	0	1
p_i	0,5	0,5

Y:

y_i	0	1
p_i	0,5	0,5

.

Эти случайные величины равны, если дополнительно справедливы равенства $(X = 0) = (Y = 0)$ и $(X = 1) = (Y = 1)$, т.е. случайная величина X принимает значение 0 тогда и только тогда, когда случайная величина Y принимает значение 0, и аналогично со значением 1.

Произвольная случайная величина допускает умножение на число. Действительно, пусть закон распределения случайной величины X имеет вид:

X :

x_i	x_1	x_2	…	x_k
p_i	p_2	p_2	…	p_k

и α – некоторое число.

Определение. *Случайной величиной* $Y = \alpha \cdot X$ *называется такая случайная величина, закон распределения которой имеет вид :*

Y :

y_i	$\alpha \cdot x_1$	$\alpha \cdot x_2$	…	$\alpha \cdot x_k$
p_i	p_2	p_2	…	p_k

Пример. Пусть закон распределения случайной величины X имеет вид:

X :

x_i	0	1	2
p_i	0,16	0,48	0,36

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

и $\alpha = 5$, $Y = \alpha \cdot X$. Тогда закон распределения Y:

Y :

y_i	0	5	10
p_i	0,16	0,48	0,36

Определение. *Случайные величины X и Y называются* ***независимыми****, если для любых i и j события $(X = x_i)$ и $(Y = y_j)$ – независимы.*

Пример. Пусть из коробки, в которой – 6 белых и 8 красных шаров, извлекается 1 шар. Рассмотрим случайные величины X – число белых шаров, Y – число красных шаров из извлеченных. События, например, $(X = 1)$ и $(Y = 1)$ – несовместны, а поэтому – независимы. Следовательно, и случайные величины X и Y независимы.

Определение. ***Суммой (разностью, произведением)*** *случайных величин X и Y называется такая случайная величина $Z = X + Y$ ($Z = X - Y$, $Z = X \cdot Y$), которая принимает значение z_k в некотором испытании, если значения x_i и y_j случайных величин X и Y в этом испытании таковы, что $z_k = x_i + y_j$ ($z_k = x_i - y_j$, $z_k = x_i \cdot y_j$).*

Пример. Пусть заданы законы распределения независимых случайных величин X и Y:

X:

x_i	0	1
p_i	0,4	0,6

Y:

y_j	0	1
p_j	0,2	0,8

Составить закон распределения случайной величины $U = X - Y$.

Решение. Удобно использовать вспомогательную таблицу вида:

y_j \ x_i	0	1
0	0	1
1	–1	0

в каждой из центральных клеток, которой записаны соответствующие произведения случайных величин X и Y. Такая таблица показывает, <u>какие</u> значения принимает случайная величина U и <u>когда</u> она принимает эти значения. Так $U = 0$ тогда и только тогда, когда $X = 0$ и $Y = 0$ или $X = 1$ и $Y = 1$. Поэтому

$$P(U = 0) = P((X = 0)(Y = 0) + (X = 1)(Y = 1)).$$

Применяя теорему сложения вероятностей для несовместных событий, теорему умножения вероятностей – для независимых событий (по условию, случайные величины X и Y – независимы), получаем

$$P(U = 0) = P(X = 0)P(Y = 0) + P(X = 1)P(Y = 1) = 0{,}4 \cdot 0{,}2 + 0{,}6 \cdot 0{,}8 = 0{,}56.$$

Для наступления каждого из двух оставшихся значений случайной величины U (-1 и 1) имеется по одной возможности. Например, $U = 1$ тогда и только тогда, когда $X = 1$ и $Y = 0$. Тогда получаем:

$$P(U = 1) = P((X = 1)(Y = 0)) = P(X = 1)P(Y = 0) = 0{,}6 \cdot 0{,}2 = 0{,}12.$$

Аналогично,

$$P(U = -1) = P((X = 0)(Y = 1)) = P(X = 0)P(Y = 1) = 0{,}4 \cdot 0{,}8 = 0{,}32.$$

Окончательно, закон распределения случайной величины U имеет вид:

U:	u_i	-1	0	1
	p_i	0,32	0,56	0,12

3.Параметры распределения дискретной случайной величины

Пусть закон распределения дискретной случайной величины X имеет вид

X:	x_i	x_1	x_2	$\dots$	x_k
	p_i	p_2	p_2	$\dots$	p_k

Определение. ***Математическим ожиданием*** *дискретной случайной величины X называется число M(X), вычисляемое по формуле*

$$\boxed{M(X)=\sum_{i=1}^{k}x_i p_i = x_1p_1+x_2p_2+\ldots+x_kp_k\,.}$$

Математическое ожидание случайной величины есть число около которого группируются значения этой случайной величины.

Механическим аналогом математического ожидания дискретной случайной величины является центр масс (центр тяжести) системы точечных масс: если в точках числовой оси с абсциссами $x_1,x_2,\ldots,x_k$ расположены точечные массы $p_1,p_2,\ldots,p_k$, то абсцисса их центра масс находится точно по формуле для $M(X)$, приведенной выше.

Пример. Пусть случайная величина X биномиально распределена с параметрами $n=3$ и $p=0{,}8$:

X:	x_i	0	1	2	3
	p_i	0,008	0,096	0,384	0,512

Тогда

$$M(X)=0\cdot 0{,}008+1\cdot 0{,}096+2\cdot 0{,}384+3\cdot 0{,}512=2{,}4.$$

Свойства математического ожидания

1. *Математическое ожидание постоянной случайной величины равно самой постоянной, т.е.*

$$M(C)=C,$$

где C – некоторое число.

Постоянной случайной величиной C называется такая случайная величина, которая принимает единственное значение равное C с вероятностью p=1. M(C)=$C\bullet 1=C$.

2. *Постоянный множитель можно выносить за знак математического ожидания, т.е.*

$$M(\alpha X)=\alpha M(X),$$

где α – произвольное число.

3. *Математическое ожидание суммы (разности) случайных величин равно сумме (разности) математических ожиданий этих случайных величин, т.е.*

$$M(X\pm Y)=M(X)\pm M(Y).$$

4. *Математическое ожидание произведения независимых случайных величин равно*

произведению их математических ожиданий, т.е.

$$M(XY) = M(X)M(Y).$$

5. *Пусть* $X_1, X_2, \dots, X_n$ *– такие случайные величины, математические ожидания которых равны между собой, т.е.* $M(X_i) = a$, *где* $i = 1, 2, \dots, n$, *и* a *– некоторое число. Тогда среднее арифметическое этих случайных величин равно их общему математическому ожиданию, т.е.*

$$M\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = a.$$

Заметим, что свойства 2 – 5 математического ожидания остаются справедливыми также для непрерывных случайных величин.

Определение. ***Дисперсией*** *дискретной случайной величины* X *называется число* $D(X)$, *определяемое равенством*

$$\boxed{\begin{aligned} D(X) = \sum_{i=1}^{k}(x_i - M(X))^2 p_i &= (x_1 - M(X))^2 p_1 + (x_2 - M(X))^2 p_2 + \dots + \\ &+ (x_k - M(X))^2 p_k. \end{aligned}}$$

Число $D(X)$ является мерой разброса значений случайной величины X около ее математического ожидания.

Пример. Пусть случайная величина X биномиально распределена с параметрами $n = 3$ и $p = 0{,}8$. Найдем дисперсию этой случайной величины.
В предыдущем примере найдено, что $M(X) = 2{,}4$. Тогда
$D(X) = (0 - 2{,}4)^2 \cdot 0{,}008 + (1 - 2{,}4)^2 \cdot 0{,}096 + (2 - 2{,}4)^2 \cdot 0{,}384 + (3 - 2{,}4)^2 \cdot 0{,}512 = 0{,}48.$

Свойства дисперсии

1. *Дисперсия постоянной случайной величины равна нулю, т.е.*

$$D(C) = 0.$$

2. *Постоянный множитель можно выносить за знак дисперсии, возводя его при этом в квадрат, т.е.*

$$D(\alpha X) = \alpha^2 D(X),$$

где α *– произвольное число.*

3. *Справедливо равенство:*

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

4. *Дисперсия суммы (разности) двух независимых случайных величин равна сумме дисперсий этих случайных величин, т.е.*

$$D(X \pm Y) = D(X) + D(Y),$$

где случайные величины X *и* Y *– независимы.*

5. *Пусть случайные величины* $X_1, X_2, \dots, X_n$ *– независимы и* $D(X_i) = \sigma^2$, *где* $i = 1, 2, \dots, n$. *Тогда*

$$D\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right) = \frac{\sigma^2}{n}.$$

Замечание. $\sqrt{D(X)}$ называется *средним квадратическим отклонением* случайной

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

величины X и обычно обозначается через σ.

Отметим также, что свойство 3 дисперсии более удобно для ее вычисления по сравнению с исходным определением дисперсии.

Пример. Пусть закон распределения случайной величины X имеет вид

X:

x_i	1	2
p_i	0,6	0,4

Найти $D(X)$, используя свойство 3 дисперсии.

Решение.

$$M(X) = 1\cdot 0{,}6 + 2\cdot 0{,}4 = 1{,}4, \qquad M(X^2) = 1^2\cdot 0{,}6 + 2^2\cdot 0{,}4 = 2{,}2,$$
$$D(X) = M(X^2) - M^2(X) = 2{,}2 - 1{,}4^2 = 0{,}24.$$

Математическое ожидание и дисперсия случайной величины называются *параметрами распределения* этой случайной величины.

Теорема. *Пусть случайная величина* $X \equiv m$ *– биномиально распределена с параметрами* n *и* p, *тогда параметры ее распределения могут быть найдены по формулам:*

$$M(m) = np, \qquad D(m) = npq.$$

Также справедливы равенства

$$M\left(\frac{m}{n}\right) = p, \qquad D\left(\frac{m}{n}\right) = \frac{pq}{n}.$$

Пример. Пусть случайная величина X биномиально распределена с параметрам $n = 3$ и $p = 0{,}8$. Тогда

$$M(X) = np = 3\cdot 0{,}8 = 2{,}4, \qquad D(m) = npq = 3\cdot 0{,}8\cdot 0{,}2 = 0{,}48.$$

Очевидно, что использование формул последней теоремы упрощает и ускоряет вычисление математического ожидания и дисперсии биномиально распределенной случайной величины по сравнению с применением исходных определений для *M(X)* и $D(X)$.

4. Функция распределения дискретной случайной величины

Определение. ***Функцией распределения*** *случайной величины X называется такая функция* $F(x)$, *значение которой в точке x численно равно вероятности того, что в произвольном испытании значение случайной величины X окажется меньше чем x, т.е.*

$$\boxed{F(x) = P(X < x).}$$

Данное определение задает функцию распределения не только для дискретных, но и для непрерывных случайных величин.

Пример. Пусть закон распределения случайной величины X имеет вид

X:

x_i	1	2
p_i	0,3	0,7

Найти функцию распределения этой случайной величины.

Решение. Найдем сначала *F(x)* для некоторых значений переменной *x*. Например,

$$F(0) = P(X < 0) = P(\varnothing) = 0,$$

так как данная случайная величина не имеет значений меньших нуля, а потому событие ($X < 0$) для нее является невозможным. Аналогично, при любом значении переменной x, которое менее или равно 1, будем иметь $F(x) = 0$. Далее имеем:

$$F(1{,}5) = P(X < 1{,}5) = P(X = 1) = 0{,}3.$$

Аналогично, при любом значении переменной x таком, что $1 < x \le 2$, будем иметь $F(x) = 0{,}3$.

$$F(2{,}5) = P(X < 2{,}5) = P(X = 1) + P(X = 2) = 0{,}3 + 0{,}7 = 1.$$

(Или, другими словами, так как все значения данной случайной величины менее 2,5, то событие ($X < 2{,}5$) является достоверным, а потому его вероятность равна 1.) Аналогично, при любом значении переменной x, которое более или равно 2, будем иметь $F(x) = 1$.

Окончательно имеем:

$$F(x) = \begin{cases} 0 & \text{при } x \le 1; \\ 0{,}3 & \text{при } 1 < x \le 2; \\ 1 & \text{при } x > 2. \end{cases}$$

График найденной функции распределения изображен на рис. 1.

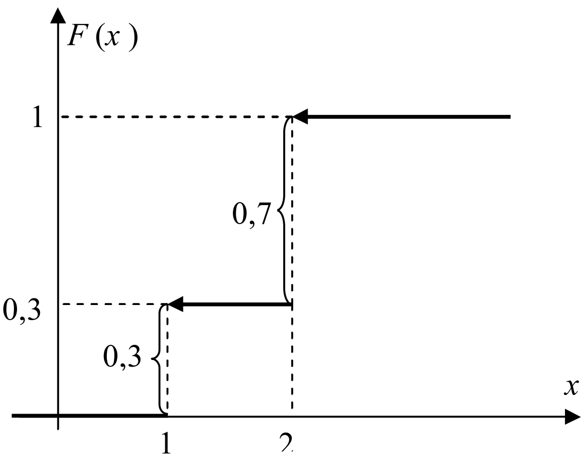


Рис. 1

Свойства функции распределения

1. Функция распределения является неубывающей функцией.
2. Область значений: $0 \le F(x) \le 1$.
3. Асимптотические свойства: $\lim\limits_{x \to -\infty} F(x) = 0$, $\lim\limits_{x \to +\infty} F(x) = 1$ (другими словами, прямые y =0 и y =1 являются асимптотами (левой и правой соответственно) графика $y = F(x)$).
4. Вероятность того, что в произвольном испытании значение случайной величины X будет принадлежать полуинтервалу $[\alpha, \beta)$, где α и β – произвольные числа, вычисляется по формуле

$$P(\alpha \le X < \beta) = F(\beta) - F(\alpha).$$

Доказательство. Значение функции распределения равна вероятности соответствующего события, но область значений вероятности есть отрезок $[0,1]$ – тем самым доказано свойство 2.

Используя определение функции распределения, получаем $\lim\limits_{x \to -\infty} F(x) = \lim\limits_{x \to -\infty} P(X < x) = P(X < -\infty)$. Но произвольное значение случайной величины принадлежит числовой прямой, поэтому событие $(X < -\infty)$ является невозможным. Вероятность невозможного события равна нулю поэтому $\lim\limits_{x \to -\infty} F(x) = 0$.

Аналогично, учитывая, что событие $(X < +\infty)$ является достоверным, а вероятность такого события равна 1, получаем $\lim\limits_{x \to +\infty} F(x) = 1$.

Нетрудно видеть, что

$$(X < \beta) = (X < \alpha) + (\alpha \le X < \beta),$$

причем события правой части этого равенства несовместны. Принимая во внимание определение функции распределения и теорему сложении вероятностей для несовместных событий, получаем

$$F(\beta) = P(X < \beta) = P(X < \alpha) + P(\alpha \le X < \beta) = F(\alpha) + P(\alpha \le X < \beta),$$

что равносильно свойству 4 т.е. $P(\alpha \le X < \beta) = F(\beta) - F(\alpha)$.

Доказательство свойства 1 мы оставляем читателю в качестве упражнения (указание: используйте рассуждении от противного и свойство 4).

III. Непрерывная случайная величина

1. Плотность распределения непрерывной случайной величины

Неформально говоря, случайная величина непрерывна, если ее значения полностью заполняют некоторый интервал. Более точно, справедливо

Определение. *Случайная величина называется* ***непрерывной****, если ее функция распределения непрерывна на всей числовой прямой и дифференцируема при всех x за исключением, быть может, отдельных значений.*

Определение. ***Плотностью распределения*** *непрерывной случайной величины X называется такая функция $\varphi = \varphi(x)$, что вероятность того, что в произвольном испытании значение случайной величины X окажется принадлежащим некоторому отрезку $[\alpha, \beta]$, вычисляется по формуле*

$$P(\alpha \le X \le \beta) = \int_{\alpha}^{\beta} \varphi(x)\,dx.$$

Принимая во внимание геометрический смысл определенного интеграла, получаем

Геометрический смысл плотности распределения. *Вероятность того, что в произвольном испытании значение случайной величины X окажется принадлежащим некоторому отрезку $[\alpha, \beta]$, численно равна площади $S(\alpha, \beta)$ под кривой плотности распределения на данном отрезке* (см. рис.1).

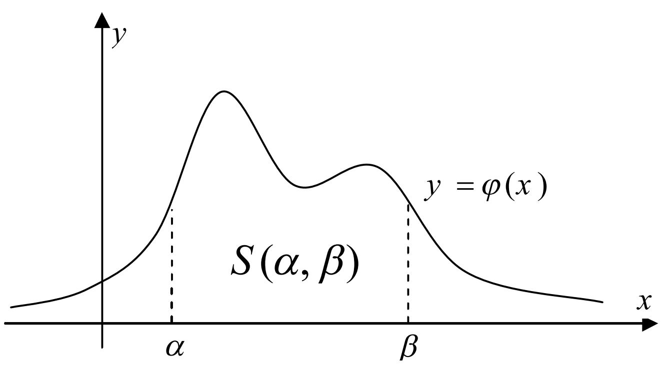


Рис. 1

Пример. Пусть плотность распределения случайной величины X имеет вид:

$$\varphi(x) = \begin{cases} 1/2 & \text{при } x \in [-1, 1]; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Найти вероятности:
а) $P(-2 \le X \le -0{,}4)$; б) $P(X \le -3)$;
в) $P(X \ge -2)$.

