ВВЕДЕНИЕ

С развитием общества экономическая система все более усложняется, усиливается статистический характер законов, описывающих социально-экономические явления. Все это предопределяет необходимость овладения методами теории вероятностей и математической статистики, как инструментом статистического анализа и прогнозирования экономических явлений и процессов.

Настоящее пособие посвящено одному из разделов классического курса высшей математики – теории вероятностей и математической статистики. В настоящее время имеется много литературы по данному курсу, однако, большинство учебников предназначены либо для студентов-математиков, которые изучают данный предмет продолжительное время, либо для широкого круга читателей. Основной особенностью данного пособия является стиль изложения, ориентированный на студентов экономических специальностей.

В пособии предлагаемого объема невозможно полностью осветить весь изучаемый теоретический материал, поэтому в каждом разделе приведены лишь необходимые теоретические сведения и формулы, отражающие количественную сторону или пространственные свойства реальных объектов и процессов, которые сопровождаются подробными решениями типовых задач, без чего невозможно успешное изучение математики.

Использование данного учебного пособия позволит студентам приобрести теоретические и практические навыки при самостоятельном решении задач. Предлагаемые в конце пособия 9 задач по 10 вариантов в каждой для самостоятельного решения позволяют проверить и укрепить приобретенные навыки.

Достоинство пособия состоит в том, что при наличии такого количества задач оно может быть использовано как задачник, а для студентов заочной формы обучения как раздаточный материал для выполнения контрольных работ по соответствующему разделу курса «Высшая математика».

В пособии указан список литературы, использование которой позволяет глубже изучить теоретический материал, рассмотренный в данном пособии.

1. ПРЕДМЕТ ТЕОРИИ ВЕРОЯТНОСТЕЙ

1.1. Основные понятия

Объектом изучения теории вероятностей служат события, носящие массовый и случайный характер.

Еще сравнительно недавно применение математики в большинстве естественных и технических наук ограничивалось классическими разделами математики, такими как дифференциальное исчисление, интегральное исчисление, дифференциальные уравнения. Вероятностные идеи и методы находили применение только в новых областях физики и в некоторых технических науках, например, в радиотехнике (обнаружение сигнала на фоне случайных помех). Последние десятилетия существенно изменили положение дела.

Вероятностные и статистические методы широко проникли в технические, технологические, экономические науки, главным образом в связи с развитием экспериментальной техники и необходимостью более тонкого анализа результата эксперимента.

Без изучения теории вероятностей немыслимо математическое образование современного инженера.

Событием (или случайным событием) называется всякий факт, который может произойти или не произойти при осуществлении определенного комплекса условий.

Обозначаются события A_1, A_2,…,A_n или A, B, C, D,… .

Примеры событий:

1. Попасть в цель при стрельбе.
2. Выиграть в лотерее.
3. Вынуть яблоко из корзины с фруктами.

Событие называется достоверным, если при осуществлении определенного комплекса условий оно обязательно произойдет.

Обозначается достоверное событие E.

Примеры:

1. При температуре + 20^0 по C вода находится в жидком состоянии.
2. Выиграть в беспроигрышной лотерее.

Событие называется невозможным, если при осуществлении определенного комплекса условий, оно не может произойти.

Обозначается невозможное событие U.

Пример: при температуре + 20^0 по C вода в сосуде находится в твердом состоянии.

Несколько событий в данном опыте образуют полную группу событий, если в результате опыта непременно должно появиться хотя бы одно из них.

Примеры:

1. Стрелок производит по мишени два выстрела.

Событие A_1 – 1 попадание.

Событие A_2 – 2 попадания.

Событие A_3 – 0 попаданий.

A_1, A_2, A_3 – полная группа событий.

2. При проверке одной страницы печатного текста.

Событие A_1 – 0 опечаток.

Событие A_2 – 1 опечатка.

Событие A_3 – 2 опечатки.

Событие A_4 – более 2-х опечаток.

A_1, A_2, A_3, A_4 – полная группа событий.

Несколько событий в данном опыте называются несовместными, если никакие два из них не могут появиться вместе.

Примеры:

1. При одном выстреле.

Событие A_1 – попадание в цель.

Событие A_2 – промах.

A_1, A_2– несовместны.

2. При одном бросании монеты

Событие A_1 – выпадение герба.

Событие A_2 – выпадение решки.

A_1, A_2 – несовместны.

Несколько событий называются равновозможными, если есть основание считать, что ни одно из этих событий не является более возможным, чем другие.

Пример. В ящике находится 10 одинаковых шаров, занумерованных от 1 до 10. При вынимании одного шара появление шара с любым из этих номеров есть события равновозможные.

Если несколько событий:

– образуют полную группу,

– несовместны,

– равновозможные,

то они называются случаями (шансами).

Пример. Опыт – бросание монеты.

A_1 – выпадение герба.

A_2 – выпадение решки.

A_1, A_2 – случаи.

1.2. Случайные события, их относительная частота и вероятность

Основной числовой характеристикой случайного события является его вероятность.

Будем много раз повторять испытание и каждый раз отмечать, произошло интересующее нас событие A или нет. Обозначим через N общее число произведенных испытаний, через N_A – число появлений события A за N испытаний, и назовем отношение $W_A = \frac{N_A}{N}$ – относительной частотой случайного события A.

Оказывается, что в различных сериях испытаний относительная частота случайного события A принимает близкие значения. Например, при многократном бросании правильной игральной кости относительная частота выпадения каждого числа очков от 1 до 6 всегда близка к числу $\frac{1}{6}$.

Отмеченное свойство относительных частот называют статистической устойчивостью. Таким образом, относительные частоты случайного события A в различных сериях группируются около определенного числа: это число называется вероятностью случайного события A и обозначается символом P(A).

Относительная частота случайного события A получается при практических испытаниях. Вычисляется после опыта.

Вероятность события получается из теоретических соображений, причем не требуется, чтобы испытания производились фактически. Вычисляют до опыта.

Случай называется благоприятным (или благоприятствующим) событию A, если появление этого случая влечет за собой событие A.

Пример. При бросании игральной кости возможно 6 случаев: появление 1, 2, 3, 4, 5, 6 очков.

Событие A – появление четного числа очков. Событию A благоприятствует 3 случая: появление 2, 4, 6 очков и не благоприятствуют остальные 3. Обозначим m – число случаев, благоприятных появлению события A.

В данном примере m = 3.

(Классическое определение вероятности). Если результаты опыта сводятся к схеме случаев, то вероятностью события A называется число $\frac{m}{n}$, где m – число случаев благоприятствующих событию A.

n – общее число случаев.

Обозначается $P(A)=\frac{m}{n}$; в дальнейшем будем считать, что имеет место схема случаев.

1.3. Аксиомы вероятностей события А

1. Вероятность достоверного события равна 1.
2. Вероятность невозможного события равна 0.
3. Вероятность случайного события есть положительное число, заключенное между нулем и единицей.

Итак, вероятность любого события А:

$$0 \le P \le 1.$$

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

где А, В – несовместные события.

Задача. Набирая номер телефона, абонент забыл последнюю цифру и набрал ее наудачу. Найти вероятность того, что набрана нужная цифра.

Решение. Событие А – набрана нужная цифра.

n = 10 (несовместны, равновозможные, образуют полную группу), где m = 1, $P(A)=\frac{1}{10}$.

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

Решение. Событие А – набраны нужные цифры.

$n=10\cdot 9=90$ (несовместны, равновозможные, образуют полную группу), m = 1, $P(A)=\frac{1}{90}$.

Задача. Монета брошена два раза. Найти вероятность того, что хотя бы один раз появится «герб».

Решение. Всего исходов 4: (г,г), (г,нг), (нг,г), (нг,нг); благоприятствующих 3: (г,г), (г,нг), (нг,г). Поэтому n = 4, m = 3, $P(A)=\frac{3}{4}$.

Задача. Брошены две игральные кости. Найти вероятность того, что сумма очков на выпавших гранях четная, причем на грани хотя бы одной из костей выпадет «6».

Решение. Всего исходов n = 36 (каждая из шести граней одной кости может сочетаться с любой из шести граней другой кости: $6\cdot 6=36$).

Благоприятствующих исходов m = 5: (6,2), (6,4), (6,6), (2,6), (4,6). Таким образом $P(A)=\frac{5}{36}$.

Задача. Среди 17 студентов группы, из которых 8 девушек, разыгрывается 7 билетов. Какова вероятность того, что среди обладателей билетов окажутся 4 девушки?

Решение. Так как всего 7 билетов, то общее число исходов $n = C_{17}^{7} = \frac{17!}{10!7!} = 19448$. Число выборов трех юношей из девяти $C_9^3 = \frac{9!}{6!3!} = \frac{7 \cdot 8 \cdot 9}{1 \cdot 2 \cdot 3} = 84$. Число выборов четырех девушек из восьми $C_8^4 = \frac{8!}{4!4!} = \frac{5 \cdot 6 \cdot 7 \cdot 8}{1 \cdot 2 \cdot 3 \cdot 4} = 70$. Комбинируя каждый способ выбора трех юношей с каждым способом выбора четырех девушек, получим $m = C_9^3 \cdot C_8^7 = 84 \cdot 70 = 5880$. Тогда $P(A) = \frac{5880}{19448} = 0{,}302$.

Суммой A + B двух событий A и B называют событие, состоящее в появлении события A или B, или обоих этих событий.

Проиллюстрируем это геометрически (рис. 1).

Пусть событие A есть попадание точки в область A.

Пусть событие B есть попадание точки в область B.

Событие A + B – попадание точки в заштрихованную область.

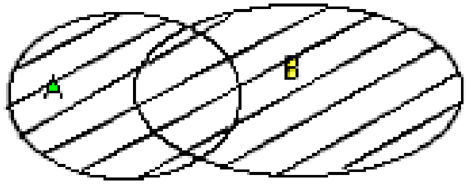


Рис. 1

Пример. Произведено 2 выстрела по цели.

Событие A – попадание при первом выстреле.

Событие B – попадание при втором выстреле.

Событие A + B – попадание при первом выстреле, или при втором, или при обоих.

Замечание. Если A и B – несовместные события, то A + B – событие, состоящее в появлении одного из этих событий, безразлично какого.

Теорема. Вероятность появления одного из нескольких попарно несовместных событий, безразлично какого, равна сумме вероятностей этих событий

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n).$$

Доказательство методом индукции.

Задача. В урне 30 шаров: 10 красных, 5 синих, 15 белых. Найти вероятность того, что при одном вынимании появится цветной шар.

Решение.

Событие A – вынимание красного шара.

Событие B – вынимание синего шара.

События A и B несовместны.

$$P(A) = \frac{10}{30} = \frac{1}{3},\ P(B) = \frac{5}{30} = \frac{1}{6},\ P(A + B) = P(A) + P(B),$$

$$P(A + B) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}.$$

Следствие 1. Если событие A_1, $A_2, \ldots, A_n$ образуют полную группу несовместных событий, то $P(A_1) + P(A_2) + \ldots + P(A_n) = 1$.

Следствие 2. Сумма вероятностей противоположных событий равна единице:

$$P(A) + P(\overline{A}) = 1.$$

Задача. Вероятность того, что день будет дождливым $p = 0{,}7$. Найти вероятность того, что день будет без осадков.

Решение.

Событие A – дождливый день.

Событие B – без осадков.

Обозначим $P(A) = p = 0{,}7$, $P(B) = q$, $q = 1 – p = 1 – 0{,}7 = 0{,}3$.

Примечание. При решении задач иногда удобно вычислять вероятность противоположного события, а затем найти вероятность данного события A по формуле

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

Задача. В партии из 10 деталей 8 стандартных. Найти вероятность того, что среди наудачу извлеченных двух деталей есть хотя бы одна стандартная.

Решение. Событие A – хотя бы одна деталь из двух извлеченных стандартная.

Событие $\overline{A}$ – обе детали нестандартные.

$$P(A) = 1 - P(\overline{A}),$$

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

$$m = 1,$$

$$n = C_{10}^{2} = \frac{10 \cdot 9}{1 \cdot 2} = 45,$$

$$P(\overline{A}) = \frac{1}{45},$$

$$P(A) = 1 - \frac{1}{45} = \frac{44}{45}.$$

1.4. Произведение событий

Произведением нескольких событий A_1, A_2, …, A_n называется событие A, состоящее в совместном появлении всех событий. Обозначение $A = A_1 \cdot A_2 \cdot ... \cdot A_n$. Иллюстрация геометрическая на рис. 2.

Событие A_1 попадание точки в область A_1

Событие A_2 попадание точки в область A_2

Событие A_3 попадание точки в область A_3.

Событие $A = A_1 \cdot A_2 \cdot A_3$ – попадание точки в выделенную область.

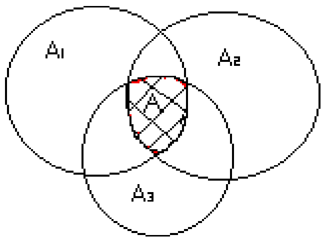


Рис. 2

Два события называются независимыми, если вероятность появления одного из них не зависит от появления или не появления другого.

Пример. Монета брошена два раза. Вероятность появления герба в первом испытании (событие A) не зависит от появления или не появления герба во втором испытании (событие B) и наоборот.

A, B – независимые события.

Несколько событий называются попарно независимыми, если каждые два из них независимы.

Задача. В ящике 100 деталей, 80 стандартных, 20 нестандартных. Наудачу берут одну деталь и не возвращают ее в ящик. Затем вынимают вторую деталь. Найти вероятность того, что при втором вынимании появилась стандартная деталь.

Решение. Событие A – появление стандартной детали при первом вынимании.

Событие B – появление стандартной детали при втором извлечении.

Если при первом вынимании появляется стандартная деталь, то

$$P(B) = \frac{79}{99}.$$

Если при первом вынимании извлекли нестандартную деталь, то

$$P(B) = \frac{80}{99}.$$

Значит вероятность события B зависит от наступления или не наступления события. A, B – зависимые события.

Условной вероятностью события B называют вероятность события, вычисленная в предположении, что событие A произошло.

Обозначение $P_A(B)$.

Задача. В урне содержится 3 белых и 3 черных шара. Из урны дважды вынимают по одному шару, не возвращая их в урну. Найти вероятность появления белого шара при втором испытании, если при первом испытании извлекли черный шар.

Решение.

Событие A – при первом испытании был извлечен черный шар.

Событие B – появление белого шара при втором испытании.

$$P_A(B) = \frac{3}{5}.$$

Теорема. Вероятность совместного появления двух событий A, B равна произведению вероятности одного из них на условную вероятность другого.

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

Несколько событий называют независимыми в совокупности (или просто независимыми), если независимы каждые два из них и независимы каждое событие и всевозможные произведения остальных.

Задача. В читальном зале имеется 6 учебников по теории вероятностей, из которых 3 в переплете. Библиотекарь взял наудачу 2 учебника. Найти вероятность того, что оба учебника окажутся в переплете.

Решение.

Событие A – и первый взятый учебник в переплете.

Событие B – и второй взятый учебник в переплете.

$$P(A \cdot B) = P(A) \cdot P_A(B),$$

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

$$P_A(B) = \frac{2}{5},$$

$$P(A \cdot B) = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{5} = 0{,}2.$$

Задача. В цехе работают 7 мужчин и 3 женщины. По табельным номерам наудачу отобраны 3 человека. Найти вероятность того, что все отобранные лица окажутся мужчинами.

Решение.

Событие А – и первым отобран мужчина.

Событие В – и вторым отобран мужчина.

Событие С – и третьим отобран мужчина.

$$P(A) = \frac{7}{10}, \; P_A(B) = \frac{6}{9}, \; P_{AB}(C) = \frac{5}{8},$$

$$P(A \cdot B \cdot C) = P(A) \cdot P_A(B) \cdot P_{AB}(C),$$

$$P(ABC) = \frac{7}{10} \cdot \frac{6}{9} \cdot \frac{5}{8} = \frac{7}{24}.$$

Теорема. Вероятность появления хотя бы одного из событий $A_1, A_2, \ldots, A_n$, независимых в совокупности, равна разности между единицей и произведением вероятностей противоположных событий, $P(A) = 1 - q_1 \cdot q_2 \cdot \ldots \cdot q_n$, где $P(\overline{A}_1) = q_1, \ldots, P(\overline{A}_n) = q_n$.

Замечание. Если $P(\overline{A}_1) = P(\overline{A}_2) = \ldots = P(\overline{A}_n) = q$, то $P(A) = 1 - q^n$.

Задача. Вероятности попадания в цель при стрельбе из трех орудий таковы: $p_1 = 0{,}8$, $p_2 = 0{,}7$, $p_3 = 0{,}9$. Найти вероятность хотя бы одного попадания при одном залпе из всех орудий.

Решение. Событие А – хотя бы одно попадание.

$$P(A) = 1 - q_1 \cdot q_2 \cdot q_3,$$

$$q_1 = 0{,}2, \; q_2 = 0{,}3, \; q_3 = 0{,}1.$$

$$P(A) = 1 - 0{,}2 \cdot 0{,}3 \cdot 0{,}1 = 0{,}994.$$

Теорема. Вероятность появления хотя бы одного из двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления.

$$P(A + B) = P(A) + P(B) - P(A \cdot B).$$

Замечание. События А, В могут быть как независимыми, так и зависимыми.

Задача. Вероятности попадания в цель при стрельбе первого и второго орудий соответственно равны: $p_1 = 0{,}7$, $p_2 = 0{,}8$. Найти вероятность попадания при одном залпе хотя бы одного из орудий.

Решение.

$$P(A \cdot B) = P(A) \cdot P(B) = 0{,}7 \cdot 0{,}8 = 0{,}56 .$$

A, B – независимы.

$$P(A + B) = 0{,}7 + 0{,}8 – 0{,}56 = 0{,}94.$$

Задача. Для разрушения моста достаточно попадания одной авиационной бомбы. Найти вероятность того, что мост будет разрушен, если на него сбросить четыре бомбы. Вероятности попадания бомб соответственно равны: 0,3; 0,4; 0,6; 0,7.

Решение. Введем события:

A_1 – попадание первой бомбы.

A_2 – попадание второй бомбы.

A_3 – попадание третьей бомбы.

A_4 – попадание четвертой бомбы.

M – мост разрушен.

$$\begin{aligned} P(M) &= 1 - P(\overline{A}_1) \cdot P(\overline{A}_2) \cdot P(\overline{A}_3) \cdot P(\overline{A}_4) = \\ &= 1 - 0{,}7 \cdot 0{,}6 \cdot 0{,}4 \cdot 0{,}3 = 0{,}9496 \end{aligned}.$$

Теорема. Если события B_1, $B_2,\ldots,B_n$ – образуют полную группу несовместных событий (гипотез) и событие A может произойти вместе с одним из этих событий, то

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) + ... + P(B_n) \cdot P_{B_n}(A) .$$

Задача. В первой коробке содержится 20 радиоламп, из них 18 стандартных, во второй коробке – 10 ламп, из них 9 стандартных. Из второй коробки наудачу взята лампа и положена в первую. Найти вероятность того, что лампа, наудачу извлеченная из первой коробки, будет стандартной.

Решение.

Событие A – из первой коробки извлечена стандартная лампа.

Событие B_1 – из второй коробки извлечена стандартная лампа.

Событие B_2 – из второй коробки извлечена нестандартная лампа.

A произойдет либо с B_1, либо с B_2.

$$P(A) = P(B_1) \cdot P_{B_1}(A) + P(B_2) \cdot P_{B_2}(A) .$$

$$P(B_1) = \frac{9}{10},\ P(B_2) = \frac{1}{10},\ P_{B_1}(A) = \frac{19}{21},\ P_{B_2}(A) = \frac{18}{21} .$$

$$P(A) = \frac{9}{10} \cdot \frac{19}{21} + \frac{1}{10} \cdot \frac{18}{21} = 0{,}9 .$$

Задача. На одном заводе на каждые 100 лампочек приходится в среднем 10 нестандартных, на втором – 15, а на третьем – 20. Продукция этих заводов составляет соответственно 50, 30, 20% всех электролампочек, приобретаемых жителями района. Найти вероятность того, что приобретенная электролампочка будет стандартной.

Решение.

Событие А – приобретенная лампочка стандартная.

B_1 – лампочка изготовлена первым заводом.

B_2 – лампочка изготовлена вторым заводом.

B_3 – лампочка изготовлена третьим заводом.

$$P(B_1) = 0,5,\ P(B^2) = 0,3,\ P(B_3) = 0,2,$$

независимо от того стандартная она или нет.

$$P_{B1}(A) = 0,9,\ P_{B2}(A) = 0,85,\ P_{B3}(A) = 0,8.$$

$$P(A) = P(B_1)\cdot P_{B_1}(A) + P(B_2)\cdot P_{B_2}(A) + P(B_3)\cdot P_{B_3}(A).$$

$$P(A) = 0,5\cdot 0,9 + 0,3\cdot 0,85 + 0,2\cdot 0,8 = 0,45 + 0,255 + 0,16 = 0,865.$$

Задача. Охотник сделал три выстрела по кабану. Вероятность попадания первым выстрелом 0,4, вторым – 0,5, третьим – 0,7. Одним попаданием кабана можно убить с вероятностью 0,2, двумя попаданиями – с вероятностью 0,6, а тремя – наверняка. Найти вероятность того, что кабан будет убит.