Решение. А) По определению плотности распределения,

$$P(-2 \le X \le -0{,}4) = \int_{-2}^{-0{,}4} \varphi(x)\,dx.$$

Вместе с тем, данная плотность распределения задана аналитически по-разному на промежутках $[-2, -1)$ и $[-1; -0{,}4]$ отрезка интегрирования. Соответственно, используя свойства определенного интеграла, получаем

$$P(-2 \le X \le -0{,}4) = \int_{-2}^{-0{,}4} \varphi(x)\,dx = \int_{-2}^{-1} 0\,dx + \int_{-1}^{-0{,}4} 1/2\,dx = \frac{1}{2}x\Big|_{-1}^{-0{,}4} = \frac{1}{2}(-0{,}4 - (-1)) = 0{,}3.$$

По геометрическому смыслу плотности распределения, полученная вероятность численно равна площади под кривой плотности распределения (см. рис. 2) на отрезке

$[-2;-0{,}4]$, т.е. равна площади фигуры, составленной из отрезка длины 1 прямоугольника со сторонами $1/2$ и 0,6.

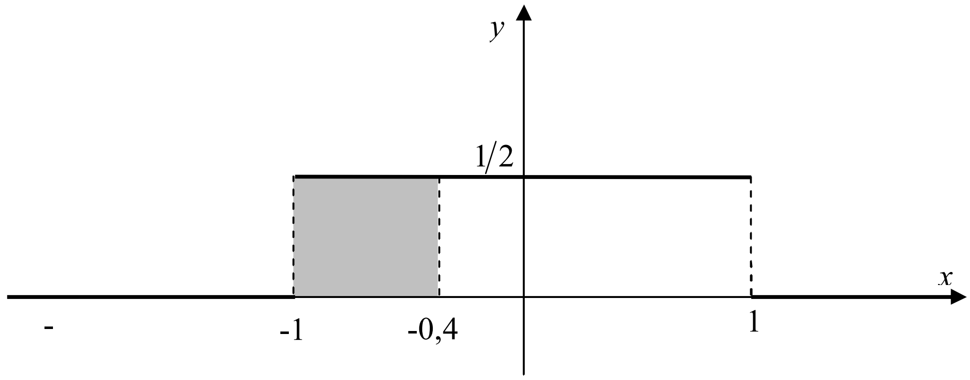


Рис. 2

Б) Неравенство $(X \le -3)$ равносильно тому, что $(-\infty < X \le -3)$. Учитывая, что на промежутке $(-\infty;-3)$ данная плотность распределения равна 0, получаем

$$P(X \le -3) = P(-\infty < X \le -3) = \int_{-\infty}^{-3} 0\,dx = 0.$$

в) Аналогично предыдущим пунктам задачи, имеем

$$P(-2 \le X < +\infty) = \int_{-2}^{+\infty} \varphi(x)dx = \int_{-2}^{-1} 0\,dx + \int_{-1}^{1} 1/2\,dx + \int_{1}^{+\infty} 0\,dx = \frac{1}{2}x\Big|_{-1}^{1} = \frac{1}{2}(1-(-1)) = 1.$$

Рассмотрение геометрического смысла результатов последних двух пунктов данного примера мы оставляем читателю в качестве упражнения. ▶

2. Свойства плотности распределения

1. Плотность распределения неотрицательна, т.е. $\varphi(x) \ge 0$ при всех *x*.
2. Интеграл от плотности распределения на всей числовой прямой равен 1, т.е.

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

(Данное свойство называется *условием нормировки* плотности распределения.)

Доказательство. Предположим противное: пусть найдется такой отрезок $[\alpha,\beta]$, что плотность распределения $\varphi(x)$ отрицательна на этом отрезке. Тогда (см. свойства определенного интеграла) имеем

$$\int_{\alpha}^{\beta} \varphi(x)dx < 0.$$

Но, по определению плотности распределения, интеграл, стоящий в левой части последнего неравенства равен $P(\alpha \le X \le \beta)$. Так как вероятность события не может быть отрицательной, приходим к противоречию, что доказывает справедливость свойства 1.

По определению плотности распределения,

$$\int_{-\infty}^{+\infty} \varphi(x)dx = P(-\infty \le X \le +\infty).$$

Но событие $(-\infty < X < +\infty)$ является достоверным, поэтому его вероятность равна 1. Тем самым доказано свойство 2.

3.Функция распределения непрерывной случайной величины

Пусть X – непрерывная случайная величина и $\varphi = \varphi(x) -$ ее плотность распределения. Используя определения функции распределения и плотности распределения, получаем

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

Обратно, если задана функция распределения непрерывной случайной величины, то (см. теорему об интеграле с переменным верхним пределом) плотность распределения этой случайной величины будет определяться равенством

$$\varphi(x) = F'(x).$$

Таким образом, имеется два равноправных способа задания непрерывной случайной величины: с помощью или плотности распределения, или функции распределения.

Пример. Пусть плотность распределения непрерывной случайной величины X имеет вид:

$$\varphi(x) = \begin{cases} \frac{1}{2} & \text{при } x \in [0;2], \\ 0 & \text{в остальных случаях.} \end{cases}$$

Найти функцию распределения.

Решение. Пусть $x < 0$. Тогда

$$F(x) = \int_{-\infty}^{x} \varphi(x) dx = \int_{-\infty}^{x} 0 dx = 0.$$

Если $x \in [0;2]$, то

$$F(x) = \int_{-\infty}^{x} \varphi(x) dx = \int_{-\infty}^{0} 0 dx + \int_{0}^{x} \frac{1}{2} dx = 0 + \frac{1}{2}x \Big|_{0}^{x} = \frac{1}{2}x .$$

Если $x > 2$, то

$$F(x) = \int_{-\infty}^{x} \varphi(x) dx = \int_{-\infty}^{0} 0 dx + \int_{0}^{2} \frac{1}{2} dx + \int_{2}^{x} 0 dx = 0 + \frac{1}{2}x \Big|_{0}^{2} + 0 = \frac{1}{2}(2-0) = 1 .$$

Таким образом, окончательно, искомая функция распределения имеет вид

$$F(x) = \begin{cases} 0 & \text{при } x < 0, \\ \frac{1}{2}x & \text{при } x \in [0;2], \\ 1 & \text{при } x > 2 \end{cases}$$

(см. рис. 3).

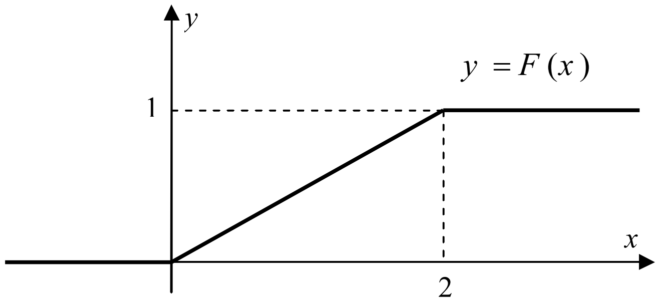


Рис. 3

4.Математическое ожидание и дисперсия непрерывной случайной величины

Формулы для вычисления математического ожидания и дисперсии непрерывной случайной величины аналогичны соответствующим формулам для дискретной случайной величины. Действительно, рассмотрим следующую таблицу.

	Дискретная случайная величина	Непрерывная случайная величина
Способ описания	Закон распределения	Плотность распределения
$M(X)$	$M(X)=\sum_{i=1}^{k} x_i p_i$	$M(X)=\int_{-\infty}^{+\infty} x\varphi(x)dx$
$D(X)$	$D(X)=\sum_{i=1}^{k}(x_i - M(X))^2 p_i$	$D(X)=\int_{-\infty}^{+\infty}(x-M(X))^2\varphi(x)dx$

Таким образом, переходя при записи этих формул от дискретной к непрерывной случайной величине, суммирование заменяется интегрированием по всей числовой оси, а вместо вероятности p_i используется плотность распределения $\varphi(x)$.

Пример. Функция распределения непрерывной случайной величины X имеет вид:

$$F(x)=\begin{cases} 0 \quad \text{при } x<0, \\ x^2/9 \quad \text{при } x\in[0,3], \\ 1 \quad \text{при } x>3. \end{cases}$$

Найти математическое ожидание и дисперсию этой случайной величины.

Решение. Для нахождения $M(X)$ и $D(X)$ нам потребуется плотность распределения данной случайной величины (см. приведенные выше формулы). Получаем:

$$\varphi(x)=F'(x)=\begin{cases} 0'=0 \quad \text{при } x<0, \\ \left(x^2/9\right)'=\dfrac{2}{9}x \quad \text{при } x\in[0,3], \\ 1'=0 \quad \text{при } x>3, \end{cases}$$

или

$$\varphi(x)=\begin{cases}\frac{2}{9}x & \text{при } x\in[0,3],\\ 0 & \text{в остальных случаях.}\end{cases}$$

Тогда имеем

$$M(X)=\int_{-\infty}^{+\infty}x\varphi(x)dx=\int_{-\infty}^{0}x\cdot 0dx+\int_{0}^{3}x\cdot\frac{2}{9}x\,dx+\int_{3}^{+\infty}x\cdot 0dx=\frac{2}{9}\cdot\frac{x^3}{3}\bigg|_0^3=2.$$

Геометрически, полученное значение математического ожидания есть абсцисса центра тяжести фигуры под графиком плотности распределения, т.е. абсцисса прямоугольного треугольника OAB (см. рис. 4; напомним, что центр тяжести треугольника есть точка пересечения медиан этого треугольника, а медианы в точке пересечения делятся в отношении 2:1, считая от вершины).

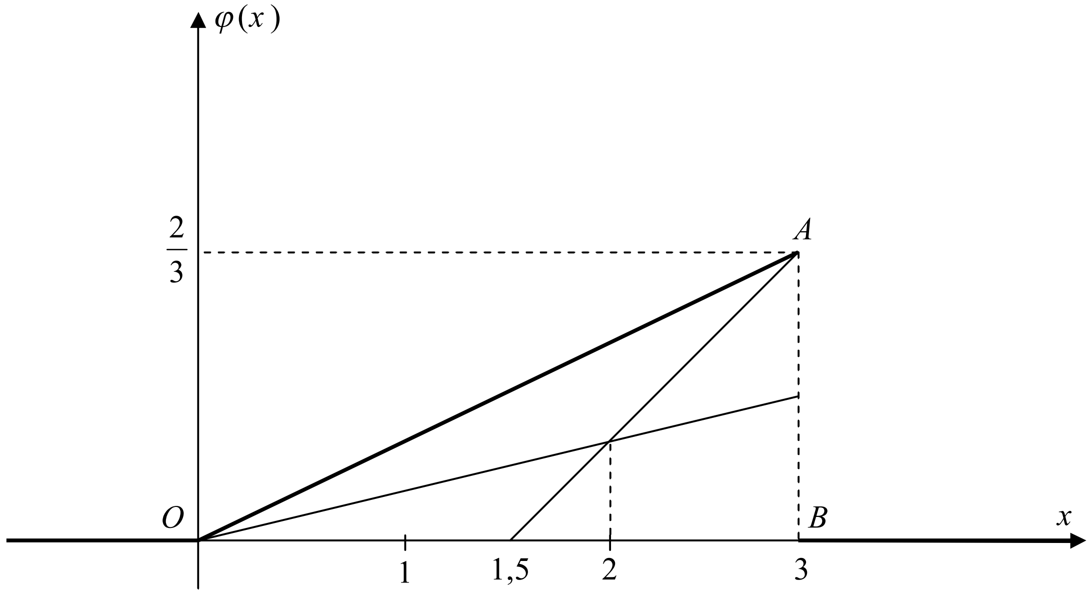


Рис. 4

Завершая решение, найдем дисперсию рассматриваемой случайной величины.

$$M(X^2)=\int_{-\infty}^{+\infty}x^2\varphi(x)dx=\int_{-\infty}^{0}x^2\cdot 0dx+\int_{0}^{3}x^2\cdot\frac{2}{9}x\,dx+\int_{3}^{+\infty}x^2\cdot 0dx=\frac{2}{9}\cdot\frac{x^4}{4}\bigg|_0^3=4{,}5,$$

$$D(X)=M(X^2)-M^2(X)=2^2-4{,}5=0{,}5.$$

5. Нормальный закон распределения

Определение. *Непрерывная случайная величина имеет* ***нормальный закон распределения*** *с параметрами a и σ, если ее плотность распределения имеет вид*

$$\varphi_N(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-(x-a)^2/2\sigma^2}.$$

Параметры a и σ нормального закона тесно связаны с параметрами распределения рассматриваемой случайной величины. Справедлива следующая теорема.

Теорема. *Пусть случайная величина X имеет нормальный закон распределения с параметрами a и σ. Тогда*

$$M(X)=a, \quad D(X)=\sigma^2.$$

Отметим, что график $\varphi_N(x)$ – результат деформации Гауссовой кривой $y=f(x)$. Рассмотрим, как изменяется этот график при изменении параметров a и σ нормального закона.

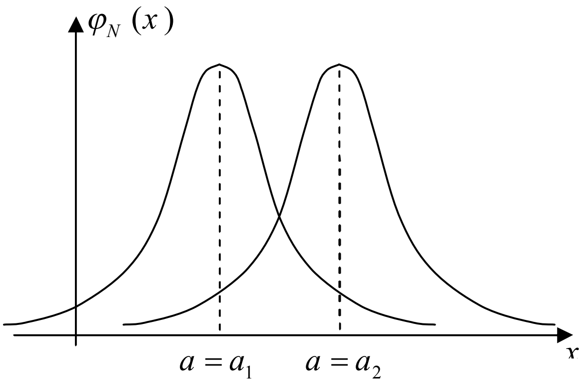


Рис.5

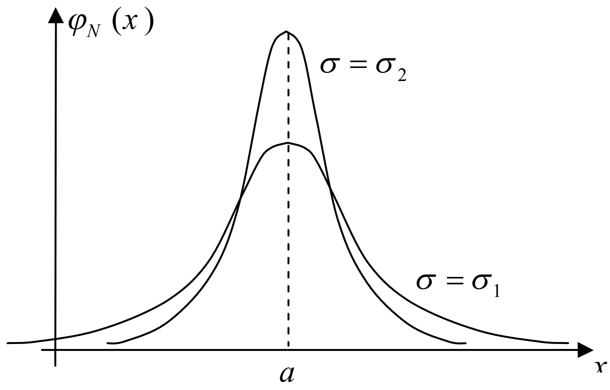


Рис. 6

На рис. 5 изображены графики $\varphi_N(x)$ при одинаковом значении параметра σ: изменение параметра a нормального закона приводит к параллельному переносу графика плотности распределения вдоль оси абсцисс.

На рис. 6 изображены графики $\varphi_N(x)$ при одинаковом значении параметра a: изменение параметра σ нормального закона приводит к “растяжению” графика вдоль оси ординат при сохранении площади под кривой равной 1 (заметим, что на рис. 6 $\sigma_2<\sigma_1$).

Теорема. *Пусть случайная величина X имеет нормальный закон распределения с параметрами a и σ. Тогда справедливы формулы:*

$$P(\alpha \le X \le \beta)=\frac{1}{2}\left(\Phi\left(\frac{\beta-a}{\sigma}\right)-\Phi\left(\frac{\alpha-a}{\sigma}\right)\right), \tag{1}$$

$$P(|X-a|\le\varepsilon)=\Phi\left(\frac{\varepsilon}{\sigma}\right), \tag{2}$$

$$F(x)=\frac{1}{2}+\frac{1}{2}\Phi\left(\frac{x-a}{\sigma}\right),$$

где $\Phi(x)$ – *функция Лапласа,* $F(x)$ – *функция распределения случайной величины X.*

Заметим, что график функции распределения $F(x)$ нормально распределенной случайной величины получается в результате деформации из графика функции Лапласа $\Phi(x)$ (см. рис. 7).

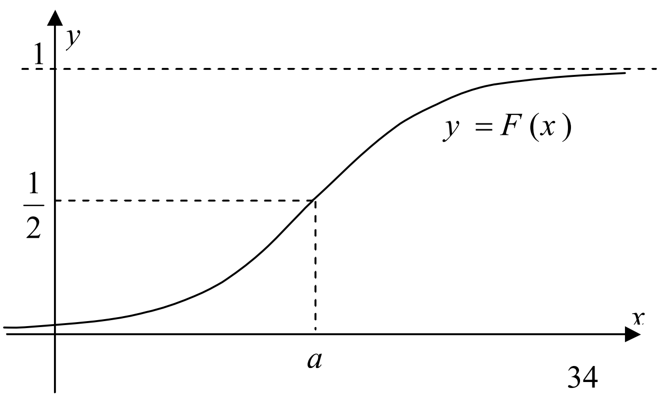


Рис. 7

Пример. Случайная величина X – ошибка измерительного прибора распределена по нормальному закону с дисперсией равной 16 мк^2. Систематическая ошибка отсутствует. Найти вероятность того, что при одном измерении

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

ошибка:
а) превзойдет по модулю 6 мк;
б) окажется в промежутке от 0,5 до 3,5 мк.

Решение. А) Отсутствие систематической ошибки означает, что значения случайной величины X группируются около нуля, поэтому $a = M(X) = 0$. Искомой является вероятность $P(|X| > 6)$. Воспользуемся переходом к противоположному событию: $P(|X| > 6) = 1 - P(|X| \le 6)$. Так как $a = 0$, то $P(|X| \le 6) = P(|X - a| \le 6)$, т.е. последняя вероятность точно того вида, что может быть вычислена по формуле (2). Используя формулу (2) при $\varepsilon = 6$, $\sigma = 4$, получаем

$$P(|X| \le 6) = \Phi(6/4) = \Phi(1{,}5) = 0{,}8664.$$

Окончательно имеем

$$P(|X| > 6) = 1 - P(|X| \le 6) = 1 - 0{,}8664 = 0{,}1336.$$

б) Искомая вероятность вычисляется по формуле (1) при $\alpha = 0{,}5,\ \beta = 3{,}5,\ a = 0,\ \sigma = 4$:

$$P(0{,}5 \le X \le 3{,}5) = \frac{1}{2}\left(\Phi\left(\frac{3{,}5-0}{4}\right) - \Phi\left(\frac{0{,}5-0}{4}\right)\right) = \frac{1}{2}\left(\Phi(0{,}875) - \Phi(0{,}125)\right) =$$

$$= \frac{1}{2}(0{,}6211 - 0{,}1034) = 0{,}2589.$$

Пример. Известна плотность распределения (дифференциальная функция) непрерывной случайной величины

$f(x) = ae^{\lambda|x|}$, $(\lambda > 0)$

Найти: а) Функцию распределения (интегральную функцию); б) коэффициент “a”
в) числовые характеристики $m_x, D_{x,}, G_{x.}$

Решение: а) Воспользуемся соотношением

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

Функция f(x) для интервала $(-\infty, 0)$ имеет вид

$$f(x) = ae^{\lambda x},\ F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{x} ae^{\lambda x}dx = \frac{a}{\lambda}\int_{-\infty}^{x} e^{\lambda x}d(\lambda x) = \frac{a}{\lambda}e^{\lambda x}\Big/_{-\infty}^{x} = \frac{a}{\lambda}e^{\lambda x},\ F(x) = \frac{a}{\lambda}e^{-\lambda x}, (x < 0).$$

Функция f(x) для интервала $(0, +\infty)$ имеет вид $f(x) = ae^{-\lambda x}$

$$F(x) = \int_{-\infty}^{0} f(x)dx + \int_{0}^{x} f(x)dx = \int_{-\infty}^{0} ae^{\lambda x}dx + \int_{0}^{x} ae^{-\lambda x}dx = \frac{a}{\lambda}e^{\lambda x}\Big/_{-\infty}^{0} - \frac{a}{\lambda}e^{-\lambda x}\Big/_{0}^{x} = (\frac{a}{\lambda} - 0) - (\frac{a}{\lambda}e^{-\lambda x} - \frac{a}{\lambda}) = \frac{2a}{\lambda} - \frac{a}{\lambda}e^{-\lambda x}$$

Итак,

$$F(x) = \begin{cases} \dfrac{a}{\lambda}e^{-\lambda x}, x < 0 \\ \dfrac{2a}{\lambda} - \dfrac{a}{\lambda}e^{-\lambda x}, x > 0 \end{cases}$$

б) Для определения параметра “a” используем свойство

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

$$\int_{-\infty}^{\infty} ae^{-\lambda|x|}dx = 1, a = \frac{1}{\int_{-\infty}^{\infty} e^{-\lambda|x|}dx} = \frac{1}{\int_{-\infty}^{0} e^{\lambda x}dx + \int_{0}^{+\infty} e^{-\lambda x}dx} = \frac{1}{\frac{1}{\lambda}e^{\lambda x}\Big/_{-\infty}^{0} + (-\frac{1}{\lambda}e^{-\lambda x})\Big/_{0}^{+\infty}} = \frac{1}{\frac{1}{\lambda} - 0 + 0 + \frac{1}{\lambda}} = \frac{\lambda}{2}$$

Итак, $a = \lambda/2, f(x) = \lambda/2\, e^{-\lambda|x|}$

Б) **Найдем числовые характеристики**.

Математическое ожидание

$$M(x) = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} x\frac{\lambda}{2}e^{-\lambda x}dx = \frac{\lambda}{2}\int_{-\infty}^{\infty} xe^{-\lambda|x|}dx = 0$$

Интеграл от нечетной функции по симметричному промежутку интегрирование равен нулю.

Дисперсия

$$D(x) = M(x^2) - M^2(x) = \int_{-\infty}^{\infty} x^2 f(x)dx - 0^2 = \int_{-\infty}^{\infty} x^2\frac{\lambda}{2}e^{-\lambda|x|}dx = \frac{\lambda}{2}\int_{-\infty}^{\infty} x^2 e^{-\lambda|x|}dx = \lambda\int_{0}^{\infty} x^2 e^{-\lambda|x|}dx$$

Интегрируя по частям, получим $D(x) = 2/\lambda^2$.

Среднее квадратическое отклонение $\sigma(x) = \sqrt{D(x)} = \sqrt{2}/\lambda$.

Пример. На перекрестке стоит автоматический светофор, в котором 1 минуту горит зеленый свет и 0,5 минут- красный, затем опять 1 минуту горит зеленый свет, 0,5 минут- красный и т.д. Некто подъезжая к перекрестку на машине в случайный момент, не связанный с работой светофора.

а/ Найти вероятность того, что он проедет перекрёсток, не останавливаясь.

Решение: Момент проезда авто машины через перекрёсток распределён равномерно в интервале, равном периоду смены цветов в светофоре.

Этот период равен 1+0,5=1,5(мин)

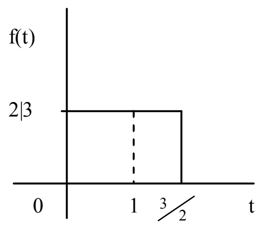


Рис.8

Для того, чтобы машина проехала через перекрёсток, не останавливаясь, достаточно, чтобы момент проезда перекрестка применяя интервал времени (0;1).

Запишем плотность распределения равномерно распределенной случайной величины.

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} .....при....x \in (\alpha, \beta) \\ 0.........при...x \notin (\alpha, \beta) \end{cases}$$

Для нашего случая

$$f(x) = \begin{cases} \frac{1}{1.5 - 0} = \frac{2}{3}, ...x \in (0, \frac{3}{2}) \\ 0..,..............x \notin (0, \frac{3}{2}) \end{cases}$$

Вероятность попадания случайной величины в заданный интервал (a,b) находим по формуле

$$P(a < x < b) = \int_a^b f(x)dx$$

Величина того, что случайная величина (момент проезда через перекрёсток) попадёт в интервал (0,1) будет равна

$$P(0 < x < 1) = \int_0^1 2\!/\!_3\, dx = \left[2\!/\!_3\, x\right]_0^1 = \frac{2}{3}$$

6.Центральная предельная теорема. Теоремы Муавра-Лапласа

Центральная предельная теорема. *Пусть случайные величины* $X_1, X_2, \ldots, X_n$ – *независимы и одинаково распределены. Тогда закон распределения их суммы* $X = X_1 + X_2 + \ldots + X_n$ *неограниченно приближается к нормальному при неограниченном увеличении числа n эти x случайных величин.*

Отметим, что центральная предельная теорема является частным случаем более общего утверждения – теоремы Ляпунова (подробнее см. учебник Е.С. Вентцель).

Следствие. *Биномиальный закон распределения неограниченно приближается к нормальному при неограниченном увеличении параметра n этого закона.*

Доказательство. Пусть случайная величина X – биномиально распределена с параметрами n и p . Рассмотрим сначала тот конкретный пример, когда X – число наступлений некоторого события A в n повторных независимых испытаниях, в каждом из которых это событие наступает с вероятностью p. Введем в рассмотрение случайные величины $X_1, X_2, \ldots, X_n$ такие, что X_i – число наступлений события A в i –□ И испытании, где $i = 1, 2, \ldots, n$. Случайная величина X_i принимает значение 1, если в i – □ И испытании событие A наступило и значение 0 – в противном случае. Сумма случайных величин X_i принимает значение m тогда и только тогда, когда число X наступлений события A в n испытаниях равно m., т.е.

$$X = X_1 + X_2 + \ldots + X_n .$$

Тогда по центральной предельной теореме для случайной величины X получаем требуемое утверждение. Аналогично данное Следствие доказывается и в общем случае.

Геометрически приближение биномиального распределения к нормальному означает, что с ростом n точки плоскости с координатами $(m, P_{m,n})$ неограниченно приближаются к кривой $\varphi_N(x)$ плотности нормального закона (здесь m – неотрицательное целое, не превосходящее n, значение $P_{m,n}$ вычисляется по формуле Бернулли; см. рис. 9).

Тогда справедливо приближенное равенство

$$P_{m,n} \approx \varphi_N(m),$$

где $a = np$, $\sigma = \sqrt{npq}$,

которое, записанное явно, и есть локальная теорема Муавра-Лапласа.

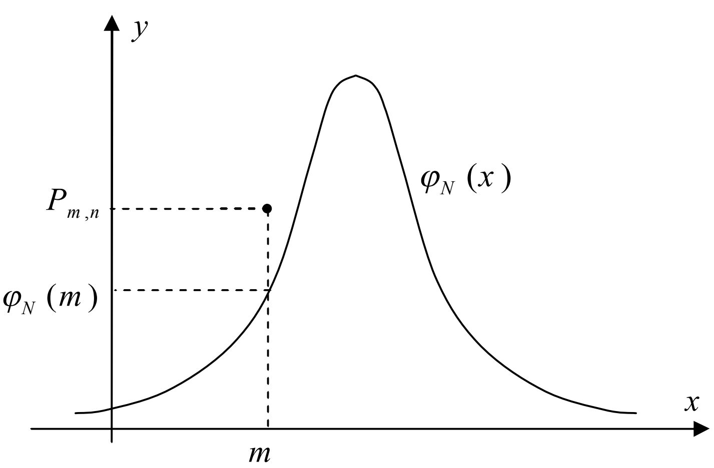


Рис.9

IV. Двумерные случайные величины

1.Совместные распределения и их параметры

Определение. *Вектор $Z=(X,Y)$, компоненты X и Y которого являются случайными величинами, называется* ***случайным вектором*** *или* ***двумерной случайной величиной.***

Пример. Пусть X – рост человека, Y – вес человека. Тогда $Z=(X,Y)$ – (непрерывная) двумерная случайная величина.

Пример. Пусть X и Y – числа попаданий в мишень первого и второго стрелков (соответственно). Тогда $Z=(X,Y)$ – (дискретная) двумерная случайная величина.

Сравнивая между собой одномерную и двумерную случайные величины, заметим, что, если результат измерения первой – точка на прямой, то результат измерения второй – точка плоскости.

Определение. *Закон распределения одной из переменных при фиксированном значении другой называется* ***условным распределением.***

Определение. *Связь между переменными называется* ***статистической****, если каждому значению одной переменной ставится в соответствие условное распределение другой переменной.*

Отметим, что задание двумерной случайной величины равносильно заданию статистической связи между переменными.

Рассмотрим сначала двумерную дискретную случайную величину.

По аналогии с одномерным случаем, закон распределения двумерной дискретной случайной величины задается с помощью таблицы вида:

$x_i \backslash y_j$	y_1	…	y_j	…	y_m
x_1	p_{11}	…	p_{1j}	…	p_{1m}
…	…	…	…	…	…
x_i	p_{i1}	…	p_{ij}	…	p_{im}
…	…	…	…	…	…
x_n	p_{n1}	…	p_{nj}	…	p_{nm}

где

$$p_{ij}=P((X=x_i)\cdot(Y=y_j)),$$

$1\le i\le n,\ 1\le j\le m.$

По аналогии с основным свойством закона распределения одномерной случайной величины, справедливо равенство

$$\sum_{i=1}^{n}\sum_{j=1}^{m}p_{ij}=1.$$

Приведенная таблица называется *совместным законом распределения* случайных величин *X* и *Y.*

Пример #. Совместный закон распределения случайных величин X и Y имеет вид:

$x_i \backslash y_j$	0	1
1	0,1	0,2
2	0,3	0,4

Найти математическое ожидание случайной величины *X.*

Решение. Прежде всего найдем закон распределения случайной величины *X.* Так как

$$P(X=1)=P((X=1)\cdot(Y=0)+(X=1)\cdot(Y=1))=0,1+0,2=0,3,$$
$$P(X=2)=P((X=2)\cdot(Y=0)+(X=2)\cdot(Y=1))=0,3+0,4=0,7,$$

то закон распределения X имеет вид:

X:

x_i	1	2
p_i	0,3	0,7

Тогда

$$M(X)=1\cdot 0,3+2\cdot 0,7=1,7.$$

Оставляем читателю в качестве упражнения проверку того, что закон распределения случайной величины Y имеет вид:

Y:

y_j	0	1
p_j	0,6	0,4

и $M(Y)=0{,}6$.

Определение. *Связь между переменными называется* ***функциональной****, если каждому значению из области определения одной переменной поставлено в соответствие однозначно определенное значение другой переменной.*

Примерами такого вида связи изобилует курс математического анализа: $y=2x+3$, $y=\sqrt{1-x}$, $y=\ln x$, $y=2^x$ и т.д. и т.д.

Определение. *Функциональная связь между значениями одной переменной и условными математическими ожиданиями другой переменной называется* ***корреляционной****.*

Определение. *График корреляционной зависимости называется* ***линией регрессии****.*

Корреляционные зависимости бывают двух видов (y по x и x по y) в зависимости от того, которая из переменных выполняет роль аргумента: x или y. Соответственно, $(x_i, M_{x_i}(Y))$ – точки корреляционной зависимости y по x и $(M_{y_j}(X), y_j)$ – точки корреляционной зависимости x по y.

Пример. По совместному закону распределения из предыдущего примера (Пример #) найти корреляционную зависимость y по x.

Решение. Применяя теорему умножения вероятностей, получаем

$$P_{X=1}(Y=0)=\frac{P((X=1)\cdot(Y=0))}{P(X=1)}=\frac{0{,}1}{0{,}3}=\frac{1}{3},$$

$$P_{X=1}(Y=1)=\frac{P((X=1)\cdot(Y=1))}{P(X=1)}=\frac{0{,}2}{0{,}3}=\frac{2}{3},$$

где вероятности, стоящие в числителях последних дробей, берутся из таблицы совместного закона распределения Примера #, вероятность $P(X=1)$ найдена в том же примере. Таким образом, условное распределение случайной величины Y при $X=1$ имеет вид:

$Y_{X=1}$:

y_j	0	1
p_j	1/3	2/3

По этому закону распределения находим условное математическое ожидание:

$$M_{X=1}(Y)=0\cdot 1/3+1\cdot 2/3=2/3.$$

Аналогично получаем:

$Y_{X=2}$:

y_j	0	1
p_j	3/7	4/7

$$M_{X=2}(Y)=0\cdot 3/7+1\cdot 4/7=4/7.$$

Собирая вместе полученные результаты, запишем корреляционную зависимость y по x в виде следующей таблицы:

x_i	1	2
$M_{x_i}(Y)$	2/3	4/7

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

Рассмотрим теперь непрерывную двумерную случайную величину.

Определение. *Функция* $\varphi(x,y)$ *называется* ***плотностью распределения*** *непрерывной двумерной случайной величины* $Z=(X,Y)$, *если для произвольных чисел* $\alpha,\beta,\gamma,\delta$ $(\alpha<\beta,\ \gamma<\delta)$ *вероятность того, что в произвольном испытании значение случайной величины* Z *попадает в прямоугольник* $\alpha\le X\le\beta,\ \gamma\le Y\le\delta$ *вычисляется по формуле*

$$P((\alpha\le X\le\beta)\cdot(\gamma\le Y\le\delta))=\int_\alpha^\beta\int_\gamma^\delta\varphi(x,y)dy\,dx.$$

Условные плотности распределения определяются формулами:

$$\varphi_x(y)=\frac{\varphi(x,y)}{\int_{-\infty}^{+\infty}\varphi(x,y)dy},\qquad \varphi_y(x)=\frac{\varphi(x,y)}{\int_{-\infty}^{+\infty}\varphi(x,y)dx}.$$

Соответственно, условные математические ожидания тогда вычисляются по формулам:

$$M_x(Y)=\int_{-\infty}^{+\infty}y\varphi_x(y)dy,\qquad M_y(X)=\int_{-\infty}^{+\infty}x\varphi_y(x)dx.$$

2. Коэффициент корреляции и его свойства

Определение. ***Коэффициентом корреляции*** ρ *случайных величин* X *и* Y *называется число, определяемое равенством*

$$\rho=\frac{M(XY)-M(X)M(Y)}{\sigma_X\sigma_Y},$$

где $\sigma_X=\sqrt{D(X)}$, $\sigma_Y=\sqrt{D(Y)}$.