Решение. Введем события: А – кабан убит, A_1 – попадание первым выстрелом, A_2 – попадание вторым выстрелом, A_3 – попадание третьим выстрелом.

Возможны следующие гипотезы, образующие полную группу несовместных событий: H_0 – три промаха, H_1 – одно попадание, H_2 – два попадания, H_3 – три попадания.

По формуле полной вероятности, вероятность того, что кабан будет убит, равна:

$$P(A) = P(H_0)P(A/H_0) + P(H_1)P(A/H_1) + P(H_2)P(A/H_2) +$$
$$+ P(H_3)P(A/H_3).$$

Нам известны условные вероятности

$$P(A/H) = 0,\ P(A/H_1) = 0,2,\ P(A/H_2) = 0,6,\ P(A/H_3) = 1$$

Найдем вероятности гипотез $P(H_0)$. $P(H_1)$, $P(H_2)$, $P(H_3)$.

Так как $H_0 = \overline{A}_1\overline{A}_2\overline{A}_3$, $H_1 = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$, $H_2 = A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3$, $H_3 = A_1A_2A_3$, события A_1, A_2, A_3 – независимы, $\overline{A}_1$, $\overline{A}_2$, $\overline{A}_3$ – независимы, то

$$P(H_0) = P(\overline{A}_1)P(\overline{A}_2)P(\overline{A}_3),$$

$$P(H_1) = P(A_1)P(\overline{A}_2)P(\overline{A}_3) + P(\overline{A}_1)P(A_2)P(\overline{A}_3) + P(\overline{A}_1)P(\overline{A}_2)P(A_3),$$

$$P(H_2)=P(A_1)P(A_2)P(\overline{A}_3)+P(A_1)P(\overline{A}_2)P(A_3)+P(\overline{A}_1)P(A_2)P(A_3),$$

$$P(H_3)=P(A_1)P(A_2)P(A_3).$$

По условию $P(A_1) = 0,4$, $P(A_2) = 0,5$, $P(A_3) = 0,7$. Следовательно,

$$P(H_0)=0,6\cdot 0,5\cdot 0,3=0,09,$$

$$P(H_1)=0,4\cdot 0,5\cdot 0,3+0,6\cdot 0,5\cdot 0,3+0,6\cdot 0,5\cdot 0,7=0,36,$$

$$P(H_2)=0,4\cdot 0,5\cdot 0,3+0,6\cdot 0,5\cdot 0,7+0,4\cdot 0,5\cdot 0,7=0,41,$$

$$P(H_3)=0,4\cdot 0,5\cdot 0,7=0,14.$$

Контроль: 0,09+0,36+0,41+0,14=1.

$$P(\grave{A})=0,09\cdot 0+0,36\cdot 0,2+0,41\cdot 0,6+0,14\cdot 1=0,458.$$

Теорема (формулы Байеса).

Если $B_1, B_2,\ldots,B_n$ – образуют полную группу несовместных событий (гипотез) и произведено испытание, в результате которого произошло событие A, то вероятность гипотезы после испытания равна:

$$P_A(B_i)=\frac{P(B_i)\cdot P_{B_i}(A)}{P(A)}.$$

Задача. Из 10 студентов, пришедших на экзамен, 3 подготовлено отлично, 4 – хорошо, 2 – посредственно и 1 – 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 вопросов. Вызванный наугад студент ответил на три произвольно заданных вопроса. Найти вероятность того, что этот студент подготовлен отлично.

Решение. Введем события: A – студент ответил на три заданных вопроса, H_1 – студент подготовлен отлично, H_2 – студент подготовлен хорошо, H_3 – студент подготовлен посредственно, H_4 – студент подготовлен плохо. Эти гипотезы образуют полную группу несовместных событий.

$$P(H_1)=\frac{3}{10},\ P(H_2)=\frac{4}{10},\ P(H_3)=\frac{2}{10},\ P(H_4)=\frac{1}{10}.$$

Контроль: 0,3+0,4+0,2+0,1=1.

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

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

$$P(A)=0,3\cdot 1+0,4\cdot 0,491+0,2\cdot 0,105+0,1\cdot 0,009=0,518.$$

Искомая вероятность $P(H_1/A) = \dfrac{0{,}3 \cdot 1}{0{,}518} = 0{,}58$.

Задача. Имеется две партии деталей, причем известно, что в одной партии все детали удовлетворяют техническим условиям, а в другой 1/4 деталей недоброкачественные. Деталь, взятая из наудачу выбранной партии, оказалась доброкачественной. Определить вероятность того, что вторая деталь из этой же партии окажется недоброкачественной, если первая деталь после проверки возвращена в партию.

Решение. Гипотезы H_1 – взята партия с недоброкачественными деталями, H_2 – взята партия доброкачественных деталей, A – первая деталь доброкачественная.

$$P(H_1) = P(H_2) = 0{,}5.$$

Контроль: $0{,}5 + 0{,}5 = 1$. $P(A/H_1) = \dfrac{3}{4}$, $P(A/H_2) = 1$.

По формуле полной вероятности $P(A) = \dfrac{1}{2} \cdot \dfrac{3}{4} + \dfrac{1}{2} \cdot 1 = \dfrac{7}{8}$.

После первого испытания вероятность того, что партия содержит недоброкачественные детали, равна:

$$P(H_1/A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{\frac{1}{2} \cdot \frac{3}{4}}{\frac{7}{8}} = \frac{3}{7}.$$

Вероятность того, что партия содержит только доброкачественные детали, равна:

$$P(H_2/A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{\frac{1}{2} \cdot 1}{\frac{7}{8}} = \frac{4}{7}.$$

Пусть событие B – при втором испытании деталь оказалась недоброкачественная. Вероятности гипотез H_1 и H_2 после первого испытания $P(H_1) = \dfrac{3}{7}$, $P(H_2) = \dfrac{4}{7}$.

Контроль $\dfrac{3}{7} + \dfrac{4}{7} = 1$.

Кроме того, $P(B/H_1) = \dfrac{1}{4}$, $P(B/H_2) = 0$.

Искомую вероятность находим по формуле полной вероятности:

$$P(B) = P(H_1)P(B/H_1) + P(H_2)P(B/H_2) = \frac{3}{7} \cdot \frac{1}{4} + \frac{4}{7} \cdot 0 = \frac{3}{28}.$$

Задача. В трех ящиках находятся однотипные изделия: в первом 10 изделий, из них 3 нестандартных, во втором 15 изделий, из них 5 нестандартных, а в третьем 20 изделий, из них 6 нестандартных. Наудачу вытащили одно изделие, и оно оказалось нестандартное. Определить вероятность того, что взятое изделие принадлежало второму ящику.

Решение.

A – нестандартное изделие вынуто из любого ящика.

B_1 – вынуто изделие из первого ящика.

B_2 – вынуто изделие из второго ящика.

B_3 – вынуто изделие из третьего ящика.

$$P(B_1) = P(B_2) = P(B_3) = \frac{1}{3}.$$

$P_A(B_2)$ – вероятность того, что взятое нестандартное изделие принадлежит второму ящику.

$$P_A(B_2) = \frac{P(B_2) \cdot P_{B_2}(A)}{P(A)},$$

$$P_{B_1}(A) = \frac{3}{10}, \; P_{B_2}(A) = \frac{1}{3}, \; P_{B_3}(A) = \frac{3}{10}.$$

$$P(A) = P(B_1)P_{B_1}(A) + P(B_2)P_{B_2}(A) + P(B_3)P_{B_3}(A).$$

$$P(A) = \frac{1}{3} \cdot \frac{3}{10} + \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{3}{10} = \frac{1}{3} \cdot \left(\frac{3}{5} + \frac{1}{3}\right) = \frac{1}{3} \cdot \frac{15}{14}.$$

$$P_A(H_2) = \frac{\frac{1}{3} \cdot \frac{1}{3}}{\frac{1}{3} \cdot \frac{14}{15}} = \frac{1}{3} \cdot \frac{15}{14} = \frac{5}{14}.$$

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

Вероятность того, что при n независимых испытаниях события A осуществится ровно k раз и, следовательно, не осуществится n–k раз вычисляется по формуле Бернулли:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k},$$

$$P_n(k) = \frac{n!}{k!(n-k)!} \cdot p^k \cdot q^{n-k},$$

где $n \le 10$, $p \ge 0{,}1$.

Задача. В урне 30 шаров: 20 белых и 10 черных. Вынули подряд четыре шара, причем каждый вынутый шар возвращается в урну перед извлечением следующего, и шары в урне перемешиваются. Какова вероятность того, что из четырех вынутых шаров будут два белых?

Решение. Вероятность вынуть белый шар в каждом испытании $\frac{20}{30}=\frac{2}{3}$, черный – $\frac{10}{30}=\frac{1}{3}$.

Так как $n = 4$, $k = 2$, тогда по формуле Бернулли получим:

$$P_4(2)=C_4^2\cdot\left(\frac{2}{3}\right)^2\cdot\left(\frac{1}{3}\right)^2=\frac{4\cdot 3}{1\cdot 2}\cdot\frac{4}{9}\cdot\frac{1}{9}=\frac{8}{27}.$$

Задача. По данным ОТК на сотню металлических брусков, заготовленных для обработки, приходится 20 с зазубринами. Какова вероятность того, что из случайно взятых 5 брусков окажется без зазубрин: а) не более двух; б) не менее четырех?

Решение. По условию:

а) $q = 0{,}2$, $p = 1 - q = 1 - 0{,}2 = 0{,}8$, $n = 5$, $k < 2$, то есть или 2, или 1, или 0.

Искомая вероятность будет равна:

$$P_5(k\le 2)=P_5(2)+P_5(1)+P_5(0)=C_5^2\cdot 0{,}8^2\cdot 0{,}2^3+C_5^1\cdot 0{,}8\cdot 0{,}2^4+$$
$$+C_5^0\cdot 0{,}8^0\cdot 0{,}2^5=10\cdot 0{,}64\cdot 0{,}008+5\cdot 0{,}8\cdot 0{,}0016+1\cdot 1\cdot 0{,}00032=0{,}519.$$

б) $p = 0{,}8$, $q = 0{,}2$, $n = 5$, $k \ge 4$, то есть или 4, или 5.

Искомая вероятность равна:

$$P_5(k\ge 4)=P_5(4)+P_5(5)=\text{Ñ}_5^4\cdot 0{,}8^4\cdot 0{,}2^3+C_5^5\cdot 0{,}8^5\cdot 0{,}2^0=0{,}737.$$

2. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ, ИХ ЧИСЛОВЫЕ ХАРАКТЕРИСТИКИ И ЗАКОНЫ РАСПРЕДЕЛЕНИЯ

2.1. Виды случайных величин. Функция распределения, плотность распределения

Случайной называется величина, которая в результате опыта может принять одно возможное значение, причем до опыта неизвестно какое.

Случайные величины обозначаются заглавными буквами – X, Y, Z, U, V, T,…, а их возможные значения – x, y, z, u, v, t,… .

Дискретной называется случайная величина, которая принимает отдельные, изолированные возможные значения с определенными вероятностями (число их конечно или бесконечное счетное множество).

Непрерывной называется случайная величина, возможные значения которой непрерывно заполняют некоторый интервал (конечный или бесконечный).

Законом распределения случайной величины называется всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими вероятностями.

Способы и формы представления закона распределения случайной величины могут быть различными.

Простейшей формой задания закона распределения дискретной случайной величины X является таблица, в которой перечислены возможные значения случайной величины и соответствующие им вероятности.

x_i	x_1	x_2	…	x_n
$p_i = p(X = x_i)$	p_1	p_2	…	p_n

где $\sum_{i=1}^{n} p_i = 1$.

Такая таблица называется рядом распределения случайной величины.

Интегральной функцией распределения называют функцию, определяющую для каждого значения x вероятность того, что случайная величина X примет значения меньше x, т.е.

$$F(x) = P(X < x).$$

Для дискретной случайной величины X, которая может принимать значения $x_1, x_2,…,x_n$, функция распределения имеет вид:

$$F(X) = \sum_{i=1}^{n} P(X < x_i),$$

где символ x_i под знаком суммы обозначает, что суммирование распространяется на все те возможные значения величины, которые по своей величине меньше аргумента x.

Случайная величина называется непрерывной, если функция распределения F(x) непрерывна и имеет производную.

Плотностью распределения непрерывной случайной величины называется функция:

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

2.2. Свойства плотности распределения

1. Плотность распределения неотрицательна:

$$f(x) \geq 0.$$

2. Функция распределения случайной величины равна интегралу от плотности в интервале от -∞ до x, т.е.

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

3. Вероятность попадания непрерывной случайной величины на участок (a, b) равна интегралу от плотности распределения, взятому по этому участку, т.е.

$$p(a < x < b) = \int_{a}^{b} f(x)dx.$$

4. Интеграл в бесконечных пределах от плотности распределения равен единице:

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

Задача 1. Дискретная случайная величина X имеет следующий ряд:

x_i	-1	0	2	2,5
$p_i = p(X < x_i)$	0,2	0,3	0,4	0,1

Найти функцию распределения.

Решение. Значения X разбивают всю числовую ось на 5 промежутков. Рассмотрим эти промежутки по порядку.

1. При $x \leq -1$, событие $X < x$ – невозможное, $F(x) = 0$.
2. При $-1 < x \leq 0$, $F(x) = P(X = -1) = 0{,}2$.

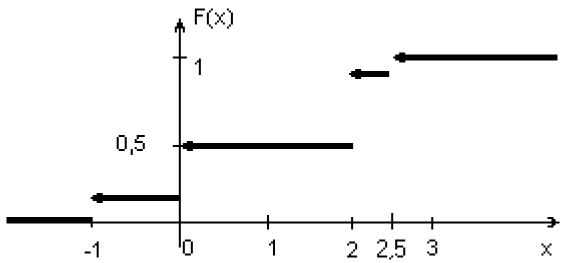
3. При $0 < x \le 2$, $F(x) = P(X = -1) + P(X = 0) = 0{,}2 + 0{,}3 = 0{,}5$.

1. При $2 < x \le 2{,}5$, $F(x) = P(X = -1) + P(X = 0) + P(X = 2) = 0{,}2 + {}$ $+ 0{,}3 + 0{,}4 = 0{,}9$

2. При $x > 2{,}5$, $F(x) = P(X = -1) + P(X = 0) + P(X = 2) + P(X = 2{,}5) =$ $= 0{,}2 + 0{,}3 + 0{,}4 + 0{,}1 = 1$.

$$\text{Итак, } F(x) = 0\begin{cases} 0 & \text{при} \quad x \le -1; \\ 0{,}2 & \text{при} \quad -1 < x \le 0; \\ 0{,}5 & \text{при} \quad 0 < x \le 2; \\ 0{,}9 & \text{при} \quad 2 < x \le 2{,}5; \\ 1 & \text{при} \quad x > 2{,}5. \end{cases}$$

График функции имеет ступенчатый вид:



Задача 2. Средняя продолжительность работы определенного типа радиоламп (в часах) имеет следующую плотность распределения:

$$f(x) = \begin{cases} \dfrac{100}{X^2}, \text{ если } x > 100; \\ 0, \text{ если } x \le 100. \end{cases}$$

Вычислить:

а) вероятность того, что лампа не будет заменена в первые 150 часов функционирования;

б) вероятность того, что лампа будет функционировать более 200 часов, и в то же время, менее 300 часов.

Решение: а) обозначим время функционирования лампы через X. Мы знаем, что P (X > 150) = 1 – P(X < 150) и что P(X < 150) = F(150).

В то же время $F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{100}^{x} \dfrac{100}{X^2}\,dx = 1 - \dfrac{100}{x}$.

Следовательно, $P(X > 150) = 1 - (1 - \dfrac{100}{150}) = \dfrac{2}{3}$.

б) $P(200 > X < 300) = \int\limits_{200}^{300} f(x)dx = \int\limits_{200}^{300} \frac{100}{x^2} dx \;\; = 1 - \int\limits_{200}^{300} \frac{100}{x} = \frac{1}{6}$.

Задача 3. Случайная величина X имеет плотность распределения:

$$f(x) = \begin{cases} A \sin \text{õ}, \text{ при} \quad 0 \le x \le \pi; \\ \quad 0, \quad \text{при} \quad x < 0, x > \pi. \end{cases}$$

Определить: а) коэффициентA; б) функцию распределения F(x); в) вероятность события $-\frac{\pi}{2} < x < \frac{2}{3}\pi$.

Решение. а) коэффициент A определяется из условия:

$$\int\limits_{-\infty}^{\infty} f(x)dx = 1, \text{ т.е. } \int\limits_{0}^{\pi} A \sin \text{õ} dx = 1, \text{ откуда } A = \frac{1}{2}.$$

б) определим функцию распределения случайной величины A по формуле: $F(x) = \int\limits_{-\infty}^{x} f(x)dx$,

1) при $x < 0$, $F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{x} 0dx = 0$;

2) при $0 \le x \le \pi$, $F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0dx + \int\limits_{0}^{x} \frac{1}{2} \sin x dx =$

$= -\frac{1}{2} \cos x \Big|_{0}^{x} = -\frac{1}{2} \cos x + \frac{1}{2}$;

3) при $x > \pi$, $F(x) = \int\limits_{-\infty}^{x} f(x)dx = \int\limits_{-\infty}^{0} 0dx + \int\limits_{0}^{\pi} \frac{1}{2} \sin x dx + \int\limits_{\pi}^{x} 0dx =$

$= -\frac{1}{2} \cos x \Big|_{0}^{\pi} = -\frac{1}{2} \cdot (-1) + \frac{1}{2} = 1$.

Итак:

$$F(x) = \begin{cases} \quad 0, \qquad \text{ïðè} \quad x < 0; \\ \frac{1}{2} - \frac{1}{2} \cos x, \text{ ïðè} \quad 0 \le x \le \pi; \\ \quad 1, \qquad \text{ïðè} \quad x > \pi. \end{cases}$$

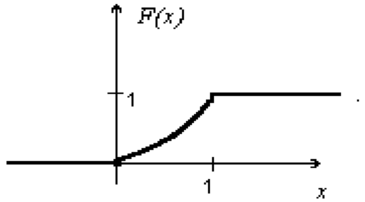
в) определим вероятность события $(-\frac{\pi}{2} < \text{õ} \le \frac{2}{3}\pi)$:

$$P(-\frac{\pi}{2} < X \le \frac{2}{3}\pi) = \int_{-\frac{\pi}{2}}^{\frac{2\pi}{3}} f(x)dx = \int_{0}^{\frac{2\pi}{3}} \frac{1}{2}\sin x dx = -\frac{1}{2}\cos x \Big|_{0}^{\frac{2\pi}{3}} = -\frac{1}{2}\cos\frac{2}{3}\pi +$$

$$+\frac{1}{2}\cos 0 = \frac{3}{4}.$$

Задача 5. Случайная величина X задана функцией распределения:

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



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

Решение. P(a < x < b) = F(b) – F(a).

$$P(0{,}25 < x < 0{,}75) = X^2{}_{x=0,75} - X^2{}_{x=0,25} = \left(\frac{3}{4}\right)^2 - \left(\frac{1}{4}\right)^2 = \frac{1}{2},$$

$$p = \frac{1}{2}; \quad q = 1 - \frac{1}{2} = \frac{1}{2}.$$

Искомую вероятность находим по формуле Бернулли:

$$P_n^k(A) = C_n^k p^k q^{n-k}.$$

$$P_4^3 = C_4^3 \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^{4-3} = 4 \cdot \left(\frac{1}{2}\right)^4 = 0{,}25.$$

Ответ: P = 0,25.

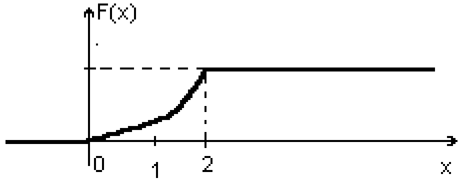
Задача. Случайная величина задана интегральной функцией (функцией распределения):

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

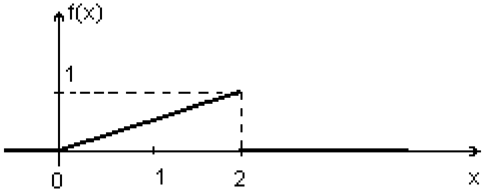
Требуется найти дифференциальную функцию (плотность вероятности); построить графики интегральной и дифференциальной функций.

Решение: $f(x) = F'(x) \Rightarrow f(x) = \begin{cases} 0, & \text{при} \quad x \le 0, \\ \dfrac{x}{2}, & \text{при} \quad 0 < x \le 2, \\ 0, & \text{при} \quad x > 2. \end{cases}$

Построим график функции F(x)



Построим график функции f(x)



2.3. Числовые характеристики случайной величины

Закон распределения полностью характеризует случайную величину с вероятностной точки зрения. При решении многих практических задач нет необходимости характеризовать случайную величину полностью, а достаточно иметь о случайной величине только некоторое общее представление, т.е. некоторые общие черты закона распределения. О каждой случайной величине необходимо прежде всего знать ее некоторое среднее значение, около которого группируются возможные значения случайной величины, а также какое-либо число, характеризующее степень разбросанности этих значений относительно среднего. Такие числа называются числовыми характеристиками случайной величины.