Коэффициент корреляции является мерой тесноты линейной связи между переменными. Величина $M(XY)-M(X)M(Y)$ называется *ковариацией* и обозначается K_{XY}.

Замечание. Из свойства математического ожидания следует, что, *если случайные величины* X *и* Y *независимы, то коэффициент корреляции* ρ *равен нулю.* Существенно, что обратное утверждение неверно, т.е. в общем случае *из условия равенства коэффициента корреляции нулю не следует, что данные случайные величины независимы.*

Теорема (Область возможных значений коэффициента корреляции). *Модуль коэффициента корреляции не превосходит 1, т.е.*

$$\boxed{-1\le\rho\le1.}$$

Теорема. *Если модуль коэффициента корреляции двух случайных величин равен 1, то между этими случайными величинами существует линейная функциональная зависимость.*

Пример. Пусть совместный закон распределения случайных величин X и Y имеет вид:

x_i \ y_j	1	2
0	0,4	0
1	0	0,6

Тогда $Y=X+1$. Оставляем читателю в качестве упражнения проверку того, что в данном случае $\rho=1$.

Из определения ковариации следует, что

$$M(XY)=M(X)M(Y)+K_{XY}.$$

Другими словами, ковариация является мерой неравенства между математическим ожиданием произведения двух случайных величин и произведением их математических ожиданий. Аналогично, применительно к дисперсии, справедливо равенство

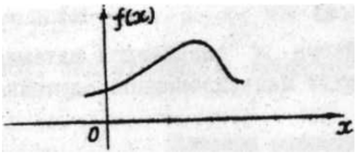
$$D(X+Y)=D(X)+D(Y)+2K_{XY}.$$

3. Вопросы для самопроверки:

1. Может ли при каком-либо значении аргумента быть:
1)функция распределения больше единицы? 2)Плотность распределения больше единицы? 3)Функция распределения отрицательной? 4)Плотность распределения отрицательной?

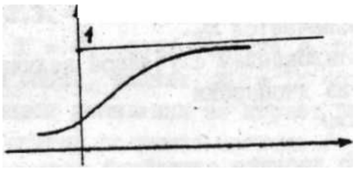
2. Какова размерность:
1)Функции распределения? 2)Плотности распределения?

2. Дан график плотности распределения *f(x)* случайной величины



Как изменится этот график, если: а) прибавить к случайной величине I; б) вычесть из случайной величины 2; в) умножить случайную величину на 2; г) изменить знак величины на обратный?

3. Дан график функции распределения *f(x)* случайной величины X



Как изменится этот график, если: а) прибавить к случайной величине I; б) умножить случайную величину на 2; в) вычесть из случайной величины 2; г) изменять знак случайной величины на обратный?

V. Закон больших чисел

1. Неравенство Чебышёва

Лемма Чебышёва. *Пусть среди значений случайной величины* Z *нет отрицательных. Тогда вероятность того, что в некотором испытании значение этой случайной величины превысит число* A *, оценивается по формуле*

$$P(Z>A)\le\frac{M(Z)}{A}.$$

Здесь M(Z)- математическое ожидание. Так как события $(Z>A)$ и $(Z\le A)$ взаимно противоположны, то $P(Z>A)=1-P(Z\le A)$, и лемма Чебышёва может быть также представлена в виде

$$P(Z \le A) \ge 1 - \frac{M(Z)}{A}.$$

Пример. В среднем в течение часа на вокзал прибывает 400 пассажиров. Оценить:
а) вероятность того, что число пассажиров, прибывших на вокзал в течение часа, будет более 420;
б) верхнюю границу для числа прибывших пассажиров, которую можно гарантировать с вероятностью не меньшей 0,9.

Решение. Пусть Z – число пассажиров, прибывающих на вокзал в течение наудачу выбранного часа. По условию, значения этой случайной величины группируются около 400. Тем самым, имеем $M(Z) = 400$. Полагая в неравенстве Чебышёва $A = 420$, получаем

$$P(Z > 420) \le \frac{M(Z)}{420} = \frac{400}{420} = 0{,}9524.$$

Из условия и второй формы записи неравенства Чебышёва следует, что

$$P(Z \le A) \ge 1 - \frac{M(Z)}{A} = 0{,}9,$$

где A – искомая верхняя граница для числа пассажиров. Таким образом, имеем равенство

$$1 - \frac{400}{A} = 0{,}9.$$

Решая это уравнение относительно A, получаем: $A = 4000$.

Неравенство Чебышёва. *Для произвольной случайной величины Y вероятность того, что в некотором испытании значение этой случайной величины будет отличаться от математического ожидания $M(Y)$ не более чем на ε (по абсолютной величине), оценивается по формуле*

$$P(|Y - M(Y)| \le \varepsilon) \ge 1 - \frac{D(Y)}{\varepsilon^2},$$

где ε – произвольное положительное число, $D(Y)$ – дисперсия.

Рассмотрим следствия из неравенства Чебышёва.

Следствие 1. *Пусть случайные величины $X_1, X_2, \ldots, X_n$ – независимы, $M(X_i) = a_i$, $D(X_i) \le C$, где $i = 1, 2, \ldots, n$, C – некоторое число. Тогда вероятность того, что среднее арифметическое этих случайных величин отличается от среднего арифметического их математических ожиданий не более чем на ε (по абсолютной величине), оценивается по формуле*

$$P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{a_1 + a_2 + \ldots + a_n}{n}\right| \le \varepsilon\right) \ge 1 - \frac{C}{n\varepsilon^2}.$$

Следствие 2. *Пусть случайные величины $X_1, X_2, \ldots, X_n$ – независимы, $M(X_i) = a$, $D(X_i) = \sigma^2$, где $i = 1, 2, \ldots, n$. Тогда вероятность того, что среднее арифметическое этих случайных величин отличается от их общего математического ожидания не более чем на ε (по абсолютной величине), оценивается по формуле*

$$P\left(\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - a\right| \le \varepsilon\right) \ge 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Следствие 3. *Пусть $X \equiv m$ – число наступлений некоторого события A в n повторных независимых испытаниях, в каждом из которых это событие наступает с вероятностью p. Тогда вероятность того, что число m наступлений события A отличается от np не более чем на ε (по абсолютной величине), оценивается по формуле*

$$P(|m - np| \le \varepsilon) \ge 1 - \frac{npq}{\varepsilon^2}.$$

Следствие 4. *Пусть $X \equiv m$ – число наступлений некоторого события A в n*

повторных независимых испытаниях, в каждом из которых это событие наступает с вероятностью p *. Тогда вероятность того, что частость* m/n *наступлений события* A *отличается от вероятности* p *не более чем на* ε *(по абсолютной величине), оценивается по формуле*

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

Последнее следствие называется также *неравенством Бернулли.*

Пример. Вероятность сделать покупку для каждого из покупателей магазина равна 0,7. Почему нельзя применить неравенство Чебышёва для оценки вероятности того, что из 1000 покупателей доля таких, которые приобретут в магазине товар, будет заключена в границах от 0,67 до 0,72? Как следует изменить левую границу, чтобы применение неравенства Чебышёва стало возможным? Решить задачу при соответствующем изменении левой границы. Найти эту же вероятность по интегральной теореме Муавра-Лапласа. Объяснить различие в полученных результатах. Сколько покупателей надо обследовать, чтобы те же границы для рассматриваемой доли можно было гарантировать с вероятностью не меньшей 0,9?

Решение. Неравенство Чебышёва позволяет оценивать вероятности попадания значения случайной величины только в границы, которые симметричны относительно математического ожидания этой случайной величины. Но в данном случае интервал (0,67; 0,72) несимметричен относительно $M(m/n)=p=0,7$, где m/n – доля покупателей, которые приобретут в магазине товар, p – вероятность приобретения товара. Соответственно, для того, чтобы применение неравенства Чебышёва стало возможным, левая граница интервала должна отстоять от $p=0,7$ ровно настолько, насколько отстоит правая, т.е. на $\varepsilon=0,72-0,7=0,02$. Неравенства $0,68\le m/n\le 0,72$ и $|m/n-0,7|\le 0,02$ – равносильны, а вероятность $P(|m/n-0,7|\le 0,02)$ оценивается по следствию 4 (неравенству Бернулли) при $p=0,7$, $\varepsilon=0,02$, $n=1000$, $q=1-p=1-0,7=0,3$:

$$P\left(\left|\frac{m}{n}-0,7\right|\le 0,02\right)\ge 1-\frac{0,7\cdot 0,3}{1000\cdot 0,02^2}=0,475.$$

Точно такая же вероятность может быть найдена по следствию 2 из интегральной теореме Муавра-Лапласа:

$$P\left(\left|\frac{m}{n}-0,7\right|\le 0,02\right)\approx\Phi\left(\frac{0,02\sqrt{1000}}{\sqrt{0,7\cdot 0,3}}\right)\approx\Phi(1,38)\approx 0,8324.$$

Очевидно, что полученные результаты не противоречат друг другу. Поясним, почему для одной и той же вероятности неравенство Чебышёва дает лишь оценку, в то время как теорема Муавра-Лапласа – точное значение. Дело в том, что неравенство Чебышёва получено без каких бы то ни было предположений о законе распределения рассматриваемой случайной величины. В результате область его применений широка, но получение точных результатов с его помощью оказывается невозможным. В свою очередь, теорема Муавра-Лапласа опирается на свойство биномиального распределения: по центральной предельной теореме, это распределение неограниченно приближается к нормальному при неограниченном увеличении числа испытаний. Использование закона распределения рассматриваемой случайной величины и позволяет уточнить окончательный результат.

Перейдем теперь к последнему заданию данной задачи. По условию и неравенству Бернулли, имеем

$$P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\ge 1-\frac{pq}{n\varepsilon^2}=0,9,$$

причем $p=0,7,\ q=0,3,\ \varepsilon=0,02$. Тогда полученное равенство

$$1-\frac{0,7\cdot 0,3}{n\cdot 0,02^2}=0,9$$

содержит единственную неизвестную: n. Решая это уравнение относительно этой неизвестной, получаем:

$$n=\frac{0,7\cdot 0,3}{0,02^2\cdot 0,1}=5250.$$

2.Теоремы Бернулли и Чебышёва

Теорема Бернулли. *Пусть* m/n *– частость наступления события* A *в* n *повторных независимых испытаниях, в каждом из которых это событие наступает с вероятностью* p *.Тогда для произвольного* $\varepsilon>0$ *вероятность того, что частость* m/n *будет отличаться от вероятности* p *не более чем на* ε *(по абсолютной величине) неограниченно приближается к 1 при неограниченном увеличении значения* n *, т.е.*

$$\lim_{n\to\infty} P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)=1.$$

Другими словами, теорема Бернулли утверждает, что *частость* m/n *наступления некоторого события сходится по вероятности к вероятности* p *наступления этого события.*

Доказательство. Учитывая, что вероятность произвольного события не превосходит 1, из неравенства Бернулли следует

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

Переходя к пределу при $n\to\infty$, получаем

$$\lim_{n\to\infty}\left(1-\frac{pq}{n\varepsilon^2}\right)\le\lim_{n\to\infty} P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\le\lim_{n\to\infty}1.$$

Крайние левый и правый пределы этого двойного неравенства равны 1. Таким образом, имеем

$$1\le\lim_{n\to\infty} P\left(\left|\frac{m}{n}-p\right|\le\varepsilon\right)\le 1,$$

что равносильно утверждению теоремы Бернулли.

Теорема Бернулли утверждает, что, если за значение вероятности p *некоторого события взять значение частости* m/n *наступления этого события, найденную по результатам* n *испытаний, то вероятность погрешности (даже сколь угодно малой) приближенного равенства* $p\approx m/n$ *будет стремиться к нулю с увеличением числа испытаний* n *.*

Теорема Чебышёва. *Пусть случайные величины* $X_1, X_2,\dots,X_n$ *независимы, одинаково распределены и* $a=M(X_1)=\ \dots\ M(X_n)$*. Тогда для произвольного* $\varepsilon>0$ *вероятность того, что среднее арифметическое этих случайных величин отличается от их общего математического ожидания не более чем на* ε *(по абсолютной величине) , неограниченно приближается к 1 при неограниченном увеличении числа* n *этих случайных величин т.е.*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1+X_2+\ldots+X_n}{n}-a\right|\le\varepsilon\right)=1.$$

Другими словами, теорема Чебышёва утверждает, что *среднее арифметическое некоторого числа случайных величин, имеющих одинаковое математическое ожидание, сходится по вероятности к их общему математическому ожиданию.*

Говоря о приложениях теоремы Чебышёва, отметим, в первую очередь, следующую возможность. *Если за значение некоторого неизвестного параметра a взять среднее арифметическое результатов $X_1, X_2, \ldots, X_n$ независимых измерений этого параметра, то вероятность погрешности (даже сколь угодно малой) приближенного равенства* $a \approx \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ *будет стремиться к нулю при неограниченном увеличении числа n этих измерений.*

Теоремы Бернулли и Чебышёва являются явными реализациями так называемого *закона больших чисел*, утверждающего, что *при проведении достаточно большого числа испытаний погрешности отдельных испытаний взаимно погашают друг друга (тем самым среднее арифметическое независимых случайных величин – результатов этих испытаний – стремится к постоянной величине при неограниченном увеличении числа испытаний).*

VI. Математическая статистика

Выборочный метод

1.Оценка неизвестного параметра. Свойства оценок

Определение. *Случайная величина H_n называется* ***оценкой*** *неизвестного параметра h, если значение этой случайной величины, найденное по результатам серии из n измерений, может быть принято за приближенное значение этого параметра т.е. если справедливо равенство $h \approx H_n$.*

Пример. Если в качестве неизвестного параметра рассматривается вероятность $p = P(A)$ наступления некоторого события A, то оценкой этого параметра служит частость $\dfrac{m}{n}$ наступлений события A в n независимых испытаниях (см. статистическое определение вероятности и теорему Бернулли).

Пример. Пусть случайные величины $X_1, X_2, \ldots, X_n$ имеют одинаковое математическое ожидание, т.е. $a \equiv M(X_1) = M(X_2) \ldots M(X_n)$. Тогда оценкой значения a общего математического ожидания таких случайных величин служит среднее арифметическое $\dfrac{X_1 + X_2 + \ldots + X_n}{n}$ этих случайных величин. Важным частным случаем рассмотренной ситуации является следующий

Пример. Оценкой некоторого параметра a служит среднее арифметическое $\dfrac{X_1 + X_2 + \ldots + X_n}{n}$ результатов $X_1, X_2, \ldots, X_n$ независимых измерений этого параметра (см. теорему Чебышёва).

При непосредственном использовании приближенного равенства $h \approx H_n$ говорят о *точечном оценивании* неизвестного параметра.

Возможно также *интервальное оценивание* неизвестного параметра. Для того, чтобы объяснить, в чем оно состоит, введем в рассмотрение следующие понятия.

Определение. *Для произвольного $\varepsilon > 0$ интервал $(H_n - \varepsilon, H_n + \varepsilon)$ называется*

доверительным интервалом*; сама величина* ε *называется в этом случае* ***предельной ошибкой выборки****.*

Определение. *Вероятность того, что неизвестное значение оцениваемого параметра накрывается доверительным интервалом, называется* ***доверительной вероятностью.***

Таким образом, если H_n – оценка параметра h, то

$$P(H_n-\varepsilon<h<H\ +\varepsilon)=P(|H_n-h|<\varepsilon)=P(|H_n-h|\le\varepsilon)$$

– доверительная вероятность (мы предполагаем, что оценка H_n является непрерывной случайной величиной).

Интервальное оценивание состоит, например, в вычислении доверительной вероятности для заданной предельной ошибки выборки. Решение задачи интервального оценивания связано с определением характера закона распределения используемой оценки H_n.

Рассмотрим теперь некоторые свойства оценок.