Математическим ожиданием дискретной случайной величины называют сумму произведений всех ее возможных значений на соответствующие им вероятности:

$$M(X) = x_1p_1 + x_2p_2 + \ldots + x_np_n$$

или

$$M(X) = \sum_{i=1}^{n} x_i p_i \ .$$

Математическим ожиданием непрерывной случайной величины X, возможные значения которой принадлежат промежутку, называют определенный интеграл

$$M(X) = \int_a^b xf(x)dx \text{, если промежуток конечен,}$$

или

$$M(X) = \int_{-\infty}^{\infty} xf(x)dx \text{, если промежуток бесконечен.}$$

Математическое ожидание случайной величины X приближенно равно среднему арифметическому наблюдаемых значений случайной величины. Причем, это приближение тем ближе, чем больше число испытаний.

Две случайные величины называются независимыми, если закон распределения одной из них не зависит от того, какие возможные значения приняла другая величина. В противном случае случайные величины зависимы.

Произведением независимых случайны величин X и Y называется случайная величина XY, возможные значения которой равны произведениям каждого возможного значения X на каждое возможное значение Y; вероятности возможных значений произведения XY равны произведениям вероятностей возможных значений сомножителей.

Свойства математического ожидания

Пусть $C = \text{const}$.

1. $M(C) = C$.
2. $M(CX) = C \cdot M(X)$.
3. $M(X + Y + \ldots + Z) = M(X) + M(Y) + \ldots M(Z)$.
4. Если X, Y,…,Z – взаимно независимые случайные величины, то $M(XY\ldots Z) = M(X) \ldots M(Z)$.

Дисперсией случайной величины называется математическое ожидание квадрата отклонения случайной величины от ее математического ожидания

$$D(X) = M[X - M(X)]^2,\ D(X) \geq 0.$$

Для дискретной случайной величины дисперсия выражается формулой

$$D(X) = \sum_{i=1}^{n} \left(x_i - M(x)\right)^2 \cdot p_i \ ,$$

а для непрерывной

$$D(X) = \int_{-\infty}^{\infty} (x - M(x))^2 \cdot f(x)dx \ .$$

Дисперсия случайной величины является очень удобной характеристикой рассеивания возможных значений случайной величины, но она не наглядна, так как имеет размерность квадрата случайной величины.

Удобнее иметь характеристику рассеивания, имеющую ту же размерность, что и случайная величина. Такой характеристикой является среднее квадратическое отклонение случайной величины $\sigma(X) = \sqrt{D(X)}$.

Свойства дисперсии

Пусть C = const.

1. $D(C) = 0$.
2. $D(CX) = C^2 \cdot D(X)$.
3. $D(X) = M(X^2) - \left[M(X)\right]^2$, $D(X) = \int_{-\infty}^{\infty} (x^2 \cdot f(x))\, dx - \left[M(x)\right]^2$.

Задача 1. Закон распределения случайной величины X таков:

x_i	1	2	3	4	5	6	7	8
$P(X = x_i)$	0,01	0,1	0,2	0,3	0,08	0,2	0,1	0,01

Найти: 1) математическое ожидание; 2) дисперсию; 3) среднее квадратическое отклонение.

Решение.

1) $M(X) = \sum_{i=1}^{n} x_i p_i$.

$$M(X) = 1 \cdot 0{,}01 + 2 \cdot 0{,}1 + 3 \cdot 0{,}2 + 4 \cdot 0{,}3 + 5 \cdot 0{,}08 + 6 \cdot 0{,}2 + 7 \cdot 0{,}1 + 8 \cdot 0{,}01 = 4{,}39;$$

2) $D(X) = \sum_{i=1}^{n} (x_i - M(x))^2 p_i$.

Запишем закон распределения случайной величины $(X – M(X))^2$:

$(x_i – M(x))^2$	11,4921	5,7121	1,9321	0,1521	0,3721	2,5921	6,8121	13,0321
p_i	0,01	0,1	0,2	0,3	0,08	0,2	0,1	0,01

$D(X)= 11,4921 \cdot 0,01 + 5,7121 \cdot 0,1 + 1,9321 \cdot 0,2 + 0,1521 \cdot 0,3 + 0,3721 \cdot$
$0,08 + 2,5921 \cdot 0,2 + 6,8121 \cdot 0,1 + 13,0321 \cdot 0,01 = 2,47789.$

3) $\sigma(X) = \sqrt{D(\tilde{O})}$,

$$\sigma(\tilde{O}) = \sqrt{2,47789} \approx 1,57 .$$

Задача 2. Охотник, имеющий 4 патрона, стреляет по цели до первого попадания или пока не израсходует все патроны. Найти математическое ожидание и дисперсию числа израсходованных патронов, если вероятность попадания 0,25.

Решение. Обозначим вероятность попадания через p, а вероятность непопадания через q; p= 0,25; q = 1 – 0,25 = 0,75.

Составим закон распределения случайной величины X – числа израсходованных патронов.

x_i	1	2	3	4
$P(x – x_i)$	p	qp	$qqp = q^2p$	$qqqp + qqqq = q^3p + q^4 = q^3(p + q) = q^3$

x_i	1	2	3	4
$p(x – x_i)$	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{9}{64}$	$\frac{27}{64}$

Контроль: $\sum_{i=1}^{n} p_i = \frac{1}{4} + \frac{3}{16} + \frac{9}{24} + \frac{27}{64} = 1$,

$$M(X) = \frac{1}{4} + 2 \cdot \frac{3}{16} + 3 \cdot \frac{9}{64} + 4 \cdot \frac{27}{64} = \frac{16 + 24 + 27 + +108}{64} = \frac{175}{64} = 2,734$$

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

Составим закон распределения случайной величины X_2.

x_i^2	1	4	9	16
p_i	$\frac{1}{4}$	$\frac{3}{16}$	$\frac{9}{64}$	$\frac{27}{64}$

$$M(X^2)=\frac{1}{4}+4\cdot\frac{3}{16}+9\cdot\frac{9}{64}+16\cdot\frac{27}{64}=\frac{16+48+94++432}{64}=\frac{577}{64}=9{,}015.$$

$$D(X)=9{,}015-(2{,}754)^2=9{,}015-7{,}475=1{,}54.$$

Задача 3. Математическое ожидание случайной величины $M(X) = a$, а дисперсия $D(X) = b$. Найти математическое ожидание и дисперсию случайных величин, считая, что X, Y, Z, V независимые:

1) $Y = -X$,
2) $Z = X + 2Y - 1$,
3) $V = 3X - Y + 2Z - 3$.

Решение.

1) $M(Y) = M(-X) = -M(X) = -a$
2) $D(Y) = D(-X) = (-1)^2 D(X) = D(X) = b$.
3) $M(Z) = M(X + 2Y - 1) = M(X) + 2MY - 1 = a - 2a - 1 = -a - 1$,
   $D(Z) = D(X + 2Y - 1) = D(X) + 4D(Y) = b + 4b = 5b$.

3) $M(V) = M(3X - Y + 2Z - 3) = 3M(X) - M(Y) + 2M(Z) - 3 =$
$= 3a + a - 2a - 2 - 3 = 2a - 5$,
$D(V) = D(3X - Y + 2Z - 3) = 9D(X) + D(Y) + 4D(Z) = 9b + b + 20b = 30b$.

Задача 4. Даны независимые случайные величины:

X:

x_i	1	2	3
p_i	0,1	0,3	0,6

и Y:

y_i	-2	-1	0
p_i	0,6	0,3	0,1

Составить распределения $X + Y$ и $X\cdot Y$.

Найти M(X), D(Y), M(Y), M(X + Y), $M(X\cdot Y)$, D(X + Y) и проверить соотношения $M(X + Y) = M(X) + M(Y)$,

$M(X\cdot Y)+M(X)\cdot M(Y)$, $D(X + Y) = D(X) + D(Y)$.

Решение. Составим закон распределения случайной величины X + Y:

X + Y	-1	0	1	2	3
p	0,06	0,18+0,03=0,21	0,01+0,09+0,36 = 0,46	0,03+0,18=0,21	0,06

Контроль: 0,06 + 0,21 + 0,46 + 0,21 + 0,06 = 1.

Составим закон распределения случайной величины $X\cdot Y$.

$X\cdot Y$	-2	-1	0	3	4	6
p	0,06+0,09=0,15	0,03	0,01+0,03+0,09= 0,1	0,18	0,19	0,36

Контроль: 0,36 + 0,18 + 0,15 + 0,03 + 0,1 = 1.

Вычислим:

1) M(X) = 1 · 0,1 + 2 · 0,3 + 3 · 0,6 = 2,5;
2) M(Y)= (-2) · 0,6 + (-1) · 0,3 + 0 · 0,1 = -1,5;
3) $D(X) = M(X_2) – [M(X)]^2$,
M(X) = 1 · 0,1 + 4 · 0,3 + 9 · 0.6 = 6,7,
D(X) = 6,7 – (2,0) = 6,7 – 6,25 = 0,45.
4) $D(Y) = M(Y)^2 – [M(Y)]^2$,
$M(Y^2)$ = 4 · 0,6 + 0,3 + 0 · 0,1 = 2,7,
$D(Y) = 2,7 – (-1,5)^2 = 2,8 – 2,25 = 0,45$.
5) M(X + Y)= (-1) · 0,06 + 0 · 0,21 + 1·0,46 + 2·0,21+3·0,06 =1,
M(X) + M(Y) = 2,5 + (-1,5) = 1.
Следовательно, выполняется соотношение:

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

6) $D(X + Y) = M[(X + Y\}^2] – [M(X + Y\}]^2$,
$M(X + Y\}]^2$ = 1 · 0,06 + 0 · 0,21 + 1 · 0,46 + 0,21 + 9 · 0,06 =1,9,
D(X + Y) = 1,9 – 1 = 0,9, D(X) + D(Y) = 0,45 + 0,45 = 0,9.
Следовательно, выполняется соотношение:

$$D(X + Y) = D(X) + D(Y).$$

7) $M(X\cdot Y)$ = (-6) · 0,36 + (-4) · 0,18 + (-3) · 0,18 + (-2) · 0,15 +
+ (-1) · 0,03 + 0 · 0,1 = 2,16 – 0,72 – 0,54 – 0,3 – 0,03 = -3,75,
$M(X)\cdot M(Y)$ = 2,5 · (-1,5) = – 3,75.
Следовательно, выполняется соотношение:

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

Задача 5. Найти M(X), если

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

Решение. Дана непрерывная случайная величина X, все ненулевые значения которой лежат на отрезке $[0, \frac{\pi}{2}]$, следовательно,

$$M(X) = \int_0^{\frac{\pi}{2}} xf(x)dx = \int_0^{\frac{\pi}{2}} x\sin x dx = -\int_0^{\frac{\pi}{2}} x d\cos x = -x\cos x \Big|_0^{\frac{\pi}{2}} + \int_0^{\frac{\pi}{2}} \cos x dx =$$

$$= \sin x \Big|_0^{\frac{\pi}{2}} = \sin\frac{\pi}{2} = 1.$$

Задача 6. Случайная величина X задана функцией распределения

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

Найти f(x), M(X), D(X), $\sigma(X)$.

Решение.

1) Так как $F'(x)=f(x)$, то

$$f(x)=\begin{cases}0, & \text{ïðè} \quad x\le 0,\\ 2x, & \text{ïðè} \quad 0<x\le 1,\\ 0, & \text{ïðè} \quad x>1.\end{cases}$$

2) Все значения случайной величины лежат в интервале [0,1], следовательно,

$$M(X)=\int_0^1 x\,f(x)\,dx=\int_0^1 x\cdot 2x dx=2\int_0^1 x^2 dx=2\cdot\frac{x^3}{3}\Big|_0^1=\frac{2}{3}.$$

3) $D(X) = M(X^2) – [M(X)]^2$,

$$M(X^2)=\int_0^1 x^2\,f(x)\,dx=\int_0^1 x^2\cdot 2x\,dx=2\int_0^1 x^3 dx=2\cdot\frac{x^4}{4}\Big|_0^1=\frac{1}{2}.$$

$$D(X)=\frac{1}{2}-\frac{4}{9}=\frac{9-8}{18}=\frac{1}{18}.$$

4) $\sigma(X)=\sqrt{D(\tilde{O})}$, $\quad \sigma(X)=\sqrt{\dfrac{1}{18}}=\dfrac{1}{3}\sqrt{\dfrac{1}{2}}=\dfrac{\sqrt{2}}{6}.$

Задача 7. В результате испытаний двух приборов (А и В) установлена вероятность наблюдения помех, оцениваемых по трехбалльной системе, которая задается в виде таблицы.

x_i	0	1	2	3
P_A	0,70	0,20	0,06	0,04
P_B	0,80	0,06	0,04	0,10

По приведенным данным выбрать лучший прибор, если лучшим является тот, который в среднем имеет меньший уровень помех.

Решение. Пусть X – уровень помех. Средний уровень для прибора A и прибора B:

$$M_A(X) = 0 \cdot 0,7 + 1 \cdot 0,20 + 2 \cdot 0,06 + 3 \cdot 0,04 = 0,44 \text{ (балла)},$$

$$M_B(X) = 0 \cdot 0,8 + 1 \cdot 0,6 + 2 \cdot 0,14 + 3 \cdot 0,10 = 0,44 \text{ (балла)}.$$

По среднему баллу приборы равноценны. В качестве дополнительного критерия используем среднее квадратическое отклонение уровня помех

$$\sigma_A(X)=\sqrt{D_A(X)}=\sqrt{M_A(X^2)-[M_A(X)]^2}=\sqrt{0{,}80-(0{,}44)^2}\approx$$

$$\approx 0{,}78 \text{ (балла)},$$

$$\sigma_B(X)=\sqrt{D_B(X)}=\sqrt{M_B(X^2)-[M_B(X)]^2}=\sqrt{0{,}12-(0{,}44)^2}\approx$$

$\approx 0{,}96$ (балла).

Таким образом, прибор А дает более устойчивые показания относительно средних и, следовательно, он лучше прибора В.

3. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Среди законов распределения вероятностей дискретных случайных величин наиболее распространенным является биномиальное распределение.

Пусть производится n независимых испытаний, в каждом из которых событие А может появиться либо не появиться. Вероятность наступления события во всех испытаниях постоянна и равна p (следовательно, вероятность непоявления q = 1 – p).

Рассмотрим в качестве дискретной случайной величины X число появлений события А в этих испытаниях.

Требуется найти закон распределения величины X.

Возможные значения X таковы: $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, ..., $x_{n+1} = n$. Соответствующие им вероятности вычисляются по формуле Бернулли:

$$P_n(k) = \tilde{N}_n^{\hat{e}} p^{\hat{e}} q^{n-\hat{e}},$$

где $k = 0, 1, 2, \ldots$.

Биномиальным называют распределение вероятностей, определяемое формулой Бернулли.

Напишем биномиальный закон в виде таблицы:

x_i	0	1	2	3	…	k	…	n
$P(x + x_i)$	q^n	npq^{n-1}	$\frac{n(n-1)}{2!}p^2q^{n-2}$	$\frac{n(n-1)(n-2)}{3!}p^3q^{n-3}$	…	$\tilde{N}_n^{\hat{e}} p^{\hat{e}} q^{n-\hat{e}}$,	…	p^nq^0

Контроль:

$$C_n^0p^0q^n + C_n^1p^1q^{n-1} + C_n^2p^2q^{n-2} + \ldots + C_n^{\hat{e}}p^{\hat{e}}q^{n-\hat{e}} + \ldots + C_n^np^nq^0 =$$

$$= (p + q)^n = 1.$$

Числовые характеристики при биномиальном законе распределения имеют вид:

1) Математическое ожидание случайной величины X: M(X) = np.
2) Дисперсия случайной величины X: D(X) = npq.

3) Среднее квадратическое отклонение случайной величины X:

$$\sigma(X) = \sqrt{D(X)} = \sqrt{n \cdot p \cdot q}\ .$$

Замечание. Если n велико и вероятность события мала ($p < 0{,}1$), то формула Бернулли непригодна. В этих случаях прибегают к асимптотической формуле Пуассона. При этом имеет место важное допущение: произведение np сохраняет постоянное значение. Обозначим np = a.

Формула Пуассона имеет вид:

$$P_n^k = \frac{a^k}{k!} \cdot e^{-a}\ .$$

Эта формула выражает закон распределения вероятностей Пуассона массовых (n велико) и редких (p мало) событий:

$$M(X) = a,\ D(X) = a,\ \sigma(X) = \sqrt{a}\ .$$

3.1. Простейший поток событий

Потоком событий называют последовательность однородных событий, которые наступают в случайные моменты времени.

Простейшим (пуассоновским) называют поток событий, который обладает свойствами стационарности, отсутствия последействий и ординарности.

Интенсивностью потока λ называют среднее число событий, которые появляются в единицу времени.

Если интенсивность потока известна, то вероятность появления k событий простейшего потока за время t определяется формулой Пуассона:

$$P_t(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}\ .$$

Задача 1. Монету подбрасывают 7 раз. Найти закон распределения числа появления герба. Найти математическое ожидание и дисперсию этой случайной величины.

Решение.

x_i	0	1	2	3	4	5	6	7
$p(x - x_i)$	$P_7(0)$	$P_7(1)$	$P_7(2)$	$P_7(3)$	$P_7(4)$	$P_7(5)$	$P_7(6)$	$P_7(7)$

$$P_n(k) = \tilde{N}_n^m p^m q^{n-m}.$$

$$p = q = \frac{1}{2},$$

$$P_7(1) = C_7^1 (\frac{1}{2})^1 \cdot (\frac{1}{2})^6 = \frac{7}{128},$$

$$P_7(2) = C_7^2 (\frac{1}{2})^2 \cdot (\frac{1}{2})^3 = \frac{21}{128},$$

$$P_7(4) = C_7^4 (\frac{1}{2})^4 \cdot (\frac{1}{2})^3 = \frac{35}{128},$$

$$P_7(5) = C_7^5 (\frac{1}{2})^5 \cdot (\frac{1}{2})^2 = \frac{21}{128},$$

$$P_7(6) = C_7^6 (\frac{1}{2})^6 \cdot (\frac{1}{2})^1 = \frac{7}{128},$$

$$P_7(7) = C_7^7 (\frac{1}{2})^7 \cdot (\frac{1}{2})^0 = \frac{1}{128},$$

Итак: нашли биномиальный закон распределения случайной величины X:

x_i	0	1	2	3	4	5	6	7
$p(x = x_i)$	$\frac{1}{128}$	$\frac{7}{128}$	$\frac{21}{128}$	$\frac{35}{128}$	$\frac{35}{128}$	$\frac{21}{128}$	$\frac{7}{128}$	$\frac{1}{128}$

Контроль: $\frac{1}{128} + \frac{7}{128} + \frac{21}{128} + \frac{35}{128} + \frac{35}{128} + \frac{21}{128} + \frac{7}{128} + \frac{128}{128} = 1$.

$$M(X) = np,\ M(X) = 7 \cdot \frac{1}{2} = 3{,}5,$$

$$D(X) = npq,\ D(X) = 7 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{7}{4} = 1{,}75.$$

Задача 2. Из большой партии изделий берут на пробу 10 штук. Известно, что доля нестандартных изделий во всей партии составляет 25%. Требуется найти вероятность того, что более пяти отобранных изделий окажутся нестандартными.

Решение. Событие A – обнаружение нестандартности у одного из отобранных изделий.

$$P(A) = p = \frac{25}{100} = 0{,}25;\ q = 1 - 0{,}25 = 0{,}75;\ n = 10.$$

X – количество нестандартных изделий среди отобранных.

Случайная величина X распределена по биномиальному закону распределения вероятностей. Ряд распределения имеет вид (с точностью до 0,0001)

x_i	0	1	2	3	4	5	6	7	8	9	10
$P(X = x_i)$	0,0563	0,1877	0,2816	0,2503	0,1460	0,0584	0,0162	0,0031	0,0004	0,0000	0,0000

$$P(X > 5) = \sum_{ê=6}^{10} Ð(Õ \geq x_6) = 0{,}0162 + 0{,}0031 + 0{,}0004 \approx 0{,}020 .$$

Задача 3. Вероятность того, что изделие не выдержит испытания, равна 0,001. Найти вероятность того, что из 5000 изделий более чем одно не выдержит испытания.

Решение.

n = 5000 (велико); p = 0,001 (мала), a = 5000 · 0,01 = 5.

P(k > 1) – ?

Имеет место распределение Пуассона.

$$P_{5000}\ (k > 1) = 1 - P(k \leq 1) = 1 - p_n(0) - p_n(1) .$$

$$Ð_n(ê) = \frac{a^k}{k!} \cdot å^{-a} .$$

$$P(0) = \frac{5^0}{0!} \cdot å^{-5} \approx 0{,}0067 \text{ (по таблице)}.$$

$$P(1) = \frac{5^1}{1!} \cdot å^{-5} \approx 0{,}0337 .$$

$$P(k > 1) = 1 – 0{,}0067 – 0{,}0337 = 1 – 0{,}0404 = 0{,}9596.$$

Задача 4. Среднее число заявок, поступающих на предприятие бытового обслуживания за 1 час, равно трем. Найти вероятность того, что за 3 часа поступит: 1) 6 заявок; 2) менее 6 заявок; 3) не менее 6 заявок.

Решение. Здесь имеет место пуассоновский поток событий.

$$P_t(k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}; \quad \lambda = 3; \quad t = 3 \div àñà, \ \lambda t = 3 \cdot 3 = 9 .$$

$$P_3(k = 6) = \frac{(9)^6}{6!} \cdot å^{-9} = 0{,}0911 \text{ (находим по таблице)}.$$

$$P_3(k < 6) = P_3(0) + P_3(1) + P_3(3) + P_4(4) + P_5(5) =$$

$$= 0{,}000123 + 0{,}001111 + 0{,}00\,4998 + 0{,}014994 + 0{,}33737 + 0{,}060727 = 0{,}115690.$$

$$P^3(k \geq 6) = 1 - P_3(k < 6) = 1 - 0{,}11569 = 0{,}884310.$$

Задача 5. Автоматическая телефонная станция получает в среднем за час 540 вызовов. Какова вероятность того, что в данную минуту она получит ровно 20 вызовов?

Решение. Так как вызовы независимы друг от друга, то число вызовов за промежуток времени t = 1 мин представляет собой пуассоновский поток событий.

$\lambda = \dfrac{540}{60} = 9;\quad t = 1$ мин.; λ – среднее число вызовов в 1 мин.

$$P_t(k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}\ .$$

$$P_1(20) = \frac{(9)^{20}}{20!} \cdot e^{-9} = 0{,}000617 \text{ (нашли по таблице).}$$

Задача 6. Пусть вероятность изготовления нестандартного изделия при некотором технологическом процессе равна 0,06. Контролер берет из партии изделие и сразу проверяет его качество. Если оно оказывается нестандартным, дальнейшие испытания прекращаются, а партия задерживается. Если же изделие оказывается стандартным, контролер берет следующее: и т.д., но всего проверяет не более пяти изделий. Составить закон распределения числа проверяемых изделий.

Решение. Искомая случайная величина X может принимать значения 1, 2, 3, 4 и 5.

p = 0,06; q = 1 – 0,06 = 0,94 – вероятность изготовления стандартного изделия.

$$P\,(X = 1) = p = 0{,}06;$$

$$P(X = 2) = q \cdot p = 0{,}94 \cdot 0{,}06 \approx 0{,}056;$$

$$P(X = 3) = q \cdot q \cdot p = q^2 \cdot p \cdot 0{,}94^2 \cdot 0{,}06 \approx 0{,}053;$$

$$P(X = 4) = q \cdot q \cdot q \cdot p = q^3 \cdot p \cdot 0{,}94^3 \cdot 0{,}06 \approx 0{,}050;$$

$$P(X = 5) = q \cdot q \cdot q \cdot q \cdot p + q^5 = q^4 (p + q) = 0{,}94^4 \approx 0{,}781.$$

Следовательно, закон распределения случайной величины X можно представить в виде ряда:

x_i	1	2	3	4	5
$P(X = x_i)$	0,06	0,056	0,053	0,050	0,781

Контроль: 0,06 + 0,056 + 0,053 + 0,50 + 0,781 = 1.

4. ЗАКОНЫ РАСПРЕДЕЛЕНИЯ НЕПРЕРЫВНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

При решении задач, которые выдвигает практика, приходится сталкиваться с различными распределениями непрерывных случайных величин. Плотности распределений непрерывных случайных величин называют также законами распределений. Часто встречаются, например, законы равномерного, показательного и нормального распределения.

4.1. Равномерное распределение

Непрерывная случайная величина X имеет равномерное распределение на отрезке [a, b], если на этом отрезке вероятность распределения случайной величины постоянна, а вне его равна нулю, т.е. если:

$$f(x) = \begin{cases} 0, & \text{при} \quad x < a, \\ C, & \text{при} \quad a \le x \le b, \\ 0, & \text{при} \quad x > b, \end{cases}$$

где C = const. Постоянная C находится из условия:

$$\int_a^b f(x)dx = 1 \ \text{èëè} \ \int_a^b cdx = 1 \Rightarrow \text{Ñ} = \frac{1}{\hat{a} - \grave{a}}.$$

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

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

График функции f(x) имеет вид:

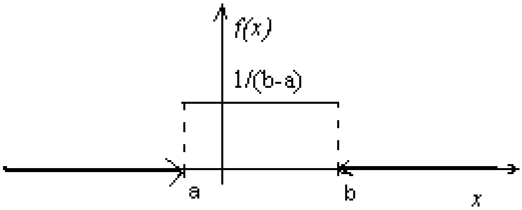
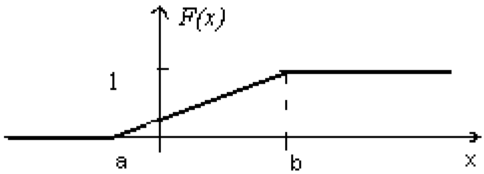


График функции F(x) имеет вид:



Со случайной величиной, имеющей равномерное распределение, мы часто встречаемся в измерительной практике при округлении отсчетов измерительных приборов до целых делений шкал.

$$M(X)=\int_a^b \tilde{o}\cdot\frac{1}{\hat{a}-\grave{a}}dx=\frac{1}{\hat{a}-\grave{a}}\left.\frac{\tilde{o}^2}{2}\right|_a^b=\frac{a+b}{2};$$

$$D(X)=\int_a^b\left(\tilde{o}-\frac{\hat{a}+\grave{a}}{2}\right)^2\cdot\frac{1}{\hat{a}-\grave{a}}d\tilde{o}=\frac{(\hat{a}-\grave{a})^2}{12};$$

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

Вероятность попадания случайной величины, имеющей равномерное распределение, в интервал $(\alpha;\beta)$ вычисляется по формуле

$$P(a<x<b)=\int_\alpha^\beta\left(\frac{1}{b-a}\right)dx=\frac{\beta-\alpha}{b-a}.$$

4.2. Показательное распределение

В практических приложениях теории вероятностей, особенно в теории массового обслуживания, исследовании операций, в физике, в биологии, в вопросах надежности и других приложениях, часто имеют дело со случайными величинами, имеющими показательное распределение.

Показательным (экспоненциальным) называют распределение вероятностей непрерывной случайной величины X, которое описывается плотностью

$$f(x)=\begin{cases}0, & \text{при} \quad x<0,\\ \lambda\cdot e^{-\lambda x}, & \text{при} \quad x\ge 0, \text{ где } \lambda>0 \ (\lambda-\text{const}).\end{cases}$$

Кривая распределения f(x) имеет вид:

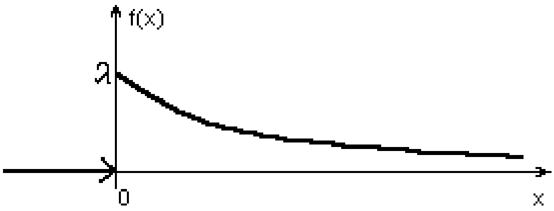
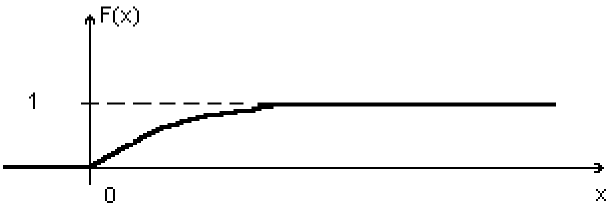


График F(x) имеет вид:



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

Вероятность попадания в интервал $(\alpha;\beta)$ непрерывной случайной величины X, распределенной по показательному закону, вычисляется по формуле: $P(\alpha < x < \beta) = e^{-\lambda\alpha} - e^{-\lambda\beta}$.

$$M(X) = \frac{1}{\lambda}, \quad D(X) = \frac{1}{\lambda^2}; \quad \sigma(X) = \frac{1}{\lambda}.$$

Таким образом, параметр λ есть величина обратная математическому ожиданию или среднему квадратическому отклонению.

4.3. Нормальное распределение

Среди распределений непрерывных случайных величин центральное место занимает нормальный закон (закон Гаусса).

Важность нормального закона распределения определяется тем, что к нему обычно приводят задачи, связанные с распределением сумм большого числа случайных величин. Случайные величины, распределенные по нормальному закону, имеют широкое распространение.

Например, случайными величинами, распределенными по нормальному закону, являются: рост мужчины (женщины), вес вылавливаемой рыбы (одного вида), дальность полета снаряда при стрельбе из орудия, винтовки, отклонения действительных размеров деталей, обработанных на станке, от проектных размеров, ошибки при измерении.

Основная особенность, выделяющая нормальный закон среди других законов, состоит в том, что он является предельным законом, к которому приближаются другие законы распределения.

Нормальным называют распределение вероятностей непрерывной случайной величины X, плотность которого имеет вид:

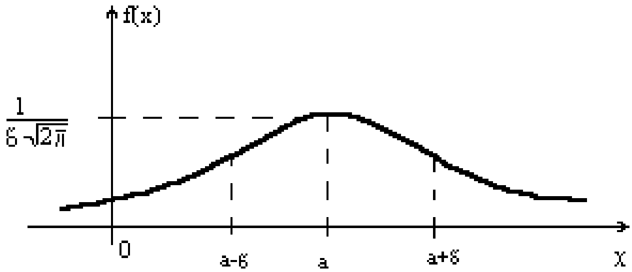
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\cdot \text{å}^{-\frac{(\text{õ}-\text{à})^2}{2\sigma^2}},$$

где a, σ – параметры нормального распределения, причем

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

т.е. параметр a равен математическому ожиданию, а параметр σ равен среднему квадратическому отклонению нормально распределенной случайной величины.

График плотности вероятностей нормального распределения называют нормальной кривой (кривой Гаусса), которая имеет вид:



Изменение величины параметра a не изменяет формы нормальной кривой, а приводит лишь к ее сдвигу вдоль оси OX: вправо, если a возрастает, и влево, если a убывает.

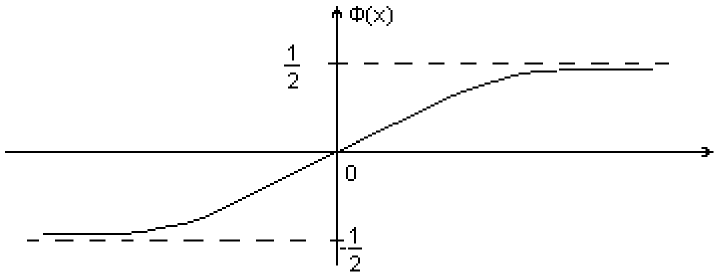
С возрастанием σ максимальная ордината нормальной кривой $\frac{1}{\sigma\sqrt{2\pi}}$ убывает, а сама кривая становится более пологой, при убывании σ максимальная ордината возрастает, а сама кривая становится более островершинной. При этом площадь, ограниченная кривой и осью OX, остается постоянной и равной 1.

Вероятность попадания в интервал $(\alpha;\beta)$ нормально распределенной случайной величины вычисляется по формуле:

$$P(\alpha < x < \beta) = \hat{O}\left(\frac{\beta - \grave{a}}{\sigma}\right) - \hat{O}\left(\frac{\alpha - \grave{a}}{\sigma}\right),$$

где $\hat{O}(\tilde{o}) = \frac{1}{\sqrt{2\pi}} \int_0^x \mathring{a}^{-\frac{\tilde{o}^2}{2}} dx$ – функция Лапласа, которая обладает свойствами:

1) $\Phi(0) = 0$;
2) $\Phi(\infty) = \frac{1}{2}$;
3) $\Phi(-x) = -\Phi(x)$ и график ее имеет вид:



Задача 1. Автобусы некоторого маршрута идут строго по расписанию. Интервал движения 5 минут. Найти вероятность того, что пассажир, подошедший к остановке, будет ожидать очередной автобус менее 3 минут.

Решение. Случайная величина Х – время ожидания автобуса распределена равномерно. Плотность распределения вероятностей вычисляется так:

$$f(x) = \begin{cases} 0, & \text{ïðè} \quad x < a, \\ \frac{1}{b-a}, & \text{ïðè} \quad a \le x \le b, \\ 0, & \text{ïðè} \quad x > b. \end{cases}$$

$b - a = 5$; тогда:

$$f(x) = \begin{cases} 0, & \text{при} \quad x < 0, \\ \frac{1}{5}, & \text{при} \quad 0 \le x \le 5, \\ 0, & \text{при} \quad x > 5. \end{cases}$$

$$P(0 < X < 3) = \int_0^3 \frac{1}{5}dx = \frac{1}{5}x\Big|_0^3 = 0{,}6\,.$$

Это вероятность того, что пассажир будет ожидать автобус менее 3 минут.

Задача 2. Плотность вероятностей случайной величины X имеет вид

$$f(x) = \begin{cases} \dfrac{1}{2l}, & \text{при} \quad |x-a| \le l, \\ 0, & \text{при} \quad |x-a| > l, \ \ l - \text{const.} \end{cases}$$

Определить: а) M(X); б) D(X).

Решение. Здесь имеет место закон равномерного распределения случайной величины на отрезке [a – l; a + l]. Распишем подробно:

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

$$M(X) = \frac{(a-\ell)+(a+\ell)}{2} = \frac{2à}{2} = à\,;$$

$$D(X) = \frac{[(a+\ell)-(a-\ell)]^2}{12} = \frac{(2\ell)^2}{12} = \frac{4\ell^2}{12} = \frac{\ell^2}{3}\,.$$

Задача 3. Непрерывная случайная величина X распределена по показательному закону, который задан дифференциальной функцией

$$f(x) = \begin{cases} 0, & \text{ïðè} \quad x < 0, \\ 3e^{-3x}, & \text{ïðè} \quad x \ge 0. \end{cases}$$

Найти: 1) M(X), 2) D(X), 3) $\sigma(X)$, 4) P(1 < X < 2)

Решение. Поскольку имеет место показательное распределение, то $\lambda = 3$.

1) $M(X) = \dfrac{1}{\lambda} = \dfrac{1}{3}$;

2) $D(X) = \dfrac{1}{\lambda^2} = \dfrac{1}{9}$;

3) $\sigma(X) = \dfrac{1}{\lambda} = \dfrac{1}{3}$.

Используя формулу: $P(\alpha < X < \beta) = e^{-\lambda\alpha} - e^{-\lambda\beta}$, $\alpha = 1, \beta = 3,$

$$P(1 < X < 3) = e^{-3} - e^{-9}.$$

Задача 4. Найти дисперсию и среднее квадратическое отклонение показательного закона, заданного функцией распределения:

$$F(x) = \begin{cases} 1 - e^{-0,4x}, \text{ при } & x \ge 0, \\ 0, \quad \text{ при } & x < 0. \end{cases}$$

Решение. Для показательного закона распределения функция распределения имеет вид:

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

Следовательно, в данной задаче λ =0,4, а

$$D(X) = \frac{1}{\lambda^2} = \frac{1}{0{,}16} = \frac{100}{16} = \frac{25}{4} = 6\frac{1}{4},$$

$$\sigma(X) = \sqrt{\frac{100}{16}} = \frac{10}{4} = \frac{5}{2} = 2{,}5.$$

Задача 5. Случайная величина X имеет плотность вероятностей:

$$f(x) = \frac{1}{5\sqrt{2\pi}} \text{å}^{-\frac{(\text{õ}-1)^2}{50}}.$$

Какой закон распределения? Найти M(X), D(X) и ее функцию распределения.

Решение. Случайная величина X распределена по нормальному закону, у которого $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{å}^{-\frac{(\text{õ}-\text{à})^2}{2\sigma^2}}$.

Сравнивая эти функции, видно, что M(X) = a = 1; $D(X) = \sigma^2 = 5^2$.

Найдем функцию распределения: $F(x) = \int\limits_{-\infty}^{x} f(x)dx$;

$$F(x) = \int\limits_{-\infty}^{x} \frac{1}{5\sqrt{2\pi}} \text{å}^{\frac{(t-1)^2}{50}} dt.$$

Сделаем замену: $\frac{t-1}{5} = Z$, $dZ = \frac{dt}{5}$; $\lim Z = -\infty$ при $t \to -\infty$;

$$F(x) = -\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x-1}{5}} \text{å}^{\frac{-z^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{0} \text{å}^{\frac{-z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x-1}{5}} \text{å}^{\frac{-z^2}{2}} dz =$$

$$= -\frac{1}{\sqrt{2\pi}} \int\limits_{0}^{-\infty} \text{å}^{\frac{-z^2}{2}} dz + \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\frac{x-1}{5}} \text{å}^{\frac{-z^2}{2}} dz = -\hat{\text{O}}(-\infty) + \hat{\text{O}}\left(\frac{x-1}{5}\right) = \frac{1}{2} + \hat{\text{O}}\left(\frac{x-1}{5}\right),$$

где $\Phi(x) = \dfrac{1}{\sqrt{2\pi}} \int\limits_0^x e^{\frac{-z^2}{2}} dz$ – функция Лапласа.

Задача 6. У нормально распределенной случайной величины X известны M(X) = 10 и D(X) = 4. Найти P(12 < X < 14).

Решение. Для нормально распределенной случайной величины плотность вероятностей имеет вид:

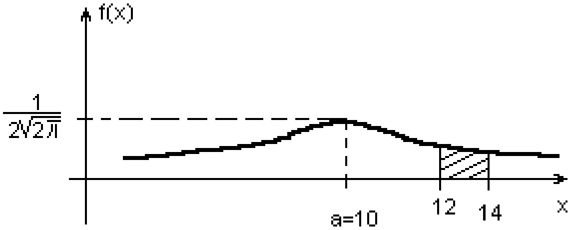
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

Из условия $M(X) = a = 10;\ D(X) = \sigma^2 = 4, \Rightarrow \sigma = 2$.

$$P(\alpha < X < \beta) = \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right);\ \alpha = 12,\ \beta = 14.$$

$$P(12 < X < 14) = \Phi\left(\frac{14-10}{2}\right) - \Phi\left(\frac{12-10}{2}\right) = \Phi(2) - \Phi(1) = 0{,}4771 - 0{,}413 =$$

$$= 0{,}1359.$$



Графически данная вероятность представляет собой площадь заштрихованной фигуры.

5. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Диалектический материализм учит, что единственным критерием правильности теоретических предложений служит практика (опыт).

На основе этого различные гипотезы могут быть приняты или отвергнуты в результате экспериментального исследования.

Для исследования объектов применяется статистический метод. Его идея заключается в том, что изучают не единичное явление, а массовую совокупность однородных явлений. В результате выделяются общие для всей совокупности статистические закономерности.

Раздел теории вероятностей, в котором рассматриваются способы обработки экспериментальных данных с целью определения вероятностных свойств различных случайных объектов, называется математической статистикой.

Математическая статистика изучает методы статистических наблюдений и анализ статистических данных.

5.1. Основные задачи математической статистики

1. Нахождение закона распределения изучаемой случайной величины по статистическим данным. Для этого достаточно найти функцию распределения.
2. Проверка правдоподобия гипотез. На основании анализа статистического материала можно только высказать предположение о законе распределения случайной величины (или о функции распределения $F(x)$). Статистический материал может подтвердить или опровергнуть эту гипотезу. Рассматриваются специальные критерии, позволяющие проверить правдоподобие гипотез.
3. Определение параметров распределения случайной величины.
4. Определение точности и надежности результатов, полученных при вычислении параметров (числовых характеристик). Здесь нужно указать с заданной вероятностью (надежностью) границы интервала, в котором лежит истинное значение числовой характеристики изучаемой случайной величины. Такой интервал называют доверительным.

5.2. Выборочный метод

Пусть производится наблюдение множества объектов, обладающих некоторыми признаками.

Генеральной совокупностью называется множество однородных объектов, обладающих некоторыми качественными или количественными признаками.

Через N обозначается число объектов генеральной совокупности (объем генеральной совокупности).

Число членов совокупности, обладающих одним значением признака, называется частотой этого признака и обозначается через m_i. Значение признака называется вариантой и обозначается $X = \{x_1, x_2, ..., x_k\}$.

Значения признака могут быть как дискретными так и непрерывными. Например, количество радиоламп определенной мощности в данной партии – дискретное значение.

Пример. В течение часа магазин продал 10 пар обуви следующих размеров 35, 38, 38, 39, 37, 36, 37, 38, 37, 39. Сгруппируем по признаку – размер, m – частота признака.

x_i	35	36	37	38	39
m	1	1	3	3	2

x_1=35, x_2=36, x_3=37, x_4=38, x_5=39 – варианты.

Упорядоченный перечень наблюдаемых вариант и соответствующих им частот называется статистическим рядом.

Полученный в примере ряд – дискретный.

Существует также непрерывные статистические ряды (для непрерывных случайных величин).

Пример. В некотором классе имеются учащиеся ростом 155–160 см, 160–165 см, 165–170 см, 170–175 см. Количество учеников соответственно 7, 10, 8, 5. Составим статистический ряд.

x_i	155–160	160–165	165–170	170–175
m	7	10	8	5

– интервальный статистический ряд.

Наиболее полное представление о генеральной совокупности будет получено в результате сплошного ее обследования, т.е. когда обследуют каждый из ее объектов относительно интересующего признака.

Однако при большом объеме N сплошное обследование бывает невозможно. В этих случаях применяют выборочный метод.

Выборочной совокупностью (выборкой) называется совокупность случайно отобранных из генеральной совокупности объектов.

Через n обозначим объем выборки. $n \leq N$.

Выборка должна удовлетворять следующим условиям:

1) быть представительной, т.е. объем выборки должен быть достаточно большим;

2) характер выборки должен быть случайным, т.е. все объекты имеют одинаковую вероятность попасть в выборку.