Определение. *Оценка* H_n *параметра* h *называется* ***несмещенной****, если математическое ожидание этой оценки равно оцениваемому параметру, т.е.*

$$M\ (H_n)=h.$$

Определение. *Оценка* H_n *параметра* h *называется* ***состоятельной****, если для произвольного* $\varepsilon>0$ *выполняется следующее предельное соотношение*

$$\lim_{n\to\infty}P(|H_n-h|<\varepsilon)=1.$$

Другими словами, оценка H_n параметра h состоятельна, если эта оценка сходится по вероятности к данному параметру

Определение. *Несмещенная оценка некоторого параметра называется* ***эффективной****, если она обладает наименьшей дисперсией среди всех несмещенных оценок, найденных по выборке заданного объема.*

Пример. Частость m/n наступления некоторого события является несмещенной, состоятельной и эффективной оценкой вероятности p этого события. Заметим, что свойства несмещенности и состоятельности частости были фактически рассмотрены нами ранее в несколько ином контексте. Действительно, несмещенность частости – равенство $M\ (m/n)=p$ – является одним из свойств биномиально распределенной случайной величины. Состоятельность частости утверждается теоремой Бернулли.

Пример. Среднее арифметическое некоторого числа независимых и одинаково распределенных случайных величин является несмещенной и состоятельной оценкой общего математического ожидания этих случайных величин. Действительно, несмещенность – есть свойство 5 математического ожидания. Состоятельность утверждается теоремой Чебышёва.

2.Первичная обработка результатов эксперимента. Характеристики вариационных рядов

Пусть произведено n независимых измерений некоторой случайной величины X : X_1 – результат первого измерения, X_2 – результат второго измерения, … , X_n – результат n-го измерения. Тогда через $\overline{X}$ обозначим *среднее арифметическое* результатов n измерений рассматриваемой случайной величины X, то есть

$$\overline{X}=\frac{X_1+X_2+\ldots+X_n}{n}.$$

Заметим, что, поскольку $X_1,X_2,\ldots,X_n$ – случайные величины, то $\overline{X}$ также является

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

случайной величиной.

Пример. Детали некоторого вида расфасованы по ящикам. Результаты обследования шести из этих ящиков (на предмет наличия в них бракованных деталей) представлены в таблице:

i	1	2	3	4	5	6
X_i	0	1	0	1	2	1

где i – номер ящика, X_i – число бракованных деталей в i -ом ящике.

Тогда

$$\bar{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{0+1+0+1+2+1}{6} = \frac{0\cdot 2 + 1\cdot 3 + 2\cdot 1}{6} = \frac{5}{6}.$$

Приведенное вычисление подсказывает возможность более компактного представления результатов обследования, а именно – использование таблицы следующего вида:

x_i	0	1	2	Σ
n_i	2	3	1	6

где x_i – число бракованных деталей в ящике; n_i – число ящиков.

Такая таблица называется *вариационным рядом.* Аналогично, в общем случае имеем

Определение. *Вариационным рядом* *признака X называется таблица вида*

x_i	x_1	x_2	$\ldots$	x_m	Σ
n_i	n_1	n_2	$\ldots$	n_m	n

где x_i – возможные значения данного признака, n_i – числа объектов, $i = 1, 2, \ldots, m$, n – число обследованных объектов ($\sum_{i=1}^{m} n_i = n$).

Отметим, что величины n_i, значения которых заполняют нижнюю строку вариационного ряда, называются *эмпирическими частотами.*

Очевидно, что признак X , для которого строится вариационный ряд, есть *случайная величина.*

В том случае, когда результаты обследования представлены вариационным рядом, формула для вычисления $\bar{X}$ имеет вид

$$\bar{X} = \frac{\sum_{i=1}^{m} x_i \cdot n_i}{n}. \qquad (1)$$

Сама величина $\bar{X}$ в этом случае называется *средней вариационного ряда* или *выборочной средней.* Появление в данном случае дополнительного эпитета *выборочный* связано с тем, что обследованные объекты выбираются из некоторой объемлющей (так называемой *генеральной*) совокупности объектов.

Напомним, что $\bar{X}$ есть случайная величина. В тех случаях, когда данные эксперимента представлены вариационным рядом, а $\bar{X}$ вычисляется по формуле (1), случайными являются эмпирические частоты n_i .

Вариационный ряд является оценкой закона распределения случайной величины (признака) X . Поясним, почему это так. По вариационному ряду построим равнозначную ему таблицу, заменяя строку эмпирических частот n_i частостями n_i/n . В результате имеем:

ABBYY PDF Transformer 2.0
Click here to buy
www.ABBYY.com

x_i	x_1	x_2	…	x_m	Σ
n_i/n	n_1/n	n_2/n	…	n_m/n	1

Учитывая, что частости n_i/n являются оценками вероятностей $p_i = P(X = x_i)$ ($p_i \approx n_i/n$, см. п.1), приходим к требуемому утверждению.

Принимая во внимание последнее замечание, получаем

$$\bar{X} = \sum_{i=1}^{m} x_i \cdot \frac{n_i}{n} \approx \sum_{i=1}^{m} x_i \cdot p_i = M(X).$$

Таким образом, *средняя вариационного ряда (выборочная средняя)* $\bar{X}$ *является* *<u>оценкой</u> математического ожидания* $M(X)$ *той случайной величины (признака)* X*, для которой построен данный вариационный ряд.* Можно доказать, что эта оценка является несмещенной и состоятельной.

Учитывая полученные результаты, аналогично построим оценку для дисперсии $D(X)$ случайной величины X:

$$D(X) = \sum_{i=1}^{m} (x_i - M(X))^2 \cdot p_i \approx \sum_{i=1}^{m} (x_i - \bar{X})^2 \cdot \frac{n_i}{n}.$$

Выражение, стоящее в правой части последнего равенства называется *выборочной дисперсией* и обозначается s^2, то есть

$$s^2 = \frac{1}{n} \cdot \sum_{i=1}^{m} (x_i - \bar{X})^2 \cdot n_i.$$

Выборочная дисперсия s^2 *– оценка для дисперсии* $D(X)$ *случайной величины* X. Можно доказать, что выборочная дисперсия s^2 является *смещенной* оценкой для $D(X)$, то есть $M(s^2) \neq D(X)$. Несмещенная оценка $\hat{s}^2$ для $D(X)$ определяется равенством

$$\hat{s}^2 = \frac{n}{n-1} \cdot s^2.$$

Заметим, что для вычисления выборочной дисперсии удобно использовать формулу – аналог свойства 3 дисперсии (см. §3):

$$s^2 = \overline{X^2} - (\bar{X})^2.$$

Определение. *Вариационный ряд называется* ***дискретным****, если число возможных значений признака – конечно, и* ***непрерывным (интервальным)****, если возможные значения признака полностью заполняют некоторый интервал.*

Вариационные ряды, которые встречались нам до сих пор в данном параграфе, являются дискретными. Рассмотрим пример интервального вариационного ряда.

Пример. По результатам обследования некоторого малого предприятия получены следующие данные о ежемесячной заработной плате его сотрудников:

$[x_{i-1}, x_i]$	5 – 15	15 – 25	25 – 35	Σ
n_i	3	5	2	10

где x_i – размер заработной платы (ден. Ед.), n_i – число сотрудников.

Для нахождения параметров непрерывного вариационного ряда – выборочной средней, выборочной дисперсии – этот вариационный ряд сначала сводится к дискретному (в результате выбора середины для каждого из рассматриваемых интервалов), после чего $\bar{X}$ и s^2 вычисляются по приведенным выше формулам.

Например, данный интервальный вариационный ряд сводится к следующему дискретному:

x_i	10	20	30	Σ
n_i	3	5	2	10

Тогда

$$\bar{X} = \frac{1}{n} \cdot \sum_{i=1}^{m} x_i \cdot n_i = \frac{1}{10} \cdot (10 \cdot 3 + 20 \cdot 5 + 30 \cdot 2) = 19.$$

$$s^2 = \frac{1}{n} \cdot \sum_{i=1}^{m} (x_i - \bar{X})^2 \cdot n_i = \frac{1}{10}((10-19)^2 \cdot 3 + (20-19)^2 \cdot 5 + (30-19)^2 \cdot 2) = 49$$

или

$$\overline{X^2} = \frac{1}{n} \cdot \sum_{i=1}^{m} x_i^2 \cdot n_i = \frac{1}{10} \cdot (10^2 \cdot 3 + 20^2 \cdot 5 + 30^2 \cdot 2) = 410,$$

$$s^2 = \overline{X^2} - (\bar{X})^2 = 410 - 19^2 = 49.$$

3. Сплошное и выборочное наблюдения

Пусть дана некоторая (*генеральная*) совокупность объектов и требуется оценить значение некоторого параметра этой совокупности (например, среднее значение прибыли для малых предприятий некоторого региона или долю выборщиков, проголосовавших за данного кандидата на выборах).

Предположим, что от полного обследования всей генеральной совокупности решили отказаться. Среди возможных причин здесь можно указать разрушение объекта в результате обследования (в том случае, когда, например, требуется узнать средний срок службы лампочек в партии, изготовленной на некотором заводе, полное обследование, конечно, даст исчерпывающую информацию, но сама совокупность перестанет существовать). Другая возможная причина – высокая стоимость полного обследования или его чрезмерная продолжительность (например, выводы экспресс-анализа результатов голосования на некоторых выборах требуется получить в кротчайшие сроки, что невозможно при тотальном обследовании). Наконец, генеральная совокупность может обладать таким свойством как «необозримость» (например, рыба некоторого вида в данном море).

Тогда из генеральной совокупности выделяют часть (*выборку*). Обследуя ее, находят значение исследуемого параметра в выборке. На основании этих результатов делают вывод о значении этого параметра во всей генеральной совокупности.

Среди основных принципов выборочного метода следует отметить ***случайность*** и ***массовость***. В самом деле, объекты в выборку следует отбирать *случайным* образом, в противном случае объективных данных о генеральной совокупности не получить. Также, следует постараться взять в выборку *так много объектов как возможно*, поскольку малая выборка будет плохо отражать свойства всей генеральной совокупности.

Определение. ***Ошибкой репрезентативности*** *называется ошибка, связанная с тем, что не все объекты генеральной совокупности попадут в выборку (и, тем самым, будут обследованы).*

Заметим, что ошибка репрезентативности выборочного метода принципиально неустранима.

В зависимости от способа формирования, выборки бывают собственно-случайные, механические, типические, серийные. В дальнейшем мы будем рассматривать лишь *собственно-случайные выборки*, которые составляются следующим образом: Предположим, что объекты генеральной совокупности некоторым образом перенумерованы. Из полной совокупности номеров случайным образом отбирают столько номеров, сколько элементов должно быть в выборке. Элементы генеральной совокупности с такими номерами и подвергаются обследованию.

Выборка называется *повторной*, если перед отбором очередного номера из полной совокупности номеров предыдущий номер возвращается назад в совокупность; в противном случае – *бесповторной.*

В данном курсе мы рассмотрим следующие из задач выборочного метода:

– оценка неизвестного значения генерального среднего;

– оценка неизвестного значения генеральной доли.

4. Оценка генеральной средней

Пусть задана генеральная совокупность объектов, для которой фиксирован некоторой числовой признак X. Требуется оценить среднее значение признака X в генеральной совокупности – генеральную среднюю $\bar{X}_0$. Для этого из генеральной совокупности выделяют часть (выборку), и по результатам ее обследования находят среднее значение признака X в выборке – выборочную среднюю $\bar{X}$, с помощью которой и выполняют оценивание неизвестного значения $\bar{X}_0$. Другими словами, <u>выборочная средняя $\bar{X}$ является оценкой генерального среднего $\bar{X}_0$</u>.

Пример. Пусть некоторая совокупность деталей обследуется на предмет их длины. Тогда $\bar{X}_0$ – средняя длина деталей в генеральной совокупности, $\bar{X}$ – средняя длина деталей в выборке, X – длина детали, взятой наудачу из генеральной совокупности.

*В том случае, когда оценивание сводится к использованию приближенного равенства $\bar{X}_0 \approx \bar{X}$, говорят о **точечном оценивании** генеральной средней* (см. §1).

Возможно также *интервальное оценивание* генеральной средней (см. §1). Для того чтобы объяснить, в чем оно состоит, введем в рассмотрение следующие понятия.

Определение. *Для произвольного $\varepsilon > 0$ интервал $(\bar{X} - \varepsilon, \bar{X} + \varepsilon)$ называется **доверительным интервалом**; величина ε называется в этом случае **предельной ошибкой выборки**.*

Определение. *Вероятность того, что неизвестное значение генеральной средней $\bar{X}_0$ накрывается доверительным интервалом, называется **доверительной вероятностью**.*

Таким образом,

$$P(|\bar{X} - \bar{X}_0| \le \varepsilon)$$

– доверительная вероятность.

***Интервальное оценивание** состоит, например, в вычислении доверительной вероятности для заданной предельной ошибке выборки.*

Как и всякая оценка, выборочная средняя $\bar{X}$ является *случайной величиной.* Действительно, элементы выборки отбираются из генеральной совокупности случайным образом, а значение $\bar{X}$ зависит от того, какие именно элементы попали в выборку. Рассмотрим свойства выборочной средней $\bar{X}$ как случайной величины.

Теорема 1. *Математическое ожидание выборочной средней $\bar{X}$ равно генеральной средней $\bar{X}_0$, то есть*

$$M(\bar{X}) = \bar{X}_0.$$

Среднее квадратическое отклонение $\sigma_{\bar{X}}$ $(\sigma'_{\bar{X}})$ выборочной средней вычисляется по формулам

$$\sigma_{\bar{X}} = \sqrt{\frac{\sigma^2}{n}}$$

– в случае повторной выборки и

$$\sigma'_{\bar{X}} = \sqrt{\frac{\sigma^2}{n} \cdot \left(1 - \frac{n}{N}\right)}$$

– в случае бесповторной,

где n – объем выборки, N – объем генеральной совокупности, $\sigma^2 = D(X)$ – дисперсия признака X для рассматриваемой генеральной совокупности (генеральная дисперсия).

Напомним, что, по определению среднего квадратического отклонения, $\sigma_{\bar{X}}$ равно корню квадратному из дисперсии выборочной средней, то есть

$$\sigma_{\bar{X}} = \sqrt{D(\bar{X})}$$

(аналогично в случае бесповторной выборки).

Замечание. *При применении на практике формул Теоремы 1 полагают, что*

$$\sigma^2 \approx s_X^2 .$$

Теорема 2. *Закон распределения выборочной средней неограниченно приближается к нормальному при неограниченном увеличении объёма выборки.*

Согласно результатам §3, для произвольной нормально распределенной случайной величины Z справедлива формула

$$P\left(\left|Z - M(Z)\right| \le \varepsilon\right) \quad \Phi\left(\frac{\varepsilon}{\sqrt{D(Z)}}\right).$$

Учитывая Теорему 2, в последнем равенстве положим $Z = \bar{X}$. Тогда, по Теореме 1, $M(Z) = M(\bar{X}) = \bar{X}_0$ и $D(Z) = D(\bar{X}) = \sigma_{\bar{X}}^2$, и приведенная формула – свойство нормального закона распределения принимает вид:

$$\boxed{P\left(\left|\bar{X} - \bar{X}_0\right| \le \varepsilon\right) \quad \Phi\left(\frac{\varepsilon}{\sigma_{\bar{X}}}\right)}.$$

Вероятность, стоящая в левой части последнего равенства называется *доверительной вероятностью* (см. выше), поэтому сама эта формула называется ***формулой доверительной вероятности***.

Теорема 3. *Выборочная средняя $\bar{X}$ является несмещенной и состоятельной оценкой генеральной средней $\bar{X}_0$.*

Пример. Для обследования средней заработной платы трехсот рабочих была образована выборка, состоящая из пятидесяти рабочих. Результаты выборочного обследования представлены в таблице:

Заработная плата в месяц, ден. Ед.	100-120	120-140	140-160	160-180	180-200	200-220	Σ
Число рабочих	5	10	19	10	4	2	50

1. Найти вероятность того, что средняя заработная плата всех рабочих отличается от средней выборочной не более чем на 5 ден. Ед. (по абсолютной величине) в случае повторной и бесповторной выборок.
2. Найти границы, в которых с вероятностью 0,9545 заключена средняя заработная плата всех рабочих.
3. Сколько рабочих надо взять в выборку, чтобы полученные в п. 2 доверительные границы можно было гарантировать с вероятностью 0,9973.

Решение. Исходный вариационный ряд является интервальным. Для нахождения его характеристик, прежде всего, сведем этот вариационный ряд к дискретному:

x_i	110	130	150	170	190	210	Σ
n_i	5	10	19	10	4	2	50

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

где x_i – возможное значение заработной платы – середина i- го интервала исходного вариационного ряда (ден. Ед.); n_i – число рабочих; $n = 50$.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{6} x_i n_i =$$

$$= \frac{1}{50}\left(110\cdot 5 + 130\cdot 10 + 150\cdot 19 + 170\cdot 10 + 190\cdot 4 + 210\cdot 2\right) = 151{,}6.$$

$$\overline{X^2} = \frac{1}{n}\sum_{i=1}^{6} x_i^2 n_i =$$

$$= \frac{1}{50}\left(110^2\cdot 5 + 130^2\cdot 10 + 150^2\cdot 19 + 170^2\cdot 10 + 190^2\cdot 4 + 210^2\cdot 2\right) = 23572.$$

$$s_x^2 = \overline{X^2} - \left(\bar{X}\right)^2 = 23572 - 151{,}6^2 = 589{,}44.$$

Для нахождения доверительной вероятности (см. п. 1 задания) воспользуемся одноименной формулой при $\varepsilon = 5$. Но сначала вычислим средние квадратические отклонения выборочной средней для каждого из рассматриваемых типов выборок.

А) Повторная выборка.

$$\sigma_{\bar{X}} = \sqrt{\frac{s_X^2}{n}} = \sqrt{\frac{589{,}44}{50}} = 3{,}433.$$

$$P\left(\left|\bar{X} - \bar{X}_0\right| \le 5\right) = \Phi\left(\frac{5}{\sigma_{\bar{X}}}\right) = \Phi\left(\frac{5}{3{,}433}\right) = \Phi\left(1{,}46\right) = 0{,}8557.$$

б) Бесповторная выборка, $N = 300$.

$$\sigma'_{\bar{X}} = \sqrt{\frac{s_X^2}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{589{,}44}{50}\left(1 - \frac{50}{300}\right)} = 3{,}134.$$

$$P\left(\left|\bar{X} - \bar{X}_0\right| \le 5\right) = \Phi\left(\frac{5}{\sigma'_{\bar{X}}}\right) = \Phi\left(\frac{5}{3{,}134}\right) = \Phi\left(1{,}60\right) = 0{,}89.$$

Доверительный интервал в данном случае:

$$\left(\bar{X} - \varepsilon, \bar{X} + \varepsilon\right) = \left(151{,}6 - 5; 151{,}6 + 5\right) = \left(146{,}6; 156{,}6\right).$$

Тем самым получаем, что: *неизвестное значение средней заработной платы всех рабочих накрывается интервалом* (146,6;156,6) *с вероятностью* 0,8557 *в случае повторной выборки и с вероятностью* 0,89 *в случае бесповторной выборки.*

В п. 2 задания искомым является доверительный интервал, для нахождения которого следует вычислить предельную ошибку выборки ε. Из условия и формулы доверительной вероятности в случае повторной выборки следует, что

$$\Phi\left(\frac{\varepsilon}{\sigma_{\bar{X}}}\right) = 0{,}9545.$$

По таблице значений функции Лапласа найдем такое значение t, что $\Phi(t) = 0{,}9545$. Имеем $t = 2$. Поскольку

$$\frac{\varepsilon}{\sigma_{\bar{X}}} = t,$$

то

$$\varepsilon = t\sigma_{\bar{X}} = 2\cdot 3{,}433 = 6{,}866.$$

Соответствующий доверительный интервал:

$$\left(\bar{X} - \varepsilon, \bar{X} + \varepsilon\right) = \left(151{,}6 - 6{,}866; 151{,}6 + 6{,}866\right) \approx \left(144{,}73; 158{,}47\right).$$

Аналогично, в случае бесповторной выборки имеем

$$\varepsilon = t\sigma'_{\overline{X}} = 2\cdot 3{,}134 = 6{,}268\,.$$

Соответствующий доверительный интервал:

$$\left(\overline{X}-\varepsilon, \overline{X}+\varepsilon\right) \quad (151{,}6-6{,}268; 151{,}6+6{,}268) \approx (145{,}33; 157{,}87).$$

Таким образом, *неизвестное значение средней заработной платы всех рабочих с вероятностью* 0,9545 *накрывается доверительным интервалом* (144,73; 158,47) *в случае повторной выборки и доверительным интервалом* (145,33; 157,87) *в случае бесповторной выборки.*

При решении п. 3 задания будем считать известными приближенные значения выборочной средней $\overline{X}$ и выборочной дисперсии s_X^2. Также используем предельные ошибки выборки ε, найденные в п. 2. Рассмотрим сначала случай повторной выборки. Из условия и формулы доверительной вероятности следует, что

$$\Phi\left(\frac{\varepsilon}{\sigma_{\overline{X}}}\right) = 0{,}9973\,.$$

По таблице значений функции Лапласа найдем такое значение аргумента t, что $\Phi(t) = 0{,}9973$: $t = 3$. Тогда

$$\frac{\varepsilon}{\sigma_{\overline{X}}} = t \ \text{ и } \ \varepsilon = t\sigma_{\overline{X}}\,.$$

Используя известную формулу для $\sigma_{\overline{X}}$ (см. Теорему 2 данного параграфа), имеем равенство:

$$\varepsilon = t\sqrt{\frac{s_X^2}{n}}\,,$$

в котором единственной неизвестной является искомый объем выборки n. Решая получившееся уравнение относительно n, получаем

$$n = \frac{t^2 s_X^2}{\varepsilon^2}\,.$$

Подставляя в правую часть последнего равенства известные величины, получаем

$$\varepsilon = \frac{3^2\cdot 589{,}44}{6{,}866^2} = 112{,}7 \approx 113$$

(заметим, что округление в данном случае, по смыслу искомой величины, следует произвести до целых, причем в <u>большую сторону</u>, чтобы обеспечить, как говорят, запас по вероятности).

Повторяя проведенные рассуждения для случая бесповторной выборки, имеем:

$$\varepsilon = t\sigma'_{\overline{X}}\,,$$

$$\varepsilon = t\sqrt{\frac{s_X^2}{n}\left(1-\frac{n}{N}\right)}\,.$$

Решая полученное уравнение относительно n, получаем

$$\varepsilon^2 = t^2 s_X^2\left(\frac{1}{n}-\frac{1}{N}\right),$$

откуда

$$\frac{1}{n} = \frac{\varepsilon^2}{t^2 s_X^2}+\frac{1}{N}\,,$$

$$n = \frac{\varepsilon^2 N + t^2 s_X^2}{t^2 s_X^2} = \frac{6{,}268^2\cdot 300 + 3^2\cdot 589{,}44}{3^2\cdot 589{,}44} = 93{,}4 \approx 94$$

(также как и выше округление здесь произведено в большую сторону).

Таким образом, *для того, чтобы с вероятностью* 0,9973 *неизвестное значение средней заработной платы всех рабочих накрывалось доверительным интервалом* (144,73; 158,47) *в случае повторной выборки, в эту выборку следует взять* 113 *рабочих.* Аналогично, *для того, чтобы с вероятностью* 0,9973 *неизвестное значение средней заработной платы всех рабочих накрывалось доверительным интервалом* (145,33; 157,87) *в случае бесповторной выборки, в выборку следует взять* 94 *рабочих.*

Замечание. *Если в задаче на выборочный метод объем генеральной совокупности много больше объема выборки (в ряде случаев это предполагается по умолчанию, а объем генеральной совокупности просто не указан), естественно считать, что* $\frac{n}{N} \approx 0$. *Как следует из формул Теоремы 1, случаи повторной и бесповторной выборок дают тогда совпадающие результаты.*

5 Оценка генеральной доли

Пусть требуется оценить долю тех объектов заданной генеральной совокупности, которые удовлетворяют некоторому условию A – *генеральную долю* p. Для этого из генеральной совокупности выделяют выборку, и по результатам её обследования находят долю тех объектов, которые удовлетворяют условию A – *выборочную долю* ω. Очевидно, что $\omega = m/n$, где n – объем выборки, m – число тех её объектов, которые удовлетворяют условию A. Выборочная доля в данном случае является той величиной, с помощью которой мы получим информацию о неизвестном значении генеральной доли.

Таким образом, *выборочная доля* ω *является оценкой генеральной доли* p.

Пример. p – доля бракованных деталей генеральной совокупности, ω – доля бракованных деталей в выборке. Условие (событие) A – деталь, взятая наудачу из генеральной совокупности – бракована.

Простейший способ оценивания – *точечное оценивание* – подразумевает использование приближенного равенства $p \approx \omega$. Как и всякая оценка, выборочная доля ω является случайной величиной. Действительно, выборка из генеральной совокупности выделяется случайным образом. Соответственно то значение, которое примет выборочная доля, будет случайным. Следующие теоремы характеризуют выборочную долю как случайную величину.

Теорема 1. *Математическое ожидание выборочной доли равно генеральной доле:*

$$M(\omega) = p.$$

Среднее квадратическое отклонение $\sigma_\omega (\sigma'_\omega)$ *выборочной доли вычисляется по формулам*

$$\sigma_\omega = \sqrt{\frac{p(1-p)}{n}}$$

– в случае повторной выборки и

$$\sigma'_\omega = \sqrt{\frac{p(1-p)}{n}\left(1-\frac{n}{N}\right)}$$

– в случае бесповторной выборки, где N *– объем генеральной совокупности.*

Напомним, что по определению среднего квадратического отклонения в случае повторной выборки имеем $\sigma_\omega = \sqrt{D(\omega)}$ (аналогично в случае бесповторной выборки).

Замечание. *При применении формул Теоремы 1 полагают*

$$p \approx \omega.$$

Теорема 2. *Закон распределения выборочной доли неограниченно приближается к*

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

нормальному закону при неограниченном увеличении объема выборки.

Подобно тому, как мы это сделали в предыдущем параграфе, как следствие Теоремы 2, получаем *формулу доверительной вероятности*:

$$\boxed{P\left(\left|\omega - p\right| \le \varepsilon\right) \quad \Phi\left(\frac{\varepsilon}{\sigma_\omega}\right)}$$

– в случае повторной выборки. Заменяя в последнем равенстве σ_ω на σ'_ω, получаем формулу доверительной вероятности в случае бесповторной выборки.

По определению, величина ε, фигурирующая в формуле доверительной вероятности, называется *предельной ошибкой выборки*. Интервал $(\omega - \varepsilon; \omega + \varepsilon)$ называется *доверительным интервалом.*

Выше было указано, в чем состоит точечная оценка генеральной доли. *Интервальное оценивание* сводится, например, к вычислению значения доверительной вероятности при заданной предельной ошибке выборки.

Теорема 3. *В случае повторной выборки выборочная доля является несмещенной и состоятельной оценкой генеральной доли.*

Пример. Выборочные данные о надое молока для 100 коров из 1000 представлены таблицей:

Надой молока, *ц*	10-20	20-30	30-40	40-50	50-60	Σ
Число коров	2	18	46	30	4	100

1.Найти вероятность того, что доля всех коров с надоем молока более 40 ц отличается от такой доли в выборке не более чем на 0,05 (по абсолютной величине), для случая повторной и бесповторной выборок.
2.Найти границы, в которых с вероятностью 0,9596 заключена доля всех коров с надоем более 40 ц.
3.Сколько коров надо обследовать, чтобы с вероятностью 0,9786 для генеральной доли коров с надоем более 40 ц можно было гарантировать те же границы что и в п.2.

Решение. Число m коров с надоем более 40ц равно 34 ($m = 30 + 4$, см. заданный вариационный ряд). Тогда $\omega = \frac{m}{n} = \frac{34}{100} = 0,34$.

Для нахождения доверительной вероятности п. 1 задания воспользуемся одноименной формулой при $\varepsilon = 0,05$.

Пусть рассматриваемая выборка – повторная. Тогда по формуле Теоремы 1, учитывая Замечание, получаем

$$\sigma_\omega = \sqrt{\frac{\omega(1-\omega)}{n}} = \sqrt{\frac{0,34 \cdot 0,66}{100}} = 0,04737.$$

Следовательно

$$P\left(\left|\omega - p\right| \le \varepsilon\right) \quad \Phi\left(\frac{\varepsilon}{\sigma_\omega}\right) \quad \Phi\left(\frac{0,05}{0,04737}\right) \quad \Phi(1,06) \quad 0,7109.$$

Аналогично, в случае бесповторной выборки:

$$\sigma'_\omega = \sqrt{\frac{\omega(1-\omega)}{n}\left(1 - \frac{n}{N}\right)} = \sqrt{\frac{0,34 \cdot 0,66}{100}\left(1 - \frac{100}{1000}\right)} = 0,04494,$$

$$P\left(\left|\omega - p\right| \le \varepsilon\right) \quad \Phi\left(\frac{\varepsilon}{\sigma'_\omega}\right) \quad \Phi\left(\frac{0,05}{0,04494}\right) \quad \Phi(1,11) \quad 0,733.$$

ABBYY PDF Transformer 2.0
Click here to buy
www.ABBYY.com

Доверительным в данном случае является интервал $(\omega-\varepsilon;\omega+\varepsilon)=(0{,}34-0{,}05;0{,}34+0{,}05)=(0{,}29;0{,}39)$. Таким образом, *неизвестное значение доли всех коров с надоем более* 40 ц *накрывается доверительным интервалом* (0,29;0,39) *с вероятностью* 0,7109 *в случае повторной выборки и с вероятностью* 0,733 *в случае бесповторной выборки.*

В п. 2 задания при заданном значении доверительной вероятности искомым является доверительный интервал. Поскольку значение выборочной доли известно, остается найти предельную ошибку выборки ε.

Пусть выборка – повторная. По условию, принимая во внимание формулу доверительной вероятности, имеем

$$P\left(\left|\omega-p\right|\le\varepsilon\right)=\Phi\left(\frac{\varepsilon}{\sigma_\omega}\right)=0{,}9596.$$

По таблице значений функции Лапласа найдем такое t, что $\Phi(t)=0{,}9596$: $t=2{,}05$. Тогда $\varepsilon/\sigma_\omega=t$ и, используя найденное выше значение σ_ω, получаем

$$\varepsilon=t\sigma_\omega=2{,}05\cdot 0{,}04737=0{,}097.$$

Соответственно, доверительным будет интервал:

$$(\omega-\varepsilon;\omega+\varepsilon)=(0{,}34-0{,}097;0{,}34+0{,}097)=(0{,}243;0{,}437).$$

Пусть выборка – бесповторная. Аналогично предыдущему, получаем предельную ошибку выборки

$$\varepsilon=t\sigma'_\omega=2{,}05\cdot 0{,}04494=0{,}092$$

и доверительный интервал:

$$(\omega-\varepsilon;\omega+\varepsilon)=(0{,}34-0{,}092;0{,}34+0{,}092)=(0{,}248;0{,}432).$$

Таким образом, *доля всех коров с надоем молока более* 40 *ц с вероятностью* 0,9596 *накрывается доверительным интервалом* (0,243; 0,437) *в случае повторной выборки и интервалом* (0,248; 0,432) *в случае бесповторной выборки.*

В п. 3 по заданным значениям доверительной вероятности и предельной ошибки выборки найдем необходимый объем выборки. Из начала решения заимствуем значение выборочной доли ω, найденное по исходному вариационному ряду.

Пусть выборка – повторная. По условию, принимая во внимание формулу доверительной вероятности, имеем:

$$P\left(\left|\omega-p\right|\le\varepsilon\right)=\Phi\left(\frac{\varepsilon}{\sigma_\omega}\right)=0{,}9786.$$

По таблице значений функции Лапласа найдем такое t, что $\Phi(t)=0{,}9786$: $t=2{,}3$. Тогда $\varepsilon/\sigma_\omega=t$ и, $\varepsilon=t\sigma_\omega$. Подставляя вместо σ_ω выражение из Теоремы 1, приходим к уравнению относительно неизвестной величины n:

$$\varepsilon=t\sqrt{\frac{\omega(1-\omega)}{n}}.$$

Решая это уравнение относительно n, подставляя в полученную формулу известные величины, завершаем решение

$$n=\frac{t^2\omega(1-\omega)}{\varepsilon^2}=\frac{2{,}3^2\cdot 0{,}34\cdot 0{,}66}{0{,}097^2}=126{,}16\approx 127$$

(заметим, что, как и ранее, округление здесь произведено в большую сторону).

Аналогично, в случае бесповторной выборки из условия и формулы доверительной вероятности следует равенство

$$\varepsilon=t\sigma'_\omega$$

или, принимая во внимание известное выражение для σ'_ω (см. Теорему 1):

$$\varepsilon = t\sqrt{\frac{\omega(1-\omega)}{n}\left(1-\frac{n}{N}\right)}.$$

Решая это уравнение относительно n, получаем

$$n = \frac{t^2\omega(1-\omega)N}{\varepsilon^2 N + t^2\omega(1-\omega)}.$$

Подставляя в правую часть последнего равенства известные значения, окончательно имеем:

$$n = \frac{2{,}3^2\cdot 0{,}34\cdot 0{,}66\cdot 1000}{0{,}092^2\cdot 1000 + 2{,}3^2\cdot 0{,}34\cdot{,}66} = 122{,}9993 \approx 123.$$

Таким образом, *в повторную выборку надо взять* 127 *коров, чтобы с вероятностью* 0,9786 *можно было утверждать, что доля всех коров с надоем молока более* 40 *ц накрывается доверительным интервалом* (0,243; 0,437). Аналогично, *в бесповторную выборку надо взять* 123 *коровы, чтобы с вероятностью* 0,9786 *можно было утверждать, что доля всех коров с надоем молока более* 40 *ц накрывается доверительным интервалом* (0,248; 0,432).