Выборка может быть повторной (перед отбором следующего объекты возвращаются в генеральную совокупность) и бесповторной.

5.3. Виды отбора

1. Отбор, не требующий расчленения на части генеральной совокупности:

а) простой случайный бесповторный отбор;

б) простой случайный повторный отбор.

2. Отбор, при котором генеральная совокупность расчленяется на части:

а) механический (генеральная совокупность механически делится на n групп и из каждой группы берется один объект);

б) типический (объекты выбираются из каждой типической части генеральной совокупности). Типическим отбором пользуются в том случае, когда обследуемый признак заметно колеблется в различных типических частях генеральной совокупности. Например, если продукция изготовляется на нескольких машинах, среди которых есть более и менее изношенные, то здесь целесообразен типический отбор;

в) серийный (объекты из генеральной совокупности выбираются не по одному, а сериями).

На практике чаще всего применяется комбинированный отбор. Например, разбивают генеральную совокупность на серии одинакового объема, затем простым случайным отбором выбирают несколько серий и из каждой серии простым случайным отбором извлекают отдельные объекты.

5.4. Статистический ряд

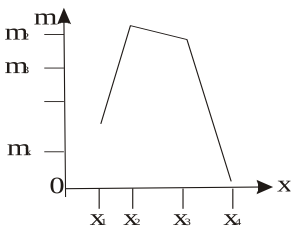
Рассмотрим дискретный статистический ряд:

x_i	x_1	$x2$	…	x_k
m_i	m_1	m_2	…	m_k,

$m_1 + m_2 + \ldots + m_k = n$ – объем выборки.

$\frac{m_i}{n} = w_i$ называется относительной частотой значения x_i, $i = \overline{1, k}$.

Ломаная, соединяющая точки (x_i; m_i) называется полигоном частот статистического ряда.



Ломаная, соединяющая точки $\left(x_i; \dfrac{m_i}{n}\right)$, называется полигоном относительных частот.

Для геометрического изображения интервального статистического ряда используется гистограмма.

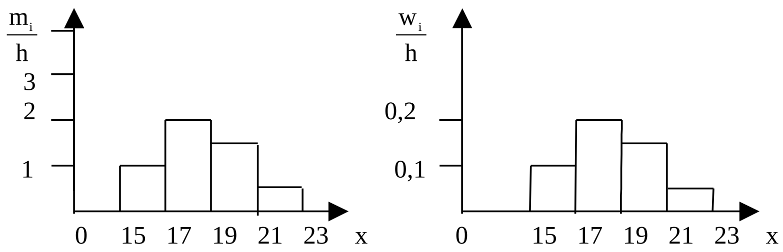
Гистограммой частот называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы статистического ряда, длиною h, а высоты равны отношению $\dfrac{m_i}{h}$ (плотность частоты).

Гистограммой относительных частот называется ступенчатая фигура, состоящая из прямоугольников, основаниями которых служат частичные интервалы, длиною h, а высоты равны $\dfrac{w_i}{h}$ (плотность относительной частоты).

Пример.

x_i	15–17	17–19	19–21	21–23
m_i	2	4	3	1

h=2, n=10, $w_i = \dfrac{m_i}{n}$



Площадь гистограммы частот равна $\sum_{i=1}^{k} \dfrac{m_i}{h} \cdot h = n$ – объему выборки.

Площадь гистограммы относительных частот равна $\sum_{i=1}^{k} \dfrac{m_i}{n \cdot h} \cdot h = \sum_{i=1}^{k} \dfrac{m_i}{n} = 1$.

5.5. Эмпирическая (статистическая) функция распределения

Рассмотрим статистическое распределение частот количественного признака X. Обозначим n – объем выборки, m – число наблюдений, при которых наблюдалось значение признака, меньшее x, т.е. X < x,

т.е. относительная частота события X < x равна $\frac{m}{n}$. Обозначим относительную частоту P^*, следовательно $P^*(X < x) = \frac{m}{n}$.

Если изменяется x, то изменяется и $\frac{m}{n}$, т.е. $P^*(X < x)$ является функцией от x.

Обозначим $F^*(x) = P^*(X < x)$.

Функция $F^*(x)$ называется эмпирической функцией распределения. Итак, по определению $F^*(x) = \frac{m}{n}$, где m – сумма частот вариант, меньших x.

Замечание. Эмпирическая функция распределения $F^*(x)$ отличается от функции распределения F(x) тем, что F(x) определяет вероятность события X < x, а $F^*(x)$ – частоту того же события.

Свойства эмпирической функции распределения аналогичны свойствам функции распределения.

График эмпирической функции распределения рассмотрим на примере:

x_i	90	95	100	105
m_i	5	10	20	5

$\sum m_i = n = 40$.

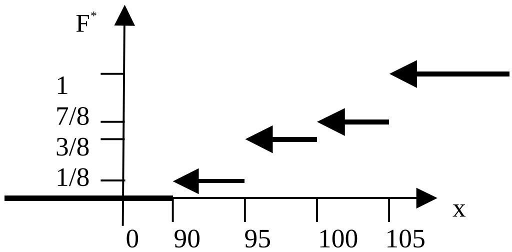
1) $x \le 90$ $F^*(x) = P^*(X < 90) = 0$;

2) $90 < x \le 95$ $F^*(x) = P^*(X < 95) = P^*(X = 90) = \frac{5}{40} = 0{,}125$;

3) $95 < x \le 100$ $F^*(x) = P^*(X < 100) = P^*(X = 90) + P^*(X = 95) = \frac{5+10}{40} = \frac{3}{8}$;

4) $100 < x \le 105$ $F^*(x) = P^*(X < 105) = \frac{5+10+20}{8} = \frac{7}{8}$;

5) $x > 105$ $F^*(x) = P^*(X \le 105) = \frac{5+10+20+5}{40} = 1$.


F*
1
7/8
3/8
1/8
x
0
90
95
100
105

6. СТАТИСТИЧЕСКИЕ ОЦЕНКИ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЯ

6.1. Характеристики статистических рядов

К характеристикам статистических рядов относятся: объем выборки или генеральной совокупности, частоты вариант, относительные частоты вариант.

Рассмотрим другие характеристики статистических рядов.

Генеральной средней $\bar{x}_Г$ называется величина:

$$\bar{x}_Г = \frac{x_1N_1 + x_2N_2 + \ldots + x_kN_k}{N},$$

где x_i – значения признака, N_i – их соответствующие частоты.

$N = \sum N_i$ – объем генеральной совокупности.

Выборочной средней $\bar{x}_в$ называется величина:

$$\bar{x}_в = \frac{x_1m_1 + x_2m_2 + \ldots + x_km_k}{n},$$

где x_i – значения признака, m_i – их соответствующие частоты.

$n = \sum m_i$ – объем выборки.

Генеральной дисперсией $Д_г$ называется величина:

$$Д_Г = \frac{\sum_{i=1}^{k} N_i(x_i - \bar{x}_Г)^2}{N}.$$

Генеральным средним квадратичным отклонением (стандартом) называется $\sigma_Г = \sqrt{Д_Г}$.

Аналогично,

$$Д_в = \frac{\sum m_i(x_i - \bar{x}_в)^2}{n},$$

$$\sigma_в = \sqrt{Д_в}.$$

Замечание. Если объем выборок достаточно велик и выборки производятся правильно, то выборочные характеристики могут служить в качестве приближенных значений генеральных характеристик (тем точнее, чем больше n). Числовые характеристики генеральной совокупности называются параметрами, числовые характеристики выборочной совокупности называются оценками этих параметров.

6.2. Требования, предъявляемые к оценкам параметров

Для того чтобы статистическая оценка давала хорошее приближение, она должна удовлетворять следующим требованиям.

1. Несмещенность – т.е. оценка параметров не должна давать систематических ошибок в сторону завышения или занижения. Обозначим параметр через θ, а его оценку через θ^*.

Оценка называется несмещенной, если $M(\theta^*) = \theta$.

Пусть различным выборкам соответствуют оценки $\theta_1^*, \theta_2^*, ..., \theta_k^*$, тогда в качестве θ^* берут среднее арифметическое этих оценок:

$$\theta^* = \frac{\theta_1^* + \theta_2^* + ... + \theta_k^*}{k}.$$

Если все $\theta_i^* > \theta$, $i = \overline{1,k}$, $M(\theta^*) > \theta$, тогда получается ошибка с избытком. Если все $\theta_i^* < \theta$, $M(\theta^*) < \theta$ – ошибка с недостатком. Условие $M(\theta^*) = \theta$ позволяет избежать этих ошибок.

2. Эффективность. Несмещенные оценки могут иметь большое рассеяние вокруг среднего значения ($M(\theta^*)$), т.е. оценка θ_1^*, полученная по одной выборке, может быть весьма удалена от $M(\theta^*)$. И если в качестве оценки параметра взять θ_1^*, будет большая ошибка. Но если бы все $\theta_i^{*\prime}$ были не сильно рассеяны вокруг $M(\theta^*)$, то ошибка была бы незначительной. Следовательно, дисперсия оценки должна быть минимальной.

Оценка называется эффективной, если она имеет минимальную дисперсию.

3. Состоятельность – оценка по вероятности должна сходиться к оцениваемому параметру. Это означает, что для $\forall\, \varepsilon > 0$ должно выполняться условие:

$$\lim_{n\to\infty} P\left(\left|\theta^* - \theta\right| < \varepsilon\right) = 1.$$

Оценкой M(x) является средняя выборочная, т.е. $M(x) \approx \overline{x}_в$.

Если по нескольким выборкам достаточно большого объема из одной и той же генеральной совокупности будут найдены выборочные средние, то они будут приблизительно равны между собой. В этом и состоит свойство устойчивых выборочных средних.

Можно показать, что $Д_в$ не может служить оценкой $Д_г$, так как $Д_в$ является смещенной оценкой, т.е. $M(Д_в) \neq Д_г$. В действительности:

$$M(Д_в) = \frac{n-1}{n} \cdot Д_г,$$

где n – объем выборки. В качестве оценки $Д_г$ берут величину $S^2 = \frac{n}{n-1} \cdot Д_в$, т.к.

$$M(S^2) = M\left(\frac{n}{n-1} \cdot Д_в\right) = \frac{n}{n-1} M(Д_в) = \frac{n}{n-1} \cdot \frac{n-1}{n} \cdot Д_г,$$

т.е. S^2 является несмещенной оценкой.

S^2 является состоятельной и эффективной оценкой.

S^2 называется исправленной дисперсией.

В качестве оценки $\sigma_Г$ берут $S = \sqrt{\frac{n}{n-1} \cdot Д_в}$.

Замечание. Вычисления $\bar{x}_в$ и $Д_в$ по указанным формулам довольно сложны и отсутствует контроль, поэтому применяют специальные методы расчета средних.

7. ТОЧНОСТЬ ОЦЕНКИ. ДОВЕРИТЕЛЬНАЯ ВЕРОЯТНОСТЬ

7.1. Доверительный интервал

Точечной называется оценка, определяемая одним числом.

Интервальной называется оценка, определяемая двумя числами – концами интервала.

Пусть найденная по данным выборки статистическая характеристика θ^* служит оценкой неизвестного параметра θ. Тогда, чем меньше $\left|\theta-\theta^*\right|$, тем точнее оценка.

Если для некоторого числа $\varepsilon > 0$ выполняется условие

$$\left|\theta-\theta^*\right| < \varepsilon, \tag{1}$$

то ε называется точностью оценки.

Статистические методы не позволяют утверждать, выполняется ли соотношение (1) или нет. Можно лишь говорить о вероятности γ, с которой это неравенство выполняется.

Надежностью (доверительной вероятностью) оценки называется вероятность γ, с которой осуществляется неравенство $\left|\theta-\theta^*\right| < \varepsilon$, т.е.

$$\gamma = P\left(\left|\theta-\theta^*\right| < \varepsilon\right). \tag{2}$$

Неравенство $\left|\theta-\theta^*\right| < \varepsilon$ эквивалентно $\theta^* - \varepsilon < \theta < \varepsilon + \theta^*$. Очевидно, что $\gamma = P\left(\theta^* - \varepsilon < \theta < \varepsilon + \theta^*\right)$.

Доверительным называется интервал $\left(\theta^* - \varepsilon; \theta^* + \varepsilon\right)$, который заключает в себе (покрывает) неизвестный параметр θ с заданной надежностью γ.

7.2. Доверительный интервал для оценки математического ожидания в случае нормального распределения при неизвестном σ

Пусть генеральная совокупность распределена нормально и σ неизвестно. Из интегральной теоремы Лапласа

$$P\left(\left|x-a\right| < \varepsilon\right) = 2\phi\left(\frac{\varepsilon}{\sigma}\right).$$

В нашем случае вместо x мы имеем $\overline{x}_в$, а вместо σ – $\sigma(\overline{x}_в) = \frac{\sigma}{\sqrt{n}}$.

Получаем соотношение: $P(|\overline{x} - a| < \varepsilon) = 2\phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right) = \gamma$.

Отсюда $\phi\left(\frac{\varepsilon \cdot \sqrt{n}}{\sigma}\right) = \frac{\gamma}{2}$. Обозначим $\frac{\varepsilon \cdot \sqrt{n}}{\sigma} = t$, t находим по таблице, следовательно, $\varepsilon = \frac{t \cdot \sigma}{\sqrt{n}}$. Получаем доверительный интервал для оценки неизвестного a:

$$\overline{x}_в - \frac{t \cdot \sigma}{\sqrt{n}} < a < \overline{x}_в + \frac{t \cdot \sigma}{\sqrt{n}}.$$

При возрастании n ε уменьшается, точность возрастает.

Пример. Случайная величина распределена нормально с $\sigma = 2$, $\overline{x}_в = 1{,}6$, $n = 64$, $\gamma = 0{,}95$. Найти доверительный интервал для оценки математического ожидания.

$\phi(t) = \frac{\gamma}{2} = \frac{0{,}95}{2} = 0{,}475$, находим по таблице $t = 1{,}96$, $\varepsilon = \frac{1{,}96 \cdot 2}{8} = 0{,}49$, следовательно, $1{,}6 < a < 1{,}6 + 0{,}49$ или $1{,}11 < a < 2{,}09$.
Доверительный интервал $(1{,}11;\ 2{,}09)$.

7.3. Метод произведений расчета выборочных средних и дисперсий (для равноотстоящих вариантов)

Рассмотрим на примере.

Дан статистический ряд:

x_i 10 11 12 13 14 15 16 17 18 19

m_i 2 3 8 13 25 20 12 10 6 1 $\qquad n = \sum m_i = 100$.

Найти $\overline{x}_в$ и $Д_в$. Воспользуемся расчетной таблицей.

x_i	m_i	u_i	$m_i u_i$	$m_i u_i^2$	$m_i(u_i + 1)^2$
10	2	-4	-8	32	18
11	3	-3	-9	27	12
12	8	-2	-16	32	8

13	13	-1	-13	13	0
14	25	0	0	0	25
15	20	1	20	20	80
16	12	2	24	48	108
17	10	3	30	90	160
18	6	4	24	96	150
19	1	5	5	25	36
	$\sum = 100$		$\sum = 57$	$\sum = 383$	$\sum = 597$

Здесь $u_i = \dfrac{x_i - C}{h}$ – условные варианты, C – варианта с наибольшей частотой, h – шаг.

В нашем примере h = 1, C = 14.

Контроль: $\sum m_i u_i^2 + 2\sum m_i u_i + n = \sum m_i (u_i + 1)^2$.

$$383 + 2 \cdot 57 + 100 = 597.$$

Вычислим условные моменты:

$$M_1^* = \frac{\sum m_i u_i}{n} \qquad M_2^* = \frac{\sum m_i u_i^2}{n}$$

$$\bar{x}_в = M_1^* \cdot h + C, \qquad Д_в = [M_2^* - (M_1^*)^2] \cdot h^2$$

$$\bar{x}_в = 0,57 \cdot 1 + 14 = 14,57 \qquad Д_в = (3,83 - 0,57^2) \cdot 1 = 3,505.$$

Замечание. Если дан статистический ряд с не равноотстоящими вариантами, то его сначала приводят к ряду с равноотстоящими вариантами. Для этого интервал, содержащий все варианты, делят на несколько равных частичных интервалов, находят середины частичных интервалов, которые и образуют последовательность равноотстоящих вариант. В качестве частот новых вариант берут общее число первоначальных вариант, которые попали в соответствующий частичный интервал.

8. ПОНЯТИЯ О КРИТЕРИЯХ СОГЛАСИЯ

8.1. Критерий Пирсона (или χ^2)

Часто закон распределения генеральной совокупности неизвестен, но есть основания предполагать, что он имеет определенный вид, то есть имеет место гипотеза или нулевая гипотеза.

Проверка гипотезы о предполагаемом законе неизвестного распределения производится при помощи специально подобранной случайной величины- критерия согласия.

Критерием согласия называется критерий проверки гипотезы о предполагаемом законе неизвестного распределения.

Ограничимся рассмотрением критерия χ^2 Пирсона о проверке гипотезы о нормальном законе распределения генеральной совокупности (критерий аналогично применяется и для других распределений, в этом достоинство критерия Пирсона).

Критерий не отвечает на вопрос: верна ли гипотеза. Он устанавливает на заданном уровне значимости согласуется или не согласуется гипотеза с данными наблюдений.

Достаточно малую вероятность α, при которой события можно считать практически невозможными, называют уровнем значимости.

На практике обычно принимают уровни значимости, заключенные между 0,01 и 0,05.

8.2. Критерий χ^2

Обычно эмпирические и теоретические частоты различаются. Случайно ли это расхождение?

Расхождение может быть случайным (незначимым) и объясняется малым числом наблюдений или другими признаками. Может быть и неслучайным (значимым) и объясняется тем, что теоретические частоты вычислены, исходя из неверной гипотезы о нормальном законе распределения генеральной совокупности.

Пусть получено эмпирическое распределение с частотами m_i и в предположении нормального распределения генеральной совокупности вычислены теоретические частоты m_i'. При уровне значимости α требуется проверить нулевую гипотезу: генеральная совокупность распределена нормально.

В качестве критерия проверки нулевой гипотезы применим случайную величину, называемую критерием Пирсона

$$\chi^2 = \sum \frac{\left(m_i - m_i'\right)^2}{m_i'} .$$

Это случайная величина, то есть в различных опытах она принимает различные, заранее неизвестные значения. Очевидно, что чем меньше различаются эмпирические и теоретические частоты, тем меньше значение χ^2, то есть χ^2 характеризует близость эмпирического и теоретического распределения.

Возведением в квадрат разности частот устраняют возможность взаимного погашения положительных и отрицательных разностей. Делением на m_i' достигают уменьшения каждого из слагаемых, в противном случае сумма была бы настолько велика, что приводила бы к отклонению нулевой гипотезы даже и тогда, когда она справедлива.

Числом степеней свободы называется число k = S – 1 – r, где S – число групп (частичных интервалов для непрерывных рядов и число различных вариант для дискретных рядов); r – число параметров предполагаемого распределения.

Так как нормальный закон, следовательно, r = 2, следовательно, k = S – 1 – 2 = S – 3.

Для закона Пуассона r = 1, следовательно, k = S – 2.

Для проверки нулевой гипотезы: генеральная совокупность распределена нормально, при заданном уровне значимости α нужно:

1) Вычислить теоретические частоты, а затем наблюдаемое (эмпирическое) значение χ^2:

$$\chi^2_{\text{íàáë}} = \sum \frac{\left(m_i - m_i'\right)^2}{m_i'} ; (*)$$

2) По таблице критических точек распределения χ^2 при заданном уровне значимости α и вычисленному числу степеней свободы k найти $\chi^2_{\text{êð}}(\alpha; k)$;

3) Если $\chi^2_{\text{íàáë}} \le \chi^2_{\text{êð}}$ – нет оснований отвергать гипотезу, если $\chi^2_{\text{íàáë}} > \chi^2_{\text{êð}}$ – нулевую гипотезу отвергают.

Замечание 1. Критерием χ^2 пользуются, если объем выборки достаточно велик $(n \ge 50)$. Каждая группа должна содержать не менее 5–

8 вариантов; малочисленные группы следует объединять в одну, суммируя частоты.

Замечание 2. Для контроля вычислений преобразуем формулу (*):

$$\chi^2=\sum\frac{m_i^2-2\cdot m_i\cdot m_i'+(m_i')^2}{m_i'}=\sum\left(\frac{m_i^2}{m_i'}-2\cdot m_i+m_i'\right),$$

но $\sum m_i=n$ и $\sum m_i'=n$, $\chi^2=\sum\frac{m_i^2}{m_i'}-n$.

Пример. При уровне значимости α=0,05 проверить гипотезу о нормальном распределении генеральной совокупности, если известны эмпирические и теоретические частоты.

m_i 6 13 38 74 106 85 30 10 4

m_i' 3 14 42 82 99 76 37 11 2

r = 2, S = 9, n = 366, следовательно, k = 9 – 3 = 6.

Составим таблицу:

m_i	m_i'	$(m_i-m_i')^2$	m_i^2	$\frac{m_i^2}{m_i'}$	$\frac{(m_i-m_i')^2}{m_i'}$
6	3	9	36	12	3
13	14	1	169	12,07	0,07
38	42	16	1444	34,38	0,38
74	82	64	5476	66,78	0,78
106	99	49	11236	113,49	0,49
85	76	81	7225	95,07	1,07
30	37	49	900	24,32	1,32
10	1	1	100	9,09	0,09
4	2	4	16	8	2
$\sum$ =366	366			375,2	$\chi^2_{\text{набл}}$ =9,2

Контроль: 375,2–366=9,2.

По таблице $\chi^2_{\text{кр}}$ (0,05;6)=12,6

$\chi^2_{\text{набл}}<\chi^2_{\text{кр}}$, следовательно, гипотеза правдоподобна (нет оснований отвергать).

9. ЭЛЕМЕНТЫ ТЕОРИИ КОРРЕЛЯЦИИ

9.1. Функциональная, статистическая и корреляционная зависимости

Рассмотрим две случайные величины X и Y.

Если каждому возможному значению случайной величины X соответствует определенное возможное значение случайной величины Y, то зависимость между X и Y называется функциональной.

Две случайных величины X и Y находятся в статистической зависимости, если изменение одной из них ведет за собой изменение распределения другой.

Частный случай статистической зависимости, при которой изменение одной из них влечет за собой изменение среднего значения другой случайной величины, называется, корреляционной зависимостью.

9.2. Условные средние

Пусть при X = x случайная величина Y имеет возможные значения $y_1, y_2, \ldots, y_k$. Тогда $\frac{y_1 + y_2 + \ldots + y_k}{k}$ называется условным средним $\overline{y}_x$.

Например: при X =2 Y: 5;6;10, следовательно, $\overline{y}_{x=2} = \frac{5+6+10}{3} = 7$.

Аналогично определяется $\overline{x}_y$. Ясно, что при этом X и Y связаны корреляционной зависимостью.

Условная средняя $\overline{y}_x$ является функцией от x, то есть (1) $\overline{y}_x = f(x)$. Это уравнение называется уравнением регрессии Y на X, функция f(x) называется регрессией Y на X, а её график называется линией регрессии.

Если уравнение выражает линейную зависимость, то корреляция называется линейной.

Замечание. Аналогично определяется уравнение регрессии X на Y:

$$\overline{x}_y = \varphi(y) \qquad (2)$$

9.3. Две основные задачи теории корреляции

1. Установление формы связи X и Y, то есть вида функции регрессии (линейная, квадратичная и т.д.)

2. Оценка тесноты (силы) корреляционной связи. Теснота оценивается по величине рассеивания значений Y вокруг условной средней $\overline{y}_x$.

9.4. Линейная корреляция и её параметры. Коэффициент регрессии

Пусть X и Y связаны линейной корреляционной зависимостью. Тогда графики уравнений (1) и (2) есть прямые линии. Пусть проведено n наблюдений и получено n пар чисел: $(x_1; y_1)$, $(x_2; y_2)$;…; $(x_n; y_n)$. Поскольку эти пары можно рассматривать как случайную выборку из генеральной совокупности, то все характеристики, найденные по этим данным называются выборочными.

Найдем выборочное уравнение прямой регрессии ($y=k \cdot x + b$)

$$y = \rho_{\frac{y}{x}} \cdot x + b \qquad (3)$$

$\rho_{\frac{y}{x}}$ называется выборочным коэффициентом регрессии Y на X.

Подберем $\rho_{\frac{y}{x}}$ и b так, чтобы точки $(x_i; y_i)$, $i = \overline{1, n}$, полученные по данным опыта, как можно ближе лежали к прямой (3).

Итак, y_i – наблюдаемая ордината, а через Y_i – обозначим ординату, вычисленную по уравнению (3), соответствующую наблюдаемому значению x_i.

Вместо $\rho_{\frac{y}{x}}$ будем временно для краткости писать просто ρ.

Подберем ρ и b так, чтобы $\sum (Y_i - y_i)^2$ была минимальной. Такой метод подбора параметров ρ и b называется методом наименьших квадратов:

$$\sum_{i=1}^{n} (Y_i - y_i)^2 = \sum_{i=1}^{n} (\rho \cdot x_i + b - y_i)^2 = F(\rho; b).$$

Найдем min функции $F(\rho; b)$:

$$\begin{cases} \dfrac{\partial F}{\partial \rho} = 2 \cdot \sum_{i=1}^{n} (\rho \cdot x_i + b - y_i) \cdot x_i = 0 \\ \dfrac{\partial F}{\partial b} = 2 \cdot \sum_{i=1}^{n} (\rho \cdot x_i + b - y_i) = 0 \end{cases},$$

следовательно,
$$\begin{cases} \rho \cdot \sum x_i^{\,2} + b \cdot \sum x_i - \sum x_i \cdot y_i = 0 \\ \rho \cdot \sum x_i + b \cdot n - \sum y_i = 0 \end{cases}.$$

Решая эту систему, находим ρ и b, подставляя которые в уравнение (3), получаем выборочное уравнение прямой регрессии.

Пример. Найти уравнение прямой линии регрессии Y на X по данным значениям X и Y.

X	0,5	1,5	2,5	3,5	4,5	5,5
Y	4	6	6,5	7,0	8,0	8,5

$n = 6$

Составим таблицу:

x_i	y_i	x_i^2	$x_i \cdot y_i$
0,5	4	0,25	2
1,5	6	2,25	9
2,5	6,5	6,25	16,25
3,5	7,0	12,25	24,5
4,5	8,0	20,25	36
5,5	8,5	30,25	46,75
$\sum = 18$	$\sum = 40$	$\sum = 71{,}5$	$\sum = 134{,}5$

Следовательно, $\begin{cases} 71{,}5 \cdot \rho + 18 \cdot b - 134{,}5 = 0 \\ 18 \cdot \rho + 6 \cdot b - 40 = 0, \end{cases}$

$$9 \cdot \rho + 3 \cdot b = 20 \text{, следовательно, } b = \frac{20 - 9 \cdot b}{3},$$

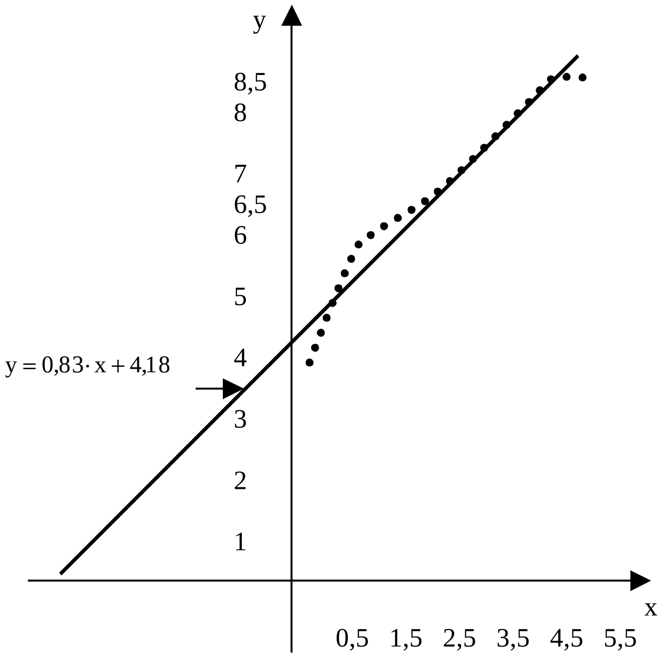
$$71{,}5 \cdot \rho + 6 \cdot (20 - 9 \cdot \rho) = 134{,}5,$$

$$17{,}5 \cdot \rho = 14{,}5,$$

$\rho \approx 0{,}83$, следовательно, $b \approx \dfrac{20 - 9 \cdot 0{,}83}{3} \approx 4{,}18$, получим искомое уравнение: $y = 0{,}83 \cdot x + 4{,}18$.

Предположим, что в некотором опыте наблюдается две случайные величины X и Y. То обстоятельство, что X и Y обусловлены одним и тем же опытом, создает между этими величинами некоторого рода связь: как принято говорить, X и Y скоррелированы (согласованы) друг с другом.

Для описания системы двух случайных величин, кроме математических ожиданий и дисперсий составляющих, используют и другие характеристики: к их числу относятся корреляционный момент и коэффициент корреляции.



Корреляционным моментом μ_{xy} случайных величин X и Y называют математическое ожидание произведения отклонений этих величин: $\mu_{xy} = M\left[(X - M(X))\cdot(Y - M(Y))\right]$.

Корреляционный момент, как свидетельствует его назначение (от латинского слова *correlatio* – соответствие, взаимосвязь), играет определенную роль при оценке зависимости X и Y. Основное свойство корреляционного момента выражается следующим предложением.

Если величины X и Y независимы, то их корреляционный момент равен нулю.

Напомним, что случайные величины X и Y называются независимыми, если закон распределения одной из них не зависит от того, какие возможные значения приняла другая величина. В противном случае случайные величины зависимы.

Из определения корреляционного момента следует, что он имеет размерность, равную произведению разностей величин X и Y. Другими словами, величина корреляционного момента зависит от единиц измерения случайных величин. Такая особенность корреляционного момента является недостатком этой числовой характеристики, поскольку сравнение корреляционных моментов различных систем случайных величин становится затруднительным. Для того чтобы устранить этот недостаток, вводят новую числовую характеристику – коэффициент корреляции.

Коэффициентом корреляции r случайных величин X и Y называют отношение корреляционного момента к произведению средних квадратических отклонений этих величин:

$$r = \frac{\mu_{xy}}{\sigma_x \cdot \sigma_y}.$$

Так как размерность μ_{xy} равна произведению размерностей X и Y, σ_x имеет размерность величины X, σ_y имеет размерность величины Y, то r – безразмерная величина. Он не зависит также и от выбора начал отсчета при изменении X и Y.

Коэффициент корреляции всегда заключен между –I и I: $-I \le r \le I$.

В случае, когда $r = \pm I$, величины X и Y связаны линейной функциональной зависимостью: $y = \rho_{\frac{y}{x}} x + b$. Угловой коэффициент $\rho_{\frac{y}{x}}$ прямой линии регрессии X и Y называют выборочным коэффициентом регрессии Y на X; он является оценкой коэффициента регрессии.

При большом числе наблюдений одно и то же значение *x* может встретиться n_x раз, одно и то же значение y – n_y раз, одна и та же пара чисел (x; y) может наблюдаться n_{xy} раз. Поэтому данные наблюдений группируют, то есть подсчитывают частоты n_x, n_y, n_{xy}. Все сгруппированные данные записывают в виде таблицы, которую называют корреляционной.

Для определения параметров уравнения прямой линии регрессии Y на X с помощью метода наименьших квадратов получается система уравнений

$$\begin{cases} \rho_{yx} \cdot \sum x^2 + b\sum x = \sum xy \\ \rho_{yx} \cdot \sum x + nb = \sum y \end{cases} \qquad (1)$$

в предположении, что значения X и соответствующие им значения Y наблюдались по одному разу. Запишем систему (1) так, чтобы она отражала данные корреляционной таблицы. Воспользуемся тождествами:

$$\sum x = n\bar{x} \qquad \text{(следствие из } \bar{x} = \frac{\sum x}{n}\text{)},$$

$$\sum y = n\bar{y} \qquad \text{(следствие из } \bar{y} = \frac{\sum y}{n}\text{)},$$

$$\sum x^2 = n\overline{x^2} \qquad \text{(следствие из } \overline{x^2} = \frac{\sum x^2}{n}\text{)},$$

$\sum xy = \sum n_{xy} \cdot xy$ (учтено, что пара чисел (x; y) наблюдалась n_{xy} раз).

Подставив правые части тождеств в систему (1) и сократив обе части второго уравнения на n, получим

$$\begin{cases} \rho_{yx} n \overline{x^2} + bn\bar{x} = \sum n_{xy} xy \\ \rho_{yx} \bar{x} + b = \bar{y} \end{cases} \qquad (2)$$

Решив эту систему, найдем параметры ρ_{yx} и b и, следовательно, искомое уравнение

$$\bar{y}_x = \rho_{yx} x + b \,.$$

Целесообразно, введя новую величину – выборочный коэффициент корреляции, написать уравнение регрессии в ином виде. Найдем b из второго уравнения системы (2): $b = \bar{y} - \rho_{yx} \cdot \bar{x}$. Подставив правую часть этого равенства в уравнение $\bar{y}_x = \rho_{yx} x + b$, получим

$$\bar{y}_x - \bar{y} = \rho_{yx}(x - \bar{x}) \,. \qquad (3)$$

Найдем из системы (1) коэффициент регрессии, учитывая, что $\overline{x^2} - (\bar{x})^2 = \sigma_x^2$:

$$\rho_{yx} = \frac{\sum n_{xy} xy - n\bar{x}\bar{y}}{n\left[\overline{x^2} - (\bar{x})^2\right]} = \frac{\sum n_{xy} xy - n\bar{x}\bar{y}}{n\sigma_x^2} \,.$$

Умножим обе части равенства на дробь $\frac{\sigma_x}{\sigma_y}$:

$$\rho_{yx} \cdot \frac{\sigma_x}{\sigma_y} = \frac{\sum n_{xy} xy - n\bar{x}\bar{y}}{n \cdot \sigma_x \cdot \sigma_y} \,. \qquad (4)$$

Обозначим правую часть равенства через $r_в$ и назовем ее выборочным коэффициентом корреляции, который является оценкой коэффициента корреляции

$$r_{â} = \frac{\sum n_{xy} xy - n\bar{x}\bar{y}}{n \cdot \sigma_x \cdot \sigma_y} \,.$$

Подставим $r_в$ в (4) $\rho_{xy} \cdot \frac{\sigma_x}{\sigma_y} = r_в$. Отсюда $\rho_{xy} = r_в \cdot \frac{\sigma_y}{\sigma_x}$. Подставим правую часть этого равенства в (3), окончательно получим выборочное уравнение прямой линии регрессии Y на X вида

$$\overline{y}_x - \overline{y} = r_в \cdot \frac{\sigma_y}{\sigma_x}(x - \overline{x}). \qquad (5)$$

Замечание. Аналогично находят выборочное уравнение прямой линии регрессии X на Y вида

$$\overline{x}_y - \overline{x} = r_в \cdot \frac{\sigma_x}{\sigma_y}(y - \overline{y}), \qquad (6)$$

где $r_в \cdot \frac{\sigma_x}{\sigma_y} = \rho_{xy}$.

Если данные наблюдений над признаками X и Y заданы в виде корреляционной таблицы с равноотстоящими вариантами, то целесообразно перейти к условным вариантам: $u_i = \frac{x_i - C_1}{h_1}$, $v_j \frac{y_i - C_2}{h_2}$, где C_1 – «ложный нуль» вариант X (новое начало отсчета); в качестве ложного нуля выгодно принять варианту, которая расположена примерно в середине вариационного ряда (условимся принимать в качестве ложного нуля варианту, имеющую наибольшую частоту); h_1 – шаг, то есть разность между двумя соседними вариантами X; C_2 – ложный нуль вариант Y; h_2 – шаг вариант Y.

В этом случае выборочный коэффициент корреляции

$$r_в = \frac{\sum n_{uv} \cdot u \cdot v - n \cdot \overline{u} \cdot \overline{v}}{n \cdot \sigma_u \cdot \sigma_v},$$

причем слагаемое $\sum n_{uv} \cdot u \cdot v$ удобно вычислять, используя табл. 3 (см. решение примера).

Величины $\overline{u}$, $\overline{v}$, σ_u, σ_v могут быть найдены либо методом произведений (при большом числе данных), либо непосредственно по формулам:

$$\overline{u} = \frac{\sum n_u \cdot u}{n}; \ \overline{v} = \frac{\sum n_v \cdot v}{n}; \ \sigma_u = \sqrt{\overline{u^2} - (\overline{u})^2}; \ \sigma_v = \sqrt{\overline{v^2} - (\overline{v})^2}.$$

Зная эти величины, можно определить входящие в уравнения регрессии (5) и (6) величины по формулам:

$$\overline{x} = \overline{u} \cdot h_1 + C_1; \ \overline{y} = \overline{v} \cdot h_2 + C_2; \ \sigma_x = \sigma_u \cdot h_1; \ \sigma_y = \sigma_v \cdot h_2.$$

Пример на отыскание выборочного уравнения прямой линии регрессии X на Y по данным корреляционной табл. 1.

Таблица 1

Y	X								
	5	10	15	20	25	30	35	40	n_y
100	2	1	-	-	-	-	-	-	3
120	3	4	3	-	-	-	-	-	10
140	-	-	5	10	8	-	-	-	23
160	-	-	-	1	-	6	1	1	9
180	-	-	-	-	-	-	4	1	5
n_x	5	5	8	11	8	6	5	2	n = 50

Перейдем к условным вариантам $u_i = \dfrac{x_i - C_1}{h_1}$, $v_i = \dfrac{y_i - C_2}{h_2}$. В качестве ложного нуля C_1 возьмем варианту x = 20, расположенную примерно в середине вариационного ряда и имеющую наибольшую частоту; шаг h_1 равен разности между двумя соседними вариантами: h_1=25–20=5. В качестве ложного нуля C_2 возьмем варианту y = 140, расположенную в середине вариационного ряда и имеющую наибольшую частоту; шаг h_2 равен разности между двумя соседними вариантами: h_2 = 120 – 100 = 20.

Итак, C_1=20, C_2=140, h_1=5, h_2=20.

Составим корреляционную таблицу в условных вариантах. Практически это делается так: в первом столбце вместо ложного нуля C_2 (варианта 140) пишут 0; над нулем последовательно записывают –1, –2; под нулем пишут 1, 2. В первой строке вместо ложного нуля C_1 (варианта 20) пишут 0; слева от нуля последовательно записывают –1, –2, –3; справа от нуля пишут 1, 2, 3, 4. Все остальные данные переписывают из первоначальной корреляционной таблицы. В итоге получим корреляционную табл. 2 в условных вариантах.

Таблица 2

v	u								
	-3	-2	-1	0	1	2	3	4	n_v
-2	2	1	-	-	-	-	-	-	3
-1	3	4	3	-	-	-	-	-	10

0	-	-	5	10	8	-	-	-	23
1	-	-	-	1	-	6	1	1	9
2	-	-	-	-	-	-	4	1	5
n_u	5	5	8	11	8	6	5	2	n = 50

Теперь для вычисления искомой суммы $\sum n_{uv} \cdot u \cdot v$ составим расчетную табл. 3. Пояснения к составлению табл. 3.

1. В каждой клетке, в которой частота $n_{uv} \neq 0$, записывают в правом верхнем углу произведение частоты n_{uv} на варианту $\bar{u}$. Например, в правых верхних углах клеток первой строки записаны произведения $2 \cdot (-3) = -6$; $1 \cdot (-2) = -2$.

2. Складывают все числа, помещенные в правых верхних углах клеток одной строки, и их сумму записывают в клетку этой же строки столбца U. Например, для первой строки U = –6 + (–2) = – 8.

3. Умножают варианту v на U и полученное произведение записывают в последнюю клетку этой же строки, т.е. в клетку столбца vU. Например, в первой строке табл. 3 v = –2, U = – 8; следовательно, $vU = (-2) \cdot (-8) = 16$.

4. Наконец, сложив все числа столбца vU, получают сумму $\sum_v vU$, которая равна искомой сумме $\sum n_{uv} uv$. Например, для табл. 3 имеем $\sum_v vU = 87$; следовательно, искомая сумма $\sum n_{uv} uv = 87$.

Для контроля аналогичные вычисления производят по столбцам: произведения $n_{uv} \cdot v$ записывают в левый нижний угол клетки, содержащей частоту $n_{uv} \neq 0$; все числа, помещенные в левых нижних углах клеток одного столбца, складывают и их сумму записывают в строку V; далее умножают каждую варианту u на V и результат записывают в клетках последней строки.

Наконец, сложив все числа последней строки, получают сумму $\sum_u uV$, которая также равна искомой сумме $\sum n_{uv} uv$. Например, для табл. 3 имеем $\sum_u uV = 87$; следовательно, $\sum n_{uv} uv = 87$.