VII. Перечень модульных вопросов теоретического курса по дисциплине «Теория вероятностей и математическая статистика»

1 модуль

1. Определение среднего квадратического отклонения случайной величины.

2. Что такое ковариация двух случайных величин?
3. Что называется коэффициентом корреляции, и каковы его свойства?
4. Что называется характеристическими функциями случайной величины? Сформулируйте их свойства.
5. Напишите неравенство Чебышева. Сформулируйте теорему Чебышева.
6. Сформулируйте теорему Бернулли.
7. Определение начальных моментов k-го порядка случайной величины .
8. Определение центральных моментов k-го порядка случайной величины .
9. Определение начальных моментов для пары дискретных и непрерывных случайных величин Х и У
10. Определение центральных моментов для пары дискретных и непрерывных случайных величин Х и У.
11. Определение дисперсии случайной величины и перечислите ее свойства.
12. Определение математического ожидания случайной величины и перечислите его свойства.
13. Сформулируйте теоремы о независимых случайных величинах. Что представляет собой распределение суммы независимых случайных величин?
14. Как найти вероятность попадания пары случайных величин в заданный прямоугольник?
15. Как найти вероятность попадания случайной величины в заданный интервал, если она распределена по нормальному закону?
16. Что называется производящей функцией случайной величины?
17. Как определяется функция распределения случайной величины, и перечислите ее свойства.
18. Как определяется функция плотности распределения вероятностей, и перечислите ее свойства.

19. Определение случайной величины. Приведите примеры.
20. Определение условной вероятности.
21. Какие события называются независимыми?
22. Дайте определение произведения (совмещения) событий. Сформулируйте теоремы умножения.
23.Напишите формулу полной вероятности, формулу Байеса.
24.Изложите схему Бернулли и напишите формулу Бернулли.
25.Сформулируйте локальную теорему Муавра-Лапласа.
26.. Сформулируйте теорему Пуассона.
27. Что называется стандартным интегралом вероятности (функцией Лапласа или функцией ошибок)?
28. Как связана дифференциальная функция распределения вероятности f(x) с интегральной функцией распределения вероятности F(x).
29. Как определяется мода, медиана случайной величины?
30. Дайте описания дискретных распределений: биномиального и пуассоновского.

31. Дайте описания непрерывных распределений: нормального, показательного, равномерного.

2 модуль

1.Что понимается под генеральной совокупностью?
2.Что называется выборкой?
3.Объясните, как получают вариационный ряд распределения.
4.Как выполняется чертеж многоугольника распределения относительных частот?
5.Как выполняется чертеж гистограммы распределения плотности относительных частот?
6.Как вычисляется средняя арифметическая выборки при малом и больших объемах ее?
7.Как вычисляется дисперсия выборки в случаях малого и большого объема ее?
8.Напишите формулу для вычисления выборочной средней.
9.Какие оценки называются точечными?
10.Дайте определения несмещенной и состоятельной оценок.
11.Какие оценки являются интервальными?
12.Метод наибольшего правдоподобия?
13.Как методом наибольшего правдоподобия пользоваться для непрерывных случайных величин?
14.Как найти доверительные интервалы для оценки математического ожидания нормального распределения?
15.Дайте определение статистической гипотезы, приведите примеры статистической проверки гипотез.
16.Опишите форму корреляционной таблицы.
17.Сформулируйте две основные задачи корреляционного анализа.
18.Как получают эмпирическую линию регрессии?
Какова форма линии регрессии при линейной корреляционной зависимости?
19.Что следует сказать о двух случайных величинах, если коэффициент корреляции равен нулю или же равен единице?
20.Дайте определение случайного процесса.
21.Что называется реализацией (или траекторией) случайного процесса?
22.Какой процесс называется процессом с независимыми приращениями?
23.Изложите сущность пуассоновского процесса.
24.Дайте определение цепи Маркова и сформулируйте ее основные свойства.
25.Что называется переходной вероятностью (вероятностью перехода)?
26.Что называется математическим ожиданием случайного процесса?
27.Что называется дисперсией и среднеквадратическим отклонением случайного процесса?
28.Что называется корреляционной функцией случайного процесса ?
29.Что называется нормированной корреляционной функцией случайного процесса ?
30.Дайте определение конечномерной n-й функции распределения случайного процесса.

31.Дайте определение дифференциальной n-й конечномерной функции распределения случайного процесса.
32.Как определяется стационарный случайный процесс.
33.Как определяются Марковские случайные процессы.
34.Что называется спектром (или спектральной плотностью) стационарного случайного процесса.
35.Что называется корреляционной функцией случайного процесса.
36.Как определяется стационарный случайный процесс, обладающий свойством эргодичности

VIII. Образцы модульных контрольных работ

1 модуль

1. Сформулируйте локальную теорему Муавра-Лапласа.
2. Дайте определение дисперсии случайной величины и перечислите ее свойства.
3. Что такое корреляция двух случайных величин?
4. Как найти вероятность попадания случайной величины в заданный интервал, если она распределена по показательному закону?
5. Напишите неравенство Чебышева. Сформулируйте теорему Чебышева.
6. В урне 30 шаров: 10 красных, 5 синих и 15 белых. Найти вероятность появления цветного шара.
7. Среди 50 изделий 20 окрашенных. Найти вероятность того, что среди наудачу извлеченных 5 изделий окажется ровно 3 окрашенных.
8. Найти дисперсию случайной величины зная закон ее распределения

X	0,1	2	10	20
P	0,4	0,2	0,15	0,25

9. Случайная величина X распределена по нормальному закону. Математическое ожидание и среднее квадратическое отклонение этой величины соответственно равны 30 и 10. Найти вероятность того, что X примет значение, принадлежащее интервалу (10;50)
10. Задана плотность совместного распределения непрерывной двумерной случайной величины *(X,У): f{x,y) = Ccosxcosy* в квадрате $0 \le x \le \pi/2$, $0 \le y \le \pi/2$ вне этого квадрата *f(x,y)*=0. Найти постоянный параметр С.

2 модуль

1. Какова вероятность попадания генеральной средней в интервал размером (±з) среднеквадратических отклонений средней выборки при нормальном распределении
2. Как найти доверительные интервалы для оценки математического ожидания нормального распределения.
3. Что следует сказать о двух случайных величинах, если коэффициент корреляции равен нулю или же единице?
4. Что называется реализацией (или траекторией) случайного процесса?
5. Сформулируйте теорему о предельных вероятностях?
6. Найти доверительный интервал для оценки математического ожидания a нормального распределения с надежностью 0,95, зная выборочную среднюю $x = 75$, объем выборки $n = 36$ и среднее квадратическое отклонение $\sigma = 6$.

5. Результаты обследования 20 семей по числу членов оказались таким: 2; 5; 3; 4; $1:3:$ 6; 2; 4; 3; 4; 1; 3; 5, 2; 3; 4; 3; 4; 3. Получить по этим данным вариационный ряд и построить

полигон распределения относительных частот.

8) Стационарна ли случайная функция $X(t)=\sin(t+\varphi)$, где φ-случайная величина, распределенная равномерно в интервале $(0;\pi)$?

9) Найти спектральную плотность стационарной случайной функции *X(t)*, зная ее корреляционную функцию $K_x(\tau)=5e^{-2|\tau|}$

6. Задана матрица перехода $P_1=\begin{pmatrix}0.4 & 0.6\\ 0.3 & 0.7\end{pmatrix}$ Найти матрицу перехода

$$P_2=\begin{pmatrix}P_{11}(2) & P_{12}(2)\\ P_{21}(2) & P_{22}(2)\end{pmatrix}$$

IX. КОНТРОЛЬНЫЕ ЗАДАНИЯ

№ I. В сумке лежат 10 мячей, пронумерованных от I до 10. Наугад вынимаются два мяча. Какова вероятность, что это будут мячи с номерами 7 и 3 ?

№2. На восьми одинаковых карточках написаны соответственно числа 2, 4, 6, 7, 8, II, 12 и 17. Наугад берутся две карточ И и из чисел образуется правильная дробь. Найти вероятность, что она сократима.

№3. Десять книг на одной полке расставляются наудачу. Определить вероятность того, что при этом три определенные книги окажутся поставленными рядом.

№ 4. Куб, все грани которого окрашены, распилен на 125 кубиков одинакового размера. Полученные кубики тщательно перемешаны. Определить вероятность того, что кубик, извлеченный наудачу, будет иметь:

А) одну окрашенную грань; *б)* две окрашенные грани; в) три окрашенные грани;

Г) ни одной окрашенной грани.

№5. Из десяти билетов – выигрышными являются два. Определить вероятность того, что среди взятых наудачу 5 билетов: а) один выигрышный; б) два выигрышных; в) хотя бы один выигрышный.

№6. В партии состоящей из 12 изделий, имеется 4 дефектных. Для контроля из партии выбирается 6 изделий. Найти вероятность того, что из них ровно 2 изделия будут дефектными.

№ 7. На складе 10 телевизоров, из них 6 с цветным изобра ИИ ием. Наудачу выбраны 2 телевизора. Найти вероятность того, что: а) оба телевизора с цветным изображением;

Б) один с цветным изображением, другой с черно-белым.

№ 8. Из двенадцати билетов, пронумерованных от I до 12, один за другим выбирают два билета (без возвращения). Какова вероятность того, что: а) оба номера четные; б) оба номера нечетные; в) первый номер четный, а второй нечетный; г) один номер четный, а другой нечетный.

№ 9. В урне 7 белых и 8 черных шаров. Одновременно наудачу извлекают два шара. Найти вероятность того, что: а) оба шара одного цвета; б) шары разных цветов.

№10. Бросаются две игральные кости. Найти вероятность следующих событий:
а/ сумма выпавших очков не равна II; б) на каждой кости выпало менее трех очков;
в) только на одной кости выпало меньше трех очков.

№ II. На 5-ти одинаковых карточках написаны соответственно буквы: а, м, р, *и* , т . Карточки тщательно перемешаны. Наудачу извлечены 3 карточки и расположены в порядке их появления. Найти вероятность того, что при этом получится слово «МИР».

№ 12. При одном цикле обзора радиолакационной станции, следящей за космическим объектом, объект обнаруживается с вероятностью $p = 0{,}6$. Обнаружение объекта в каждом цикле происходит независимо от других. Найти вероятность того, что при 3-ех циклах объект будет обнаружен.

№13. При включении зажигания двигатель начинает работать с вероятностью 0,95. Найти вероятность того: а) что двигатель начинает работать при втором включении зажигания; б) что для ввода двигателя в работу придется включить зажигание не более двух раз.

№ 14. Монета бросается 6 раз. Найти вероятность того, что выпадает больше гербов, чем цифр.

№15. В первой урне находятся 5 белых, II красных и 8 черных шаров, во второй урне 10 белых, 8 красных и 6 черных шаров, отличающихся только цветом. Из каждой урны извлекают по одному шару. Какова вероятность, что оба шара одного цвета?

№ 16. Каждое из 4-х несовместных событий может произойти соответственно с вероятностями 0,012; 0,010; 0,006 и 0,002. Определить вероятность того, что в результате опыта произойдет хотя бы одно из этих событий.

№ 17. При работе электронной вычислительной машины время от времени возникают неисправности (сбои). Поток сбоев можно считать простейшим. Среднее число сбоев за сутки равно 1,5. Найти вероятности следующих событий: A – за двое суток не будет ни одного сбоя; B – в течение суток произойдет хотя бы один сбой; C – за неделю работы произойдет не менее 3 сбоев.

№ 18. Радиоаппаратура состоит из 1000 электроэлементов. Вероятность отказа одного элемента в течение одного года работы равна 0,001 и не зависит от состояния других элементов. Какова вероятность отказа двух и не менее двух электроэлементов за год?

№ 19. Найти вероятность того, что среди 200 изделий окажется более трех бракованных, если в среднем бракованные изделия составляют 1%.

№20. Корректура в 500 страниц содержит 500 опечаток. Найти вероятность того, что на странице не меньше трех опечаток.

№ 21. Вероятность того, что изделие не выдержит испытания, равна 0,001. Найти вероятность того, что из 5000 изделий более чем одно не выдержит испытания.

№ 22. Покупатель посещает близлежащие магазины с равной вероятностью. Вероятность того, что он сделает покупку в 1-ом магазине равна 0,4, во втором 0,45, в третьем 0,6. После третьего посещения он совершил покупку. Найти вероятность, что он купил вещь во 2-ом магазине.

№ 23. Прибор состоит из двух узлов. Прибор выходит из строя, если откажет хотя бы один из узлов. Надежность (вероятность безотказной работы) первого узла равна 0,9, второго 0,95. Прибор после испытания вышел из строя. Найти вероятность, что отказал только первый узел, а второй исправен.

№ 24.Из 18 стрелков пятеро попадают в мишень с вероятностью 0,8; семеро – с вероятностью 0,7; четверо – с вероятностью 0,6 и двое – с вероятностью 0,5. Наудачу выбранный стрелок произвел выстрел, но в мишень не попал. К какой из групп вероятнее всего принадлежал этот стрелок?

№ 25. Два автомата производят детали, которые поступают на общий конвейер. Производительность первого автомата втрое больше производительности второго. Вероятности изготовления годных деталей 1-ым и 2-ым автоматами соответственно равны 0,9 и 0,6. С конвейера взята наудачу I деталь. Найти вероятность того, что она годная.

№ 26. Оптовая база снабжает 10 магазинов, от каждого из которых может поступить заявка на очередной день с вероятностью 0,4 независимо от заявок других магазинов. Найти наивероятнейшее число заявок в день и вероятность получения этого числа заявок.

№ 27. Для данного баскетболиста вероятность забросить мяч в корзину при броске равна 0,4. Произведено 10 бросков. Найти наивероятнейшее число попаданий и соответствующую вероятность.

№ 28. Вероятность попадания, стрелком в десятку равна 0,7, а в девятку – 0,3. Определить вероятность того, что данный стрелок при трех выстрелах наберет не менее 29 очков.

№ 29. Вероятность появления некоторого события в каждом из восемнадцати независимых опытов равна 0,2. 0пределить вероятность появления этого события по крайней мере три раза.

№30. Вероятность того, что лампа останется исправной после1000 часов работы, равна 0,2. Какова вероятность того, что хотя бы одна из трех ламп останется исправней после 1000 часов работы?

№ 31. Сколько семян надо отобрать для определения процента всхожести, чтобы с вероятностью 0,977 можно было бы утверждать, что отклонение частости доброкачественных семян от их доли, равной 0,9 не превышало по абсолютной величине 0,03?

№ 32. Доля брака в некоторой продукции составляет 3%. В партии 800 изделий. Каково наиболее вероятное число бракованных изделий в этой партии и какова соответствующая вероятность?

№ 33. Вероятность попадания в цель при одном выстреле равна 0,6. Произведено 100 выстрелов. Какова вероятность числа попаданий: а) не меньше 20; б) не больше 75; в) от 45 до 75?

№ 34. Вероятность наступления события в каждом отдельном испытании равна 0,9. Сколько надо произвести испытаний, чтобы с вероятностью 0,6826 можно было ожидать, что относительная частота появления события отклонится от его вероятности не более чем на 0,03.

№ 35. Вероятность появления события в каждом из независимых испытаний равна 0,8. Найти вероятность того, что в 100 испытаниях событие появится не менее 70 и не более 80 раз.

№ 36. Вероятность появления события в каждом из независимых испытаний равна 0,8. Найти вероятность того, что событие наступит 120 раз в 144 испытаниях.

№ 37. Вероятность попадания в цель при I выстреле равна 0,2. Определить вероятность поражения цели при 6 выстрелах, если для этого нужно не менее 2 попаданий.

№ 38. Вероятность того, что студент А решит задачу, равна 0,75, а вероятность того, что студент В решит ту же задачу, равна 0,80. Найти вероятность того, что задача будет решена, если оба студента будут решать ее независимо друг от друга.

№ 39. Для изготовления детали необходимо три основных операции. Вероятность брака на первой операции равна 0,01, на второй – 0,02, на третьей – 0,03. Предполагая появление брака на отдельных операциях событиями независимыми, определить вероятность изготовления стандартной детали.

№ 40. Наборщик пользуется двумя кассами. В первой кассе *90%*, а во второй *80%* отличного шрифта. Найти вероятность того, что наудачу извлеченная литера из наудачу взятой кассы будет отличного качества.

№41. Какова должна быть вероятность изготовления изделия удовлетворяющего стандарту, чтобы с вероятностью, равной 0,90 можно было бы утверждать, что среди 20 изготовленных изделий хотя бы одно не удовлетворяет стандарту.