Для нахождения выборочного коэффициента корреляции вычислим величины $\bar{u}$, $\bar{v}$, σ_u, σ_v. Их можно вычислить методом произведений;

однако, поскольку числа u_i, v_i малы, вычислим $\bar{u}$ и $\bar{v}$, исходя из определения средней, а σ_u и σ_v – используя формулы

$$\sigma_u = \sqrt{\overline{u^2} - (\bar{u})^2}, \qquad \sigma_v = \sqrt{\overline{v^2} - (\bar{v})^2}.$$

Таблица 3

v	u								U = $\sum n_{uv} u$	vU
	-3	-2	-1	0	1	2	3	4		
-2	- 6 2 -4	-2 1 -2	-	-	-	-	-	-	-8	16
-1	-9 3 -3	-8 4 -4	-3 3 -3	-	-	-	-	-	-20	20
0	-	-	-5 5 0	0 10 0	8 8 0	-	-	-	3	0
1	-	-	-	0 1 1	-	12 6 6	3 1 1	4 1 1	19	19
2	-	-	-	-	-	-	12 4 8	4 1 2	16	32
V = $\sum n_{uv} v$	-7	-6	-3	1	0	6	9	3		$\sum_v vU = 87$
uV	21	12	3	0	0	12	27	12	$\sum_u uV = 87$	контроль

Найдем $\bar{u}$ и $\bar{v}$, используя табл. 2

$$\bar{u} = \frac{\sum n_u u}{n} = \frac{5\cdot(-3)+5\cdot(-2)+8\cdot(-1)+11\cdot 0+8\cdot 1+6\cdot 2+5\cdot 3+2\cdot 4}{50} = 0{,}2\,;$$

$$\bar{v} = \frac{\sum n_v v}{n} = \frac{3\cdot(-2)+10\cdot(-1)+23\cdot 0+9\cdot 1+5\cdot 2}{50} = 0{,}06\,.$$

Вычислим вспомогательные величины $\overline{u^2}$ и $\overline{v^2}$:

$$\overline{u^2} = \frac{\sum n_u u^2}{n} = \frac{5\cdot 9+5\cdot 4+8\cdot 1+11\cdot 0+8\cdot 1+6\cdot 4+5\cdot 9+2\cdot 16}{50} = 3{,}64\,;$$

$$\overline{v^2} = \frac{\sum n_v v^2}{n} = \frac{3\cdot 4+10\cdot 1+23\cdot 0+9\cdot 1+5\cdot 4}{50} = 1{,}02\,.$$

Посчитаем σ_u и σ_v:

$$\sigma_u = \sqrt{\overline{u^2}-(\bar{u})^2} = \sqrt{3{,}64-0{,}04} = 1{,}897\,;$$

$$\sigma_v = \sqrt{\overline{v^2}-(\bar{v})^2} = \sqrt{1{,}02-0{,}0036} = 1{,}008\,.$$

Найдем искомый выборочный коэффициент корреляции:

$$r_{\hat{a}} = \frac{\sum n_{uv}\cdot u\cdot v - n\cdot\bar{u}\cdot\bar{v}}{n\cdot\sigma_u\cdot\sigma_v} = \frac{87-50\cdot 0{,}2\cdot 0{,}06}{50\cdot 1{,}897\cdot 1{,}008} = 0{,}904\,.$$

Учитывая, что $C_1 = 20$, $C_2 = 140$, $h_1=5$, $h_2 = 20$, проведем вычисления $\bar{x}$ и $\bar{y}$:

$$\bar{x} = \bar{u}\cdot h_1 + C_1 = 0{,}2\cdot 5 + 20 = 24\,;$$

$$\bar{y} = \bar{v}\cdot h_2 + C_2 = 0{,}06\cdot 20 + 140 = 141{,}2\,.$$

Найдем σ_x и σ_y:

$$\sigma_x = h_1\cdot\sigma_u = 5\cdot 1{,}897 = 9{,}485\,;\ \ \sigma_y = h_2\cdot\sigma_v = 20\cdot 1{,}008 = 20{,}16\,.$$

Выборочное уравнение прямой линии регрессии Y на X имеет вид:

$$\bar{y}_x - \bar{y} = r_{\hat{a}}\cdot\frac{\sigma_y}{\sigma_x}\cdot(x-\bar{x}).$$

Подставив найденные величины в уравнение (1), получим искомое уравнение:

$$\bar{y}_x - 141{,}2 = 0{,}904\cdot\frac{20{,}16}{9{,}485}\cdot(x-24);$$

или окончательно

$$\bar{y}_x = 1{,}92x + 100{,}86\,.$$

10. НЕЛИНЕЙНАЯ КОРРЕЛЯЦИЯ

Цель работы: исследование нелинейной корреляционной зависимости случайной величины Y от неслучайной величины X по результатам эксперимента.

Задание. Для установления корреляционной зависимости между величинами X и Y (где Y – случайная величина, X – неслучайная величина) проведены эксперименты, результаты которых записаны в табл. 1. Требуется: 1. Найти условные средние $\overline{y_i}$ и построить эмпирическую линию регрессии Y по X (ломанную). 2. Найти уравнение регрессии Y по X методом наименьших квадратов, принимая в качестве сглаженной линии параболу

$$\hat{y}_x = a \cdot x^2 + b \cdot x + c \tag{1}$$

и затем построить ее на одном чертеже с эмпирической линией регрессии. 3. Оценить тесноту корреляционной зависимости Y по X. 4. Проверить адекватность уравнения регрессии Y по X.

Таблица 1

Исходные данные

x_i	$x_1=1$	$x_2=2$	$x_3=3$	$x_4=4$	$x_5=5$	$x_6=6$
	10	10	19	30	45	79
	9	16	21	28	44	75
y_{ij}	11	15	18	31	42	80
	12	13	20	28	48	73
	10	14	18	33	43	82

Решение

1. Условные средние вычислим по формуле

$$\overline{y_i} = \frac{1}{5}\sum_{j=1}^{5} y_{ij} \tag{2}$$

Например, $\overline{y_1} = \frac{1}{5} \cdot$ (10+9+11+12+10)=10,4.

Результаты вычислений представлены в табл. 2

Таблица 2

Условные средние $\overline{y_i}$

x_i	1	2	3	4	5	6
$\overline{y_i}$	10,4	13,6	19,2	30,0	44,4	77,8

По точкам (x_i, $\overline{y_i}$) строим эмпирическую (ломанную) линию регрессии (рис. 1).

2. Теоретическое уравнение регрессии Y по X будем искать в виде (1). Неизвестные параметры a, b, c будем находить из системы линейных уравнений

$$\begin{cases} a\cdot\sum\limits_{i=1}^{n} x_i^4 + b\cdot\sum\limits_{i=1}^{n} x_i^3 + c\cdot\sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i^2\cdot\overline{y_{x_i}}, \\ a\cdot\sum\limits_{i=1}^{n} x_i^3 + b\cdot\sum\limits_{i=1}^{n} x_i^2 + c\cdot\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i\cdot\overline{y}_{x_i}, \\ a\cdot\sum\limits_{i=1}^{n} x_i^2 + b\cdot\sum\limits_{i=1}^{n} x_i + c\cdot n = \sum\limits_{i=1}^{n} \overline{y}_{x_i}; \end{cases} \qquad (3)$$

в которой n = 6. Для определения коэффициентов этой системы составим расчетную табл. 3, где в последней строке записаны суммы элементов соответствующих столбцов.

Используя последнюю строку табл. 2, имеем: $\sum\limits_{j=1}^{5} x_i^4$ =2275; $\sum\limits_{j=1}^{6} x_i^2\cdot\overline{y_i}$ =4628,4 и т.д.

Таблица 3

Расчетная таблица

i	x_i	$\overline{y_i}$	x_i^2	x_i^3	x_i^4	$x_i\overline{y_i}$	$x_i^2\overline{y_i}$
1	1	10,4	1	1	1	10,4	10,4
2	2	13,6	4	8	16	27,2	54,4
3	3	19,2	9	27	81	57,6	172,8
4	4	30,0	16	64	256	120,0	480,0
5	5	44,4	25	125	625	222,0	1110,0
6	6	77,8	36	216	1296	466,8	2800,8
$\sum$	21	195,4	91	441	2275	904	4628,4

В результате получаем систему:

$$\begin{cases} 2275\cdot a+441\cdot b+91\cdot c=4628{,}4, \\ 441\cdot a+91\cdot b+21\cdot c=904, \\ 91\cdot a+21\cdot b+6\cdot c=195{,}4. \end{cases} \tag{4}$$

Систему (4) решаем методом Гаусса. Первое уравнение системы (4) делим на 2275. Получаем уравнение (*). Второе уравнение системы (4) разделим на 441. И из полученного результата вычтем уравнение (*). получим уравнение (**). третье уравнение системы (4) разделим на 91, и из полученного результата вычтем уравнение (*). Имеем линейную систему:

$$\begin{cases} a+0{,}193846\cdot b+0{,}04\cdot c=2{,}034461, & (*) \\ \quad 0{,}0125032\cdot b+0{,}007619047\cdot c=0{,}0154256, & (**) \\ \quad 0{,}03692323\cdot b+0{,}02593406\cdot c=0{,}1127917. & (***) \end{cases}$$

Теперь уравнение (**) разделим на 0,0125032. Получим уравнение (α) уравнение (***) разделим на 0,03692323, и из полученного результата вычтем уравнение (α) получим уравнение (β) имеем систему линейных уравнений:

$$\begin{cases} a+0{,}193846\cdot b+0{,}04\cdot c=2{,}034461, & (*) \\ \qquad b+0{,}609364\cdot c=1{,}233732, & (\alpha) \\ \qquad\quad 0{,}09301282\cdot c=0{,}1127917. & (\beta) \end{cases}$$

Из уравнения (β) находим c, подставляя его в (α), из уравнения (α) находим b. Подставляя найденные b и c в (*), находим a:

$$\begin{cases} c=19{,}57830\approx 19{,}578, \\ b=-10{,}69659\approx -10{,}697, \\ \text{à}=3{,}32432\approx 3{,}325. \end{cases}$$

Найденные a, b, c подставляем в (1), получим теоретическое уравнение регрессии:

$$\hat{y}_x=3{,}325\cdot x^2-10{,}697\cdot x+19{,}578\,. \tag{5}$$

Вычислив ординаты теоретической линии регрессии по формуле (5) для значений X, заданных в таблице исходных данных 1 (эти результаты записаны в табл. 4), строим теоретическую линию регрессии на одном чертеже с эмпирической линией регрессии (рис. 1).

Таблица 4

Ординаты теоретической линии регрессии

x_i	1	2	3	4	5	6
$\hat{y}_i$	12,207	11,484	17,411	30,000	49,216	75,093

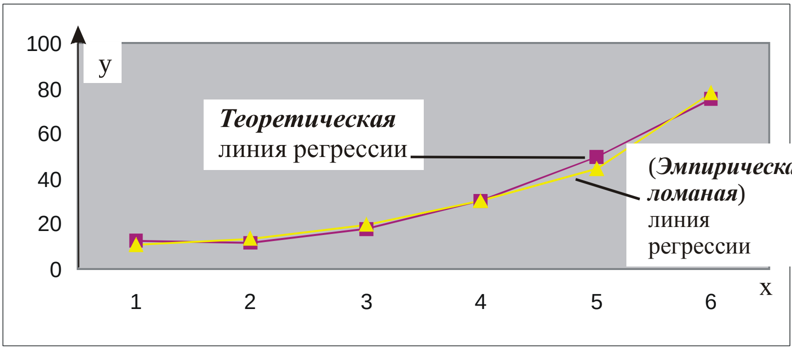


Рис. 1. Эмпирическая ломаная и теоретическая линии регрессии

Наглядно убеждаемся, что теоретическая линия регрессии хорошо сглаживает эмпирическую линию регрессии, чем подтверждается, что система (4) составлена и решена верно.

3. Оценим тесноту корреляционной зависимости.

а) находим общее среднее по формуле (2):

$$\bar{y} = \frac{1}{6}\sum_{j=1}^{6}\bar{y}_i = \frac{1}{6}\cdot(10{,}4+\ldots+77{,}8)=32{,}57.$$

Здесь значения $\bar{y}_i$ взяты из табл. 2.

б) Находим внешнюю дисперсию $S^2_{âíåø.}$ по формуле:

$$S^2(\bar{y}_x) = S^2_{âíåø.} = \frac{1}{n}\sum_{j=1}^{n}\bar{y}_i^{\,2} - \bar{y}^2, \qquad (6)$$

где $$\bar{y} = \frac{1}{n}\sum_{j=1}^{n}\bar{y}_i, \qquad (7)$$

$$S^2_{âíåø.} = \frac{1}{6}\sum_{j=1}^{6}\bar{y}_i^{\,2} - \bar{y}^2 = \frac{1}{6}\cdot(10.4^2+\ldots+77.8^2) - 32.57^2 = 537.$$

в) Для нахождения усредненной внутренней дисперсии $S^2_{\hat{a}\acute{i}\acute{o}\grave{o}\eth.}$ сначала вычисляем внутренние дисперсии S_i^2 для каждого x_i по формуле:

$$S^2_{\hat{a}\acute{i}\acute{o}\grave{o}\eth.} = \frac{1}{n}\sum_{i=1}^{n} S_i^2(y), \tag{8}$$

где

$$S_i^2(y) = \frac{1}{m}\sum_{j=1}^{n}\left(y_{ij} - \overline{y}_i\right)^2, \tag{9}$$

в которой m = 5:

$$S_i^2(y) = \frac{1}{5}\sum_{j=1}^{5} y_{ij}^2 - \overline{y}_i^{\,2}, \tag{10}$$

$$i = 1, 2, 3, 4, 5, 6.$$

Для примера вычислим S_1^2, составив предварительно расчетную таблицу (табл. 5).

Таблица 5

Расчетная таблица S_1^2

y_{1j}	10	9	11	12	10	
y_{1j}^2	100	81	121	144	100	$\sum y_{1j}^2$ =546

$$S_1^2 = \frac{1}{5}\sum_{j=1}^{5} y_{1j}^2 - \overline{y}_1^{\,2} = \frac{1}{5}\cdot 546 – 10.4^2 = 1{,}04.$$

Значения всех внутренних дисперсий S_i^2 для каждого x_i представлены в табл. 6.

Таблица 6

Значения дисперсий S_i^2

x_i	1	2	3	4	5	6
S_i^2	1,04	4,24	1,36	3,60	4,24	10,26

Усредненная внутренняя дисперсия

$$S^2_{\text{внутр.}} = \frac{1}{6}\sum_{i=1}^{6} S_i^2 = \frac{1}{6}\cdot(1{,}04+4{,}24+\ldots+10{,}26) = 4{,}24.$$

г) Вычисляем корреляционное соотношение по формуле (11):

$$\eta = \sqrt{\frac{S^2_{\text{внеш.}}}{S^2_{\text{внеш.}} + S^2_{\text{внутр.}}}}\,, \qquad (11)$$

$$\eta = \sqrt{\frac{537}{537 + 4{,}24}} = 0{,}996.$$

Вывод: корреляционная зависимость – тесная.

4. Проверяем адекватность уравнения регрессии.

а) Вычисляем остаточную дисперсию по формуле (12):

$$S^2_{\text{ост.}} = \frac{1}{n-l}\sum_{i=1}^{n}\left(\overline{y}_i - \hat{y}_i\right)^2, \qquad (12)$$

$$S^2_{\text{ост.}} = \frac{(0.4-12.207)^2 + (3{,}6-11{,}484)^2 + \ldots + (7{,}8-75{,}093)^2}{6-3} = \frac{31{,}21}{3} =$$

$$=10{,}40.$$

Здесь n (число значений X) равно 6, l – количество параметров уравнения регрессии (5) – равно 3, значения $\overline{y}_i$ взяты из табл. 2, значения $\hat{y}_i$ взяты из табл. 4.

б) Вычисляем «дисперсию воспроизводимости средних» по формуле (13):

$$S^2_{\text{воспр.ср.}} = \frac{1}{m}\cdot\frac{m}{m-1}\; S^2_{\text{внутр.}} = \frac{m}{m-1}\; S^2_{\text{внутр.}}, \qquad (13)$$

$$S^2_{\text{воспр.ср.}} = \frac{1}{5-1}\cdot 4{,}24=1{,}06.$$

в) Вычисляем величину $F_{\text{эмп.}}$ по формуле (14):

$$F_{\text{эмп.}} = \frac{S^2_{\text{ост.}}}{S^2_{\text{воспр.ср.}}}\,, \qquad (14)$$

$$F_{\text{эмп.}} = \frac{10{,}40}{1{,}06} = 9{,}8.$$

г) Величина $F_{эмп.}$ имеет распределение Фишера с $k_1 = n - l$ и $k_2 = n \cdot (m-1)$ числами степеней свободы (n – число задаваемых экспериментатором значений величины X, m – число проводимых опытов, l – число коэффициентов в уравнении регрессии).

По заданному уровню значимости $\alpha = 0{,}05$ и числом степеней свободы $k_1 = n - l$ и $k_2 = n \cdot (m-1)$ из таблицы критических точек распределения Фишера находим $F_{крит}$.

Если $F_{эмп.} < F_{крит.}$, уравнение регрессии адекватно.

Если $F_{эмп.} > F_{крит.}$, расхождение между теоретической и эмпирической линиями регрессии значимо, уравнение неадекватно, следует взять многочлен более высокого порядка.

Находим $F_{крит.}$ при уровне значимости $\alpha = 0{,}05$.

В нашем случае n (число значений X) равно 6, m (число значений Y в каждом столбце табл. 1) равно 5, l (число параметров уравнения регрессии(5)) равно 3. Следовательно, числа степеней свободы соответственно равны $k_1 = n - l = 6 - 3 = 3$, $k_2 = n \cdot (m-1) = 6 \cdot 4 = 24$. По таблице критических точек распределения Фишера для $\alpha = 0{,}05$, $k_1 = 3$, $k_2 = 24$ находим $F_{крит} = 3{,}01$.

Так как 9,8>3,01, то $F_{эмп.} > F_{крит.}$.

Вывод: теоретическое уравнение регрессии неадекватно. В качестве уравнения регрессии необходимо принять многочлен 3-й степени $y = b_3 \cdot x^3 + b_2 \cdot x^2 + b_1 \cdot x + b_0$ и все расчеты сделать заново.

(В работе пересчета делать не надо).

11. КОНТРОЛЬНЫЕ ЗАДАНИЯ ПО ВЫСШЕЙ МАТЕМАТИКЕ

Задача № 1

1. В студии три телевизионные камеры. Для каждой камеры вероятность того, что она включена в данный момент, равна 0,6. Найти вероятность того, что в данный момент включена хотя бы одна камера.
2. Чему равна вероятность того, что при бросании трех игральных костей шесть очков появится хотя бы на одной из костей?
3. Предприятие изготавливает 95% стандартных деталей, причем из них 86% первого сорта. Найти вероятность того, что взятое наудачу изделие, изготовленное на этом предприятии, окажется первого сорта.
4. В коробке шесть одинаковых занумерованных кубиков. Наудачу по одному извлекают все кубики. Найти вероятность того, что номера извлеченных кубиков появится в возрастающем порядке.
5. Библиотека состоит из 10 различных книг, причем 5 книг стоят 4 рубля каждая 3 книги – по 1 рублю, 2 книги – по 3 рубля. Найти вероятность того, что взятые наудачу две книги стоят вместе 5 рублей.
6. Для некоторой местности среднее число ясных дней в июле равно 25. Найти вероятность того, что первые 2 дня июля будут ясными.
7. На каждой из пяти одинаковых карточек напечатана одна из следующих букв «а», «м», «р», «т», «ю». Карточки тщательно перемешаны. Найти вероятность того, что на четырех вынутых по одной карточке можно прочесть слово «юрта».
8. В замке на общей оси 5 дисков, каждый из которых разделен на 6 секторов с различными написанными на них буквами. Замок открывается только в том случае, если каждый диск занимает одно определенное положение относительно корпуса замка. Найти вероятность того, что при произвольной установке дисков замок можно открыть.
9. Четырехтомное сочинение расположено на полке в случайном порядке. Найти вероятность того, что тома стоят в должном порядке справа налево или слева направо.
10. На складе имеется 10% пальто размера 44; 20% – размера 46; 30% – размера 48; 25% – размера 50 и остальное пальто выше размера 50. Какова вероятность того, что наугад взятое пальто окажется: а) не более размера 48; б) не менее 48; в) размера 46 или 48?

Задача № 2

1. Три покупателя посетили магазин. Вероятность того, что они совершат покупку, соответственно для них равна 0,9; 0,7; 0,8. Найти веро-

ятность того, что: 1) все трое сделают покупки; 2) все трое ничего не купят; 3) только один из них совершит покупку; 4) хотя бы один из них что-то купит.

2. Четыре пловца взяли старт на соревнованиях по плаванию. Вероятность уложиться в рекордное время у первого пловца равна 0,95, у второго – 0,92, у третьего – 0,9, у четвертого – 0,88. Найти вероятность того, что: 1) все пловцы станут рекордсменами; 2) только два пловца станут рекордсменами.

3. Устройство состоит из 3 независимо работающих элементов. Вероятность безотказной работы в течение часа первого элемента равна 0,95; для второго – 0,98 и для третьего – 0,9. Найти вероятность того, что в течение часа будут работать: 1) 2 элемента; 2) 1 элемент; 3) все 3 элемента.

4. Некоторый механизм состоит из 3 деталей типа А, пяти деталей типа В, двух деталей типа С, шести деталей типа Д и четырех деталей типа Е. Вероятность повреждения детали типа А равна 0,02; типа В – 0,05; типа С – 0,1; типа Д – 0,03 и типа Е – 0,09. Механизм вышел из строя. Какого типа деталь вероятнее всего повреждена?

5. Вероятность появления брака на первом станке равна 0,02; на втором – 0,01; на третьем – 0,03. Производительность первого станка вдвое больше третьего, а производительность второго станка в четыре раза больше производительности первого станка. Детали, изготовленные на трех этих станках, хранятся на одном складе. Кладовщик взял наудачу одну деталь, она оказалась стандартной. На каком из станков вероятнее всего она была изготовлена?

6. В продажу поступило 1000 пальто с трех фабрик. С первой фабрики поступило 300 пальто, среди них 10 второсортных. Со второй фабрики поступило 450 пальто, среди них 12 второсортных. С третьей фабрики поступило 250 пальто, среди них 8 второсортных. Покупателю подали для примерки пальто первого сорта. На какой фабрике вероятнее всего пошито это пальто?

7. В первой коробке 35 радиоламп, среди них 4 нестандартные. Во второй коробке 20 радиоламп, среди них 1 нестандартная. В третьей коробке 45 радиоламп, в том числе 5 нестандартных. Из третьей коробки взяли наудачу 1 радиолампу и переложили в первую коробку. Затем из второй коробки была наудачу взята радиолампа и переложена в первую коробку. После этого из первой коробки наудачу извлекли радиолампу. Какова вероятность того, что эта лампа стандартная?

8. Рабочий обслуживает три станка, работающих независимо друг от друга. Вероятность того, что в течение часа первый станок потребует внимания рабочего, равна 0,02; для второго станка такая вероятность равна 0,1, а для третьего – 0,15. Какова вероятность того, что в течение одного часа: а) ни один из станков не потребует внимания рабочего;

б) все три станка потребуют внимания рабочего; в) какой-нибудь один станок потребует внимания рабочего?

9. Три стрелка произвели залп по цели. Вероятность попадания в цель первым стрелком равна 0,7; для второго и третьего стрелков эти вероятности соответственно равны 0,8 и 0,9. Найти вероятность того, что: а) только один из стрелков поразит цель; б) только 2 стрелка поразят цель; в) все три стрелка поразят цель.

10. Покупатель приобрел пылесос и полотер. Вероятность того, что пылесос выдержит гарантийный срок, равна 0,95, для полотера такая вероятность равна 0,9. Найти вероятность того, что гарантийный срок выдержат: а) оба прибора; б) только один из них; в) хотя бы один из них.

Задача № 3

1. Найти вероятность того, что событие А появится в пяти испытаниях не менее двух раз, если в каждом испытании вероятность появления события А равна 0,6.
2. Произведено 8 независимых испытаний, в каждом из которых вероятность появления события А равна 0,1. Найти вероятность того, что событие А появится: а) два раза; б) хотя бы два раза.
3. Монету бросают 5 раз. Найти вероятность того, что герб выпадает: а) менее трех раз; б) не менее трех раз.
4. В партии из 100 деталей 6 нестандартных. Из партии выбирается наугад 10 деталей. Определить вероятность того, что среди этих 10 деталей будет ровно две стандартные.
5. В партии 6 деталей, среди которых 4 стандартные. Какова вероятность того, что среди наудачу отобранных трех деталей: а) менее двух стандартных; б) не менее двух стандартных.
6. Производится 5 независимых испытаний, в каждом из которых вероятность непоявления события А равна 0,1. Найти вероятность того, что событие А появится: а) два раза; б) не менее четырех раз; в) менее четырех раз.
7. Устройство, состоящее из пяти независимо работающих элементов, включается за время T. Вероятность отказа каждого из них за это время равна 0,2. Найти вероятность того, что откажут: а) три элемента; б) не менее четырех элементов; в) хотя бы один элемент.
8. Вероятность хотя бы одного попадания при двух выстрелах равна 0,99. Найти вероятность четырех попаданий при пяти выстрелах.
9. В семье пять детей. Найти вероятность того, что среди этих детей не более двух мальчиков; б) не менее двух и не более трех мальчиков. Вероятность рождения мальчика принять равной 0,51.
10. По цели производится 5 независимых выстрелов. Вероятность попадания в цель при одном выстреле равна 0,4. Для получения зачета

по стрельбе требуется не менее трех попаданий. Найти вероятность получения зачета.

Задача № 4

Найти закон распределения указанной дискретной случайной величины X и ее функцию распределения F(x). Вычислить математическое ожидание M(X), дисперсию D(X) и среднее квадратическое отклонение $\sigma(X)$. Построить график функции распределения F(x).

1. Автомобиль должен проехать по улице, на которой установлено четыре независимо работающих светофора. Каждый светофор с интервалом в 2 минуты подает красный и зеленый сигналы. СВ X – число остановок автомобиля на этой улице.

2. Производится три выстрела по мишени. Вероятность поражения мишени первым выстрелом равна 0,4; вторым – 0,5; третьим – 0,6. СВ X – число поражений мишени.

3. Вероятность безотказной работы в течение гарантийного срока для телевизора первого типа равна 0,9; второго типа – 0,7; третьего типа – 0,8. СВ X – число телевизоров по одному из каждого типа, проработавших гарантийный срок.

4. Вероятность поражения цели при одном выстреле равна 0,6. СВ X – число поражений цели при четырех выстрелах.

5. Вероятность выпуска прибора, удовлетворяющего требованиям качества, равна 0,9. В контрольной партии 3 прибора. СВ X – число приборов, удовлетворяющих требованиям качества.

6. Вероятность перевыполнения плана СУ-1 равна 0,9; для СУ-2 – 0,8; для СУ-3 – 0,7. СВ X – число СУ перевыполнивших план.

7. 90% панелей, изготовляемых на железобетонном заводе, – высшего сорта. СВ X – число панелей высшего сорта из четырех, взятых наудачу.

8. Вероятность отказа прибора за время испытания на надежность равна 0,2. СВ X – число приборов, отказавших в работе, среди пяти испытываемых.

9. Вероятность поступления вызова на АТС в течение одной минуты равна 0,4. СВ X – число вызовов, поступивших на АТС за 4 минуты.

В первой коробке 10 сальников, из них два бракованных, во второй коробке 16 сальников, из них 4 бракованных, в третьей – 12, из них три бракованных. СВ X – число бракованных сальников при условии, что из каждой коробки взято по одному сальнику.

Задача № 5

В задачах 1–10 независимые случайные величины X и Y заданы законами распределения. Требуется:

1) составить закон распределения случайной величины Z ;
2) вычислить M(X), M(Y), M(Z), D(X), D(Y), D(Z);
3) проверить справедливость указанного свойства.

1.

x:	–2	1	4
y:	0,1	0,5	0,4

y:	0	2	3
p:	0,2	0,3	0,5

Z = X + Y,
3. D(X + Y) = D(X) + D(Y).

2.

x:	–1	0	2
y:	0,5	0,3	0,2

y:	–2	0	3
p:	0,2	0,5	0,3

Z = X + Y,
3. M(Z) = M(X) + M(Y).

3.

x:	–1	0	1
y:	0,2	0,5	0,4

y:	0	1	2	3
p:	0,1	0,2	0,3	0,4

Z = X + Y,
3. D(Z) = D(X) + D(Y).

4.

x:	–1	0	1
y:	0,3	0,2	0,5

y:	0	2	3
p:	0,1	0,3	0,6

Z = X + Y,
3. M(Z) = M(X) + M(Y).

5.

x:	–3	–1	1
y:	0,3	0,3	0,4

y:	–1	0	1
p:	0,2	0,1	0,7

$$Z = \frac{X+Y}{2},$$

$$3.\ D\left(\frac{X+Y}{2}\right) = \frac{1}{4}\left[D(X) + D(Y)\right].$$

6.

x:	–2	0	4
y:	0,1	0,2	0,7

y:	–4	–2	2
p:	0,2	0,4	0,4

$$Z = \frac{1}{2} \cdot (X + Y),$$

$$3.\ M(Z) = \frac{1}{2} \cdot \left[M(X) + M(Y)\right].$$

7.

x:	–4	0	4
y:	0,2	0,5	0,3

y:	0	2	4
p:	0,1	0,5	0,4

$$Z = \frac{X+Y}{2},$$

$$3.\ D\left(\frac{X+Y}{2}\right) = \frac{1}{4}\left[D(X) + D(Y)\right].$$

8.

x:	1	3	5
y:	0,1	0,5	0,4

y:	–1	1	3
p:	0,1	0,6	0,3

$$Z = \frac{1}{2} \cdot (X + Y),$$

$$3.\ M(Z) = \frac{1}{2} \cdot \left[M(X) + M(Y)\right].$$

9.

x:	–3	–1	1
y:	0,3	0,3	0,4

y:	–1	0	1
p:	0,2	0,1	0,7

$$Z = X + Y,$$

$$3.\ D(X + Y) = D(X) + D(Y).$$

10.

x:	1	0	2
y:	0,1	0,3	0,6

y:	–2	0	3
p:	0,2	0,5	0,3

Z + X + Y,
3. D(X + Y) = D(X) + D(Y).

Задача № 6

Дана функция распределения F(x) случайной величины X. Найти плотность распределения вероятностей f(x), математическое ожидание M(X), дисперсию D(X) и вероятность попадания случайной величины X на отрезок [a; b]. Построить графики функций F(x) и f(x).

$$1.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \frac{1}{8}\cdot x^3,\ \text{åñëè}\ \ 0\le x\le 2, \\ 1,\ \text{åñëè}\ \ x>2. \quad a=0,\ b=1. \end{cases}$$

$$2.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \frac{1}{33}\cdot(2\cdot x+3\cdot x^2),\ \text{åñëè}\ \ 0\le x\le 3, \\ 1,\ \text{åñëè}\ \ x>3. \quad a=1,\ b=2. \end{cases}$$

$$3.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \frac{1}{9}\cdot x^2,\ \text{åñëè}\ \ 0\le x\le 3, \\ 1,\ \text{åñëè}\ \ x>3. \quad a=0,\ b=1. \end{cases}$$

$$4.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \frac{1}{24}\cdot(x^2+2\cdot x),\ \text{åñëè}\ \ 0\le x\le 4, \\ 1,\ \text{åñëè}\ \ x>0. \quad a=0,\ b=1. \end{cases}$$

$$5.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \frac{1}{10}\cdot(x^3+x),\ \text{åñëè}\ \ 0\le x\le 2, \\ 1,\ \text{åñëè}\ \ x>2. \quad a=0,\ b=1. \end{cases}$$

$$6.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \dfrac{1}{20}\cdot(x^2+x),\ \ \text{åñëè}\ \ 0\le x\le 4, \\ 1,\ \text{åñëè}\ \ x>4. \quad a=0,\ b=3. \end{cases}$$

$$7.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<\dfrac{3\pi}{4}, \\ \cos 2x,\ \ \text{åñëè}\ \ \dfrac{3\pi}{4}\le x\le \pi, \\ 1,\ \text{åñëè}\ \ x>\pi. \quad a=\dfrac{3\pi}{4},\ b=\dfrac{5\pi}{4}. \end{cases}$$

$$8.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ 1-\cos x,\ \ \text{åñëè}\ \ 0\le x\le \dfrac{\pi}{2}, \\ 1,\ \text{åñëè}\ \ x>\dfrac{\pi}{2}. \quad a=0,\ b=\dfrac{\pi}{3}. \end{cases}$$

$$9.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \dfrac{1}{96}\cdot(x^3+8x),\ \ \text{åñëè}\ \ 0\le x\le 4, \\ 1,\ \text{åñëè}\ \ x>4. \quad a=0,\ b=2. \end{cases}$$

$$10.\ F(x)=\begin{cases} 0,\ \text{åñëè}\ \ x<0, \\ \dfrac{1}{9}\cdot(x+1)^2,\ \ \text{åñëè}\ \ -1\le x\le 2, \\ 1,\ \text{åñëè}\ \ x>2. \quad a=1,\ b=2. \end{cases}$$

Задача № 7

Найти доверительный интервал для оценки математического ожидания a нормального распределения с надежностью 0,95, зная выборочную среднюю $\bar{x}$, объем выборки n и среднее квадратическое отклонение σ .

1. $\bar{x}$ =75,17, n=36, σ =6.
2. $\bar{x}$ =75,16, n=49, σ =7.
3. $\bar{x}$ =75,15, n=64, σ =8.
4. $\bar{x}$ =75,14, n=81, σ =9.

5. $\bar{x}$ =75,13, n=100, σ =10.

6. $\bar{x}$ =75,12, n=121, σ =11.

7. $\bar{x}$ =75,11, n=144, σ =12.

8. $\bar{x}$ =75,10, n=169, σ =13.

9. $\bar{x}$ =75,09, n= 196, σ =14.

10. $\bar{x}$ =75,08, n=225, σ =15.

Задача № 8

Найти выборочное уравнение прямой линии регрессии Y на X по данным корреляционной таблицы.

1.

Y	X						
	10	15	20	25	30	35	n_y
30	2	6	-	-	-	-	8
40	-	4	4	-	-	-	8
50	-	-	7	35	8	-	50
60	-	-	2	10	8	-	20
70	-	-	-	5	6	3	14
n_x	2	10	13	50	22	3	n = 100

2.

Y	X						
	15	20	25	30	35	40	n_y
5	4	2	-	-	-	-	6
10	-	6	4	-	-	-	10
15	-	-	6	45	2	-	53
20	-	-	2	8	6	-	16
25	-	-	-	4	7	4	15
n_x	4	8	12	57	15	4	n = 100

3.

Y	X						
	5	10	15	20	25	30	n_y
20	1	5	-	-	-	-	6
30	-	5	3	-	-	-	8
40	-	-	9	40	2	-	51
50	-	-	4	11	6	-	21
60	-	-	-	4	7	3	14
n_x	1	10	16	55	15	3	n = 100

4.

Y	X							
	18	23	28	33	38	43	48	n_y
125	-	1	-	-	-	-	-	1
150	2	1	5	-	-	-	-	8
175	-	3	2	12	-	-	-	17
200	-	-	1	8	7	-	-	16
225	-	-	-	-	3	3	-	6
250	-	-	-	-	-	1	1	2
n_x	2	5	8	20	10	4	1	n = 50

5.

Y	X							
	5	10	15	20	25	30	35	n_y
100	-	-	-	-	-	6	1	7
120	-	-	-	-	-	4	2	6
140	-	-	8	10	5	-	-	23
160	3	4	3	-	-	-	-	10
180	2	1	-	1	-	-	-	4
n_x	5	5	11	11	5	10	3	n = 50

6.

Y	X						
	4	9	14	19	24	29	n_y
10	2	3	-	-	-	-	5
20	-	7	3	-	-	-	10
30	-	-	2	50	2	-	54
40	-	-	1	10	6	-	17
50	-	-	-	4	7	3	14
n_x	2	10	6	64	15	3	n = 100

7.

Y	X						
	10	15	20	25	30	35	n_y
6	4	2	-	-	-	-	6
12	-	6	2	-	-	-	8
18	-	-	5	40	5	-	50
24	-	-	2	8	7	-	17
30	-	-	-	4	7	8	19
n_x	4	8	9	52	19	8	n = 100

8.

Y	X						
	5	10	15	20	25	30	n_y
8	2	4	-	-	-	-	6
12	-	3	7	-	-	-	10
16	-	-	5	30	10	-	45
20	-	-	7	10	8	-	25
24	-	-	-	5	6	3	14
n_x	2	7	19	45	24	3	n = 100

9.

Y	X						
	2	7	12	17	22	27	n_y
10	2	4	-	-	-	-	6
20	-	6	2	-	-	-	8
30	-	-	3	50	2	-	55
40	-	-	1	10	6	-	17
50	-	-	-	4	7	3	14
n_x	2	10	6	64	15	3	n = 100

10.

Y	X						
	11	16	21	26	31	36	n_y
25	2	4	-	-	-	-	6
35	-	6	3	-	-	-	9
45	-	-	6	45	4	-	55
55	-	-	2	8	6	-	16
65	-	-	-	4	7	3	14
n_x	2	10	11	57	17	3	n = 100

Задача № 9

Цель работы: исследование нелинейной корреляционной зависимости случайной величины Y от неслучайной величины X по результатам эксперимента.

Задание. Для установления корреляционной зависимости между величинами X и Y (где Y – случайная величина, X – неслучайная величина) проведены эксперименты, результаты которых представлены в таблице. Требуется:

1. Найти условные средние $\overline{y_i}$ и построить эмпирическую линию регрессии Y по X (ломаную).

2. Найти уравнение регрессии Y по X методом наименьших квадратов, принимая в качестве сглаживающей линии параболу:

$$\hat{y}_x = a \cdot x^2 - b \cdot x + c\,,$$

и затем построить ее на одном чертеже с эмпирической линией регрессии.

3. Оценить тесноту корреляционной зависимости Y по X.

4). Проверить адекватность уравнения регрессии Y по X.

1.

x_i	100	200	300	400	500
y_{ij}	210 221 249 271 294	251 251 294 317 331	250 291 317 320 345	300 320 351 372 380	371 329 342 373 370

2.

x_i	11	12	13	14	15
y_{ij}	135 152 159 213 231	200 200 210 236 274	215 239 215 280 311	261 251 245 290 333	250 971 258 319 322

3.

x_i	1,5	2,1	3,5	4,7	5,2
y_{ij}	31,6 50,2 29,6 50,2 32,6 50,2	56 56 54 53 36 35	12,8 10,4 9,7 9,7 10,2 11,2	11,29 10,62 12,06 11,15 10,36 12,53	12,19 13,29 17,90 12,01 13,08 16,03

4.

x_i	10	21	35	42	54
y_{ij}	11,6 10,2 12,4 13,6 11,0 11,2	45 45 52 52 52 54	31,05 49,21 30,60 47,22 30,61 50,01	8,10 9,13 11,23 10,72 8,15 8,17	9,55 8,55 9,13 11,32 9,48 9,22

5.

x_i	1,1	1,2	1,3	1,4	1,5
y_{ij}	2,71 2,53 2,14 3,72 2,40	2,32 2,50 3,06 3,17 3,75	3,19 2,72 2,63 2,41 2,25	3,21 2,95 3,36 3,27 3,31	8,10 6,52 5,63 6,28 6,07

6.

x_i	10	20	30	40	50
y_{ij}	1,23 1,25 2,28 1,28 1,92	2,25 1,30 1,31 1,27 2,27	1,33 2,32 2,32 1,32 1,28	1,71 1,55 1,50 1,53 1,52	2,23 2,30 1,85 2,32 1,86

7.

x_i	2	4	6	8	10	12
y_{ij}	1,2 1,2 1,0 1,4 1,3	1,1 1,1 1,2 1,2 0,9	0,6 0,8 0,6 0,6 0,7	0,6 0,4 0,6 0,3 0,6	1,1 1,2 1,0 1,0 1,1	1,5 1,3 1,6 1,4 1,4

8.

x_i	5,1	7,1	10,1	12,1	15,1
y_{ij}	9,27 9,15 8,87 8,84	5,30 6,20 3,25 3,77	4,3 5,1 6,4 9,7	8,1 8,23 7,5 7,1	9,27 8,18 8,87 9,34

9.

x_i	0,1	0,2	0,3	0,4	0,5
y_{ij}	9,38 5,0 3,23 3,61 1,38	4,61 5,38 5,23 6,65 4,03	9,61 3,33 11,0 3,83 3,38	7,51 9,61 4,62 7,93 8,53	9,61 10,0 5,72 8,54 8,73

10.

x_i	10	20	30	40	50
y_{ij}	164,9 135,7 123,5 172,9	93,42 71,35 84,0 91,73	40,55 77,95 52,06 35,84	72,87 61,28 31,83 31,42	61,28 31,83 81,34 31,45

12. КОНТРОЛЬНЫЕ ВОПРОСЫ ДЛЯ САМОСТОЯТЕЛЬНОЙ ОЦЕНКИ КАЧЕСТВА ОСВОЕНИЯ ДИСЦИПЛИНЫ

1. Число перестановок, размещений, сочетаний.
2. Классификация событий.
3. Основные определения теории вероятностей.
4. Сумма событий, произведение.
5. Классическое определение вероятности события.
6. Аксиомы вероятностей.
7. Теоремы сложения и умножения вероятностей (формулировки).
8. Вероятность появления хотя бы одного из n независимых событий.
9. Формула полной вероятности события, формулы Байеса.
10. Повторные испытания, формулы Бернулли.
11. Локальная и интегральная формулы Лапласа.
12. Определения и классификация случайных величин.
13. Ряд распределения. Многоугольник распределения.
14. Функция распределения случайной величины, плотность распределения.
15. Числовые характеристики случайных величин и их вероятностный смысл.
16. Показательный закон распределения и его числовые характеристики.
17. Биномиальный закон распределения, числовые характеристики.
18. Закон Пуассона.
19. Равномерный закон распределения, функция распределения и числовые характеристики.
20. Простейший поток событий.
21. Нормальный закон распределения.
22. Кривая Гаусса.
23. Влияние параметров нормального распределения на вид кривой Гаусса.
24. Основные задачи математической статистики. Генеральная совокупность, выборка.
25. Полигон частот, гистограмма.
26. Эмпирическая функция распределения.
27. Точечные оценки параметров. Свойства точечных оценок.
28. Интервальные оценки параметров.
29. Доверительная вероятность, доверительные интервалы.
30. Понятие критерия согласия. Критерий Пирсона.
31. Парная линейная регрессия. Метод наименьших квадратов.
32. Коэффициент корреляции.
33. Корреляционное отношение, теснота связи.

СПИСОК ЛИТЕРАТУРЫ

Вентцель Е.С. Теория вероятностей. – М.: Наука, 1964.

Вентцель Е.С., Овчаров Л.А. Теория вероятностей: Задачи и упражнения. – М.: Наука, 1973.

Гмурман В.Е. Введение в теорию вероятностей и математическую статистику. – М.: Высш. шк., 1972.

Гмурман В.Е. Руководство к решению задач по теории вероятностей и математической статистике. – М.: Высш. шк., 1975.

Гурский Е.И. Теория вероятностей математической статистики. – М.: Высш. шк., 1971.

Карасев А.И. Теория вероятностей и математическая статистика. – М.: Статистика. 1979.

Кремер Н.Ш. Теория вероятностей и математическая статистика. – М.: ЮНИТИ, 2000.

Румшинский Л.З. Элементы теории вероятностей. – М.: Наука, 1970.

ОГЛАВЛЕНИЕ