№ 42. Вероятность того, что книга имеется в фондах первой библиотеки равна 0,5, во второй – 0,7, и в третьей – 0,4. Определить вероятность наличия книги в фондах хотя бы одной библиотеки.

№ 43. Вероятность попадания авиабомбы в цель равна 0,2.Найти вероятность поражения цели тремя бомбами, если 2% сброшенных бомб не взрываются.

№ 44. Минное заграждение поставлено в 4 линии. Вероятность подрыва корабля, идущего без мер предосторожности, на первой ли ИИ равна 0,60, на второй – 0,73, на третьей – 0,70, на четвертой – 0,65. Найти вероятность подрыва корабля при форсировании минного поля. Какова вероятность при прохождении четырех линий подорваться только на одной из них.

№ 45. Прибор состоит из 4 узлов, вероятность безотказной работы в течение смены для каждого узла равна 0,8. Узлы выходят из строя независимо один от другого. Найти вероятность того, что за смену откажут: а) два узла; б) не менее двух узлов.

№46. Средний диаметр стволов деревьев на некоторой делянке равен 35 см., среднее квадратическое отклонение 5 см. Считая, что диаметр ствола – случайная величина, распределенная нормально, определить: а) процент стволов, имеющих диаметр свыше 30 см; б) размер, который не превзойдет диаметр ствола дерева с вероятностью 0,95.

№ 47. В каждой из двух урн содержится 2 черных и 8 белых шаров. Из первой урны наудачу извлечен шар и переложен во вторую урну, после чего из второй урны наудачу извлечен шар. Найти вероятность того, что шар, извлеченный из второй урны окажется белым.

№48. Всхожесть семян равна 90%. Для опыта отбирается 5 семян. Определить вероятность того, что будет не менее 4-х всходов.

№49. Вероятность выживания бактерий после радиоактивного облучения равна 0,004. Найти вероятность того, что после облучения из 500 бактерий останется не менее 3 бактерий.

№ 50. На участке две бригады. Вероятность выполнения плана первой бригадой равна 0,8, а вероятность выполнения плана второй – 0,9. Требуется найти: а) вероятность выполнения плана участком; б) вероятность выполнения плана только одной бригадой участка; в) вероятность выполнения плана хотя бы одной бригадой участка.

ЗАДАЧА № 51-60

Известна плотность распределения величины

$f(x) = \frac{2}{a}(1 - \frac{x}{a})$ при $x \in (0, a)$

$f(x) = 0$ при $x \notin (0, a)$

Требуется найти:

а/ функцию распределения F(x); б) вероятность попадания величины x в интервал $(\frac{a}{2}, a)$;

в) характеристики случайной величины x: m_x, D_x, G_x.

Постройте график функций f((x) и F(x). Значение параметра “a” даны в таблице

номера задач	№51	№52	№53	№54	№55	№56	№57	№58	№59	№60
Значение параметра	6	3	12	4	9	8	10	5	2	1

Задачи №61-70

Известна плотность распределения величины

$f(x) = \frac{b}{a}\sqrt{a^2 - x^2}$ при $x \in (-a, a)$

$f(x) = 0$ при $x \notin (-a, a)$

Найти: I) параметр “b”; 2) функцию распределения F(x); 3) математическое ожидание, дисперсию, среднее квадратическое отклонение величины X ; 4) построить график функции F(x)- плотности распределения.

Значение “a” дано в таблице

Номера задач	№61	№62	№63	№64	№65	№66	№67	№68	№69	№70
	2	3	5	4	6	3,2	1,6	1,4	1,8	2,2

ЗАДАЧИ № 71-80

Найти математическое ожидание, дисперсию (двумя способами), среднее квадратическое отклонение дискретной случайной величины x, если известен ее ряд распределения:

№71	x	12,0	14	20	32	35
	p	0,1	0,25	0,15	0,4	0,1
№72	x	42	46	53	60	72
	p	0,2	0,15	0,3	0,25	0,1
№73	x	65	72	88	94	100
	p	0,12	0,3	0,25	0,15	0,18
№74	x	8	10	14	18	20
	p	0,1	0,15	0,25	0,3	0,2
№75	x	24	30	32	40	45
	p	0,15	0,2	0,3	0,25	0,1
№76	x	100	120	130	150	180
	p	0,1	0,3	0,25	0,2	0,15
№77	x	4	8	10	16	20
	p	0,15	0,2	0,3	0,25	0,1
№78	x	34	40	42	48	50
	p	0,12	0,15	0,3	0,25	0,18
№79	x	60	64	70	72	76
	p	0,08	0,22	0,3	0,25	0,15
№80	x	80	82	86	88	92
	p	0,18	0,2	0,3	0,22	0,1

ЗАДАЧИ №81-90

Функциями распределения (интегральная функция) имеет вид

$f(x)=0$ при $x \le a$

$f(x)=1-\frac{a^3}{x^3}$ при $x > a$

Найти: а) плотность распределения, (дифференциальную функцию); б) математическое ожидание и дисперсию; в) вероятность падания x в интеграл (α,β).
Значения параметра “a” и интервал даны в таблице

номера задач	“a”	(α,β).
№81	2,4	(3;4)
№82	3,0	()3;5
№83	4,2	(5;6)
№84	5,0	(6;7)
№85	6,0	(7;8)
№86	8,0	(8;9)
№87	1,2	(2;3)
№88	10	(10;11)
№89	1,4	(2;4)
№90	2,5	(3;4)

Задача № 91-100

В ОТК завода были изменены диаметры 100 втулок в мм (результаты приведены в таблице). Составьте статистический ряд, постройте полигон; найдите статистическую функцию распределения, начертите её график; вычислите несмещенные оценки математического ожидания и дисперсии.

Таблица1 (для варианта 1,3,5,7,9)

21	24	24	22	23	23	24	23	26	23
24	25	25	24	24	25	26	27	25	25
25	24	26	26	25	26	24	26	25	26
24	23	24	26	23	21	25	28	24	27
26	22	22	24	24	25	24	25	25	25
25	24	25	26	25	24	23	23	26	23
24	25	26	27	27	24	24	27	25	25
26	25	24	26	25	28	25	26	26	22
24	24	25	28	22	21	24	25	25	26
26	25	24	26	25	24	25	23	25	23

Таблица2 (для вариантов 0,2,4,6,8)

25	28	22	21	25	26	25	29	24	25
24	24	24	27	25	24	27	26	22	26
26	23	26	24	25	24	26	27	24	24
22	24	22	29	25	26	24	25	25	24
26	22	25	25	24	27	22	26	24	26
26	24	24	26	25	26	26	24	23	24
25	25	25	27	24	27	21ё	25	27	25
26	25	24	24	25	28	26	27	24	24
25	27	26	26	27	25	24	25	25	27
24	25	27	25	25	28	23	26	26	24

Задача № 91-100

По данным n независимых равноточных изменений некоторой величины среднее арифметическое результатов измерений $\overline{x_{â}}$ и исправленное среднее квадратическое отключение S. Оценить истинное значение “a” измеренной величины с надёжностью γ .

Значения $n, \overline{x_{\hat{a}}}, S, \gamma$ даны в таблице

Номера задач	n	$\overline{x_{\hat{a}}}$	S	γ
№ 91	16	42,8	4,0	0,999
№92	25	31,2	3,0	0,95
№93	49	25,4	2,0	0,996
№94	64	14,2	3,5	0,993
№95	81	27,5	1,6	0,9973
№96	100	38,4	3,2	0,9876
№97	25	27,2	2,4	0,9975
№98	49	34,8	1,6	0,9962
№99	16	41,4	2,2	0,968
№100	64	40,6	3,1	0,9876

Задача № 101-110

№ 101 Найти выборочное уравнение прямой:

$$\overline{y_x} - \overline{y} = r\frac{G_y}{G_x}(x - \overline{x})$$

Регрессия y и x по данной корреляционной таблице

ABBYY PDF Transformer 2.0
Click here to buy
www.ABBYY.com

x / y	15	20	25	30	35	40	n_x
30	3	3	-	-	-	-	6
40	-	5	4	-	-	-	9
50	-	-	8	40	2	-	50
60	-	-	5	10	6	-	21
70	-	-	-	4	7	3	14
n_y	3	8	17	54	15	3	n=100

№ 102 Найти выборочное уравнение прямой:

$$\overline{y_x} - \overline{y} = r\frac{G_y}{G_x}(x - \overline{x})$$

Регрессия y и x по данной корреляционной таблице

y / x	5	10	15	20	25	30	n_x
15	2	4	-	-	-	-	6
25	-	3	7	-	-	-	10
35	-	-	5	30	10	-	45
45	-	-	7	10	8	-	25
55	-	-	-	5	6	3	14
n_y	2	7	19	45	24	3	n=100

№ 103 Найти выборочное уравнение прямой:

$$\overline{y_x} - \overline{y} = r\frac{G_y}{G_x}(x - \overline{x})$$

Регрессия y и x по данной корреляционной таблице

y / x	10	15	20	25	30	35	n_x
15	1	4	-	-	-	-	5
25	-	7	3	-	-	-	10
35	-	-	2	50	2	-	54
45	-	-	1	10	6	-	17
55	-	-	-	4	7	3	14
n_y	1	2	6	64	15	3	n=100

№ 104 Найти выборочное уравнение прямой:

$$\overline{y_x} - \overline{y} = r\frac{G_y}{G_x}(x - \overline{x})$$

Регрессия y и x по данной корреляционной таблице

y / x	15	20	25	30	35	40	n_x
30	3	3	-	-	-	-	6
40	-	5	4	-	-	-	9
50	-	-	8	40	2	-	50
60	-	-	5	10	6	-	21
70	-	-	-	4	7	3	14
n_y	3	8	17	54	15	3	n=100

№ 105 Найти выборочное уравнение прямой:

$$\overline{y_x}-\overline{y}=r\frac{G_y}{G_x}(x-\overline{x})$$

Регрессия у и х по данной корреляционной таблице

y / x	10	15	20	25	30	35	n_x
35	3	5	-	-	-	-	4
45	-	4	2	-	-	-	8
55	-	-	5	40	5	-	50
65	-	-	2	8	7	2	19
75	-	-	-	4	7	8	19
n_y	3	7	9	52	19	10	n=100

№ 106 Найти выборочное уравнение прямой:

$$\overline{y_x}-\overline{y}=r\frac{G_y}{G_x}(x-\overline{x})$$

Регрессия у и х по данной корреляционной таблице

y / x	5	10	15	20	25	30	n_x
20	3	4	-	-	-	-	8
30	-	5	-	-	-	-	8
40	-	-	7	35	8	-	50
50	-	-	2	10	8	-	20
60	-	-		5	6	3	14
n_y	3	9	13	50	22	3	n=100

№ 107 Найти выборочное уравнение прямой:

$$\overline{y_x}-\overline{y}=r\frac{G_y}{G_x}(x-\overline{x})$$

Регрессия у и х по данной корреляционной таблице

y / x	5	10	15	20	25	30	n_x
30	1	5	-	-	-	-	6
40	-	5	3	-	-	-	8
50	-	-	9	40	2	-	51
60	-	-	4	11	6	-	21
70	-	-		4	7	3	14
n_y	1	10	16	55	15	3	n=100

№ 108 Найти выборочное уравнение прямой:

$$\overline{x_y}-\overline{x}=r\frac{G_x}{G_y}(y-\overline{y})$$

Регрессия у и х по данной корреляционной таблице

y / x	10	15	20	25	30	35	n_x
20	1	5	-	-	-	-	6

30	-	6	4	-	-	-	10
40	-	-	7	40	3	-	50
50	-	-	2	10	8	-	20
60	-	-	-	5	6	3	14
n_y	10	11	13	55	17	3	n=100

№ 109 Найти выборочное уравнение прямой:

$$\overline{x_y} - \overline{x} = r\frac{G_x}{G_y}(y - \overline{y})$$

Регрессия y и x по данной корреляционной таблице

$\frac{y}{x}$	10	15	20	25	30	35	n_x
20	1	5	-	-	-	-	6
30	-	6	4	-	-	-	10
40	-	-	7	40	3	-	50
50	-	-	2	10	8	-	20
60	-	-	-	5	6	3	14
n_y	1	11	13	55	17	3	n=100

Найти выборочное уравнение прямой:

$$\overline{x_y} - \overline{x} = r\frac{G_x}{G_y}(y - \overline{y})$$

Регрессия y и x по данной корреляционной таблице

$\frac{y}{x}$	5	10	15	20	25	30	n_x
30	1	5	-	-	-	-	6
40	-	5	3	-	-	-	8
50	-	-	9	40	2	-	51
60	-	-	4	11	6	-	21
70	-	-		4	7	3	14
n_y	1	10	16	55	15	3	n=1000

№ 110 Найти выборочное уравнение прямой:

$$\overline{x_y} - \overline{x} = r\frac{G_x}{G_y}(y - \overline{y})$$

Регрессия y и x по данной корреляционной таблице

$\frac{y}{x}$	5	10	15	20	25	30	n_x
15	2	2	-	-	-	-	6
25	-	2	2	-	-	-	8
35	-	-	3	50	2	-	55
45	-	-	1	10	6	-	17

ABBYY PDF Transformer 2.0
Click here to buy
www.ABBYY.com

55	-	-		4	7	3	14
n_y	2	10	6	64	15	3	n=100

X. Правила выполнения контрольных заданий

При выполнении работ надо придерживаться указанных ниже правил. Работы, выполненные без соблюдения этих правил, не зачитываются и возвращаются студенту для переработки.

1.Контрольные задания следует выполнить в тетради (отдельной для каждой работы) чернилами любого цвета, кроме красного, оставляя поля для замечаний рецензента.

2. В заголовке работы должны быть ясно написаны фамилия студента, его инициалы, учебный номер (шифр), номер контрольной работы. Заголовок работы надо поместит на обложке тетради. Для студентов заочной формы обучения следует указать дату отсылки работы в университет и почтовый адрес студента.

3. Решения задач следует располагать в порядке номеров, указанных в заданиях, сохраняя номера задач.

4. Перед решением каждой задачи надо выписать полностью ее условие. В том случае, когда несколько задач имеют общую формулировку, следует, переписывая условие задачи, заменить общие данные контрольными из соответствующего номера.

5. Решения задач излагать подробно и аккуратно, объясняя все действия.

6. Контрольные задания, выполненные не по своему варианту не зачитываются.

Студент решает задачи варианта, номер которого совпадает с последней цифрой его учебного номера (шифра) по таблице №1, если предпоследняя цифра шифра нечетная и по таблице №2, если – четная.

Таблица № 1

Варианты	Номер задачи					
1	1	17	41	71	91	101
2	2	18	42	72	92	102
3	3	19	43	73	93	103
4	4	20	44	74	94	104
5	5	21	45	75	95	105
6	6	26	46	76	86	96
7	7	27	47	77	87	97
8	8	28	48	78	88	98
9	9	29	49	79	89	99
0	10	30	50	80	90	100

Таблица № 2

Варианты	Номер задачи					
1	11	31	51	61	81	111
2	12	32	52	62	82	112
3	13	33	53	63	83	113
4	14	34	54	64	84	114
5	15	35	55	65	85	115

ABBYY PDF Transformer 2.0 Click here to buy www.ABBYY.com

6	16	36	56	66	106	116
7	22	37	57	67	107	117
8	23	38	58	68	108	118
9	24	39	59	69	109	119
0	25	40	60	70	110	120

Л И Т Е Р А Т У Р А

7. Е.С. Вентцель Теория вероятностей 2003.
8. В.Е. Гмурман: Теория вероятностей и математическая статистика. 2001.
9. В.Ф. Чудесенко: Сборник заданий по специальным курсам высшей математики (типовые расчеты) 1983. (1998)
10. Сборник задач по математике для ВТУЗов в 4-х частях, под редакцией Б.П. Демидовича, А.В. Ефимова. Часть 3, 1990.
11. Л.З. Румшиский Элементы теории вероятностей. М.- Наука, 1970.
12. Б.В. Гнеденко Курс теории вероятностей.- М.: Физматгиз, 1988.-406с.

Дополнительная

13. Е.С. Вентцель, Л.А. Овчаров Теория вероятностей и ее инженерные приложения. 2003.
14. В.Е. Гмурман. Руководство к решению задач по теории вероятностей и математической статистике. 2002.
15. П.Е. Данко, А.Г. Попов, Т.Я. Кожевникова: Высшая математика в упражнениях и задачах. Часть 2; 2002.
10.В.Н. Тутубалин. Теория вероятностей и случайных процессов. 1992.

Справочная

11. Сигорский В.П. Математический аппарат инженера 1975.
12.И.Н. Бронштейн, К.А. Семендяев Справочник по математике для инженеров и учащихся втузов. 1986.
13.Г. Корн, Т. Корн Справочник по математике для научных работников и инженеров. 1984.

СОДЕРЖАНИЕ

V. ЗАКОН БОЛЬШИХ ЧИСЕЛ

VI. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА