Электронный научный журнал
«Инновации в образовании»
Специальный выпуск

Ю.В.Родионов, А.Д.Нахман

ВВЕДЕНИЕ В СТОХАСТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
машин, агрегатов и процессов

Учебно-методическое пособие

Издательская платформа
Российской академии естествознания
2017

УДК 372.851

Рецензенты:

Доктор технических наук, доцент кафедры «Техническая механика и детали машин» ФГБОУ ВО «Тамбовский государственный технический университет» С.В.Плотникова,

доцент кафкдры общеобразовательных дисциплин ТОГОАУ ДПО «Институт повышения квалификации работников образования» И.Ю.Иванова

Рекомендовано Ученым советом ФГБОУ ВО «Тамбовский государственный технический университет» в качестве учебно-методического пособия

Родионов, Ю.В. Введение в стохастическое моделирование машин, агрегатов и процессов: учебное-методическое пособие / Ю.В.Родионов, А.Д.Нахман – «Инновации в образовании». Специальный выпуск. Издательская платформа Российской академии естествознания. – 2017. –92 с.

Начальная подготовка в области стохастического моделирования (и, в частности, моделирования машин, агрегатов и процессов) предполагает ознакомление обучающихся с основными понятиями и фактами комбинаторики, теории вероятностей и математической статистики, а также типовыми задачами на вычисление вероятностей и анализ распределений случайных величин. Материал сформирован в виде кейса, содержащего глоссарий, обучающий и контрольный модули. Теоретический блок содержит основные понятия и утверждения с полными обоснованиями, а также образцы решения типовых задач. В составе контрольного модуля представлены теоретические упражнения, задания для самостоятельного тренинга и итоговые тесты с ответами. Материал адресован студентам инженерных и экономических направлений обучения.

На пособие получено положительное заключение Методического совета ФГБОУ ВО «Тамбовский государственный технический университет» ; протокол №1 от 23 марта 2017 г.

ВВЕДЕНИЕ

1."Машины, агрегаты и процессы" - область науки и техники, включающая разработку научных и методологических основ конструирования, производства, ремонта и эксплуатации машин, агрегатов и процессов. В её основе лежат теоретические и экспериментальные исследования, и, в частности, исследования параметров машин и агрегатов и их взаимосвязей при комплексной механизации основных и вспомогательных процессов и операций. Важное место здесь отводится стохастическому моделированию, и, в частности, методам обработки экспериментальных данных. Активно используются такие понятия, как вероятность безотказной работы, функция распределения времени безотказной работы, плотность распределения, понятие интенсивности (относительной частоты) отказов и др. ([1]).

Одним из примеров стохастического подхода в теории машин и агрегатов является расчет валов, существенно влияющих на надежность насосных агрегатов, которые в процессе перекачки нефти и нефтепродуктов испытывают значительные динамические нагрузки, вызванные вибрацией. В связи с этим возникает необходимость развития расчетных (том числе – вероятностно-статистических) методов оценки ресурса валов с целью получения показателей надежности насосов на начальной стадии их эксплуатации. Необходимость развития такого подхода в теории надежности конструкций и машин указана во многих работах (см., напр., [2]).

Введение в стохастическое моделирование машин, агрегатов и процессов предполагает ознакомление обучающихся с основными понятиями и фактами вероятностно-статистической теории. Здесь важное место принадлежит различным практико-ориентированным задачам, порою выходящим далеко за рамки означенного научно-технического направления. Именно таким задачам и посвящено настоящее пособие.

2. Стохастическое моделирование. Математические модели служат средством познания свойств и закономерностей поведения объекта-оригинала путем создания его образа (объекта-заместителя), выраженного в математической форме.

Математическое моделирование будем понимать как процесс установления соответствия данному реальному объекту его математической модели, исследования модели средствами математики и интерпретации результата в терминах исходной предметной области.

Стохастическая (недетерминированная, вероятностная) модель есть математическая модель, для которой параметры, условия функционирования и характеристики состояния моделируемого объекта представлены случайными величинами и связаны случайными зависимостями.

Традиционные задачи теории вероятностей и математической статистики могут рассматриваться как простейшие задачи стохастического моделирования, поскольку предполагают анализ (средствами математики) процессов и явлений, носящих случайный характер.

Так, речь может идти:

-о задачах на нахождение вероятности случайного события, которое может наступить в результате проведения данного опыта;

- о задачах на анализ распределения некоторой случайной величины (в частности, о нахождении параметров её распределения, числовых характеристик, вероятностей принятия значений в заданном интервале и др.);

- о задачах на анализ эмпирических распределений (в частности, о нахождении точечных или интервальных оценок параметров соответствующего теоретического распределения и др.)

3. Актуальность стохастических знаний и умений. Стохастические знания, вероятностное мышление, умение строить научно обоснованные прогнозы, становятся как нельзя востребованными в современных условиях. Изучение различного рода случайных явлений, стохастических отклонений от нормы является важным средством предотвращения чрезвычайных ситуаций, техногенных катастроф, выпуска некачественной и ненадежной продукции и т.п.

Методы теории вероятностей и математической статистики находят все большее применение, например, к анализу ошибок разного рода измерений, а также в физике, биологии, экологии, социологии, в телефонии и процессах обслуживания, адаптивного управления и т. д.

По этой причине стохастические знания становятся неотъемлемым и крайне востребованным компонентом содержания образования – как общего, так и профессионального, чем определяется ***актуальность*** *данной работы.*

В школьном курсе (уровень общего среднего образования) , а также на начальной ступени бакалаврской подготовки стохастическое моделирование ограничивается рамками практико-ориентированных и прикладных задач. В процессе изучения «смежных» дисциплин возникают и решаются уже профессинально-ориентированные и даже квазипрофессиональные задачи. Более того, в процессе освоения вузовского курса математики и смежных дисциплин совокупность вероятностно статистических знаний и умений будущего специалиста (бакалавра, магистра) должна перерасти в стохастическую компетенцию.

4. Понятие стохастической компетенции является сравнительно новым**.** Его возникновение связано с двумя факторами;

1)утверждением в отечественной системе образования *компетентностного подхода*;

2) наличием все более возрастающего общественного интереса к теории вероятностей, математической статистике, теории случайных процессов и к применению вероятностно-статистических методов в самых разнообразных областях науки, техники, производства и экономики.

Стохастической компетенцией мы будем называть способность к математической и практической деятельности, связанной с овладением основными комбинаторики, понятиями и фактами теории вероятностей, математической статистики и теории случайных процессов.

Стохастическая компетентность, в нашем понимании, есть *готовность* к такой деятельности, умение мобилизовать знания в области стохастики на использование их в новых условиях, в решении практических, производственных и др. задач.

Условиями для формирования соответствующей компетенции/компетентности являются:

- знание концептуальных основ стохастики;
- владение разнообразными методами вероятностно-статистического анализа окружающих явлений;
- использование методологии современной науки, осмысление глубокого внутреннего единства эмпирического и теоретического уровней познания мира случайного.

В структуре стохастической компетенции/компетентности мы выделяем следующие основные компоненты.

1) *Мотивационно-ценностный*: мотивация, заинтересованное отношение к математической деятельности.

2) *Когнитивный*: знания, умения, в области теории вероятностей и математической статистики.

3) *Операциональный*: опыт практического применения математических знаний, закрепление умений на уровне навыков.

4) *Рефлексивный*: включение в математическую деятельность, рефлексия

математической деятельности (в частности, самоконтроль, самоанализ и самооценка).

В соответствии с вышеотмеченными компонентами стохастической компетенции, в данной работе представлены:

- задачи практико-ориентированного содержания, призванные повысить *мотивацию* обучающихся к математической деятельности;
- основные теоретические сведения и теоретические упражнения в области теории вероятностей и математической статистики, способствующие формированию *когнитивного* компонента компетенции;
- задачи с решениями (обсуждениями решений) и для самостоятельного решения, способствующие активному обучению студентов и выработке *операционального* компонента компетенции;
- тестовые задания (с ответами) для самоконтроля, самооценки и самоанализа, способствующие *рефлексии* соответствующей математической деятельности.

Данный интегрированный подход находится в русле *инновационных форм образовательной деятельности*, чем и определяется ***новизна*** *настоящего учебного материала.*

В то же время, в содержание пособия включены как традиционные понятия и факты, так и традиционные типы задач, что свидетельствует о достаточно высокой степени ***преемственности*** материала.

5. О данном пособии. Его структура – следующая:

- терминологическое поле (основные понятия и факты, сформулированные в краткой и доступной форме с целью первичного ознакомления);
- обучающий модуль (содержит основные теоретические сведения и задачи для активного обучения);
- контрольный модуль.

При этом контрольный модуль содержит:

- теоретические упражнения (упражнения в доказательствах или выявлениях некоторых фактов теории, способствующие формированию исследовательских навыков обучающихся);
- задачи для самостоятельного решения (в основном, задачи из теории вероятностей и математической статистики практико-ориентированного содержания);
- итоговые тесты по каждому изучаемому разделу.

Представление материала форме кейса и значительное количество авторских задач определяют ***оригинальность*** пособия, его отличие от других изданий по основам стохастики.

Таблицы и графики, иллюстрирующие теоретический и практический материал, соответствуют дидактическому принципу ***наглядности*** в обучении.

Достаточно широкий охват учебного материала и демонстрация (в ходе решения задач) глубоких внутренних связей между его разделами позволяют говорить о представленной ***целостной картине соответствующей области знаний.***

6. Концепция развития математического образования, ФГОС и стохастическая компетенция. В соответствии с Концепцией развития математического образования в Российской федерации и введением *новых ФГОС общего и профессионального образования* стохастическим знаниям отводится роль неотъемлемого компонента инновационного содержания образования – как общего, так и профессионального. Изучение вероятностно-статистического материала необходимо уже в школьном курсе в рамках самостоятельной содержательно-методической линии. Согласно ФГОС основного общего образования изучение предметной области «Математика и информатика» должно «обеспечить осознание значения математики … как части общечеловеческой культуры, универсальном языке науки, позволяющем описывать и изучать реальные процессы и явления». Изучение математики должно способствовать развитию логического мышления, формированию первичных навыков математического моделирования, применению математических знаний при решении различных задач и оцениванию полученных результатов, развитию математической интуиции.

В следующей таблице 1 представлен ***список учебных целей*** и отражена *динамика развития компонент* «знать-уметь» в структуре стохастической компетенции в контексте требований ФГОС к уровню подготовки выпускников основной и старшей школ.

	Знать/понимать	***Уметь***
Основная школа	• вероятностный характер многих закономерностей окружающего мира;	• извлекать информацию, представленную в стандартных формах (таблицы, диаграммы, графики) и

	● примеры статистических закономерностей и выводов;	строить соответствующие формы; ● решать комбинаторные задачи путем систематического перебора возможных вариантов, а также с использованием правила умножения; ● вычислять средние значения результатов измерений; ● находить частоту события, используя собственные наблюдения и готовые статистические данные; ● находить вероятности случайных событий в простейших случаях; ● использовать приобретенные знания и умения в практической деятельности и повседневной жизни для анализа реальных числовых данных, представленных в виде диаграмм, графиков, таблиц; решения учебных и практических задач, требующих систематического перебора вариантов; сравнения шансов наступления случайных событий, оценки вероятности случайного события в практических ситуациях, сопоставления модели с реальной ситуацией; понимания статистических утверждений.
Старшая школа	● значение математической науки, в том числе, методов стохастики, для решения задач, возникающих в теории и практике; значение практики для развития самой математической науки; ● детерминированный и вероятностный характеры различных процессов и закономерностей окружающего мира, значение статистических данных для прогнозирования явлений и процессов .	● решать простейшие комбинаторные задачи с использованием известных формул; ● вычислять вероятности событий на основе подсчета числа исходов и свойств вероятности (простейшие случаи); ● анализировать массив данных, в том числе применять простейшие методы интерполяции и экстраполяции; ● использовать приобретенные знания и умения в практической деятельности и повседневной жизни для: анализа

		информации статистического характера, прогнозирования наступления событий на основе вероятностно-статистических методов; применения полученных умений в решении задач из смежных дисциплин.

Таблица 1. Динамика развития компонент стохастической компетенции

В ФГОС высшего профессионального образования структуры формируемых компетенций дополнены компонентом «владеть». С нашей точки зрения, *данный компонент соответствует приобретению первичного опыта соответствующей деятельности и, следовательно, переходу в «нечетко-пограничную зону» между компетенцией и компетентностью* (компетенция «в действии»). Приведем в качестве примера (таблица 2), требования к структуре результата обучения, способствующего формированию профессиональной компетенции (ПК-1) «использование основных законов естественнонаучных дисциплин в профессиональной деятельности, применение методов… моделирования, теоретического и экспериментального исследования» (ФГОС направления подготовки бакалавров 270800.62 «Строительство»). Здесь стохастическая компетенция/компетентность выступает в качестве одной из подсистем, интегрированных в указанную компетенцию ПК-1.

Структура результата обучения		
Обучающийся знает:	***Обучающийся умеет:***	***Обучающийся владеет:***
- фундаментальные основы теории вероятностей математической статистики: • понятия вероятности и относительной частоты события и основные формулы для их вычисления; • методы точечного и	*Самостоятельно использовать математический аппарат, содержащийся в специальной литературе, расширять свои математические познания:* • решать прикладные задачи (в том числе, связанные с профессиональной	*- первичными навыками и основными методами решения математических (стохастических) задач из общеинженерных и специальных дисциплин профилизации и методами вероятностного моделирования:* • численно прогнозирует

интервального оценивания параметров распределения; • понятия и виды случайных процессов; • регрессионные методы моделирования экспериментальных зависимостей.	деятельностью) с помощью стохастических методов; • использовать алгоритмы проверки статистических гипотез; • строить вероятностно-статистические модели реальных ситуаций.	степень объективной возможности того или иного явления, характера протекания того или иного процесса; • интерпретирует статистические данные в терминах вариационных рядов и оценок параметров распределения; • разрабатывает регрессионные модели статистических зависимостей

Таблица 2. Требования к структуре результата обучения

Ввиду ***практической ориентированности*** стохастической компетенции/компетентности ее компоненты «знать», «уметь» и, в особенности, «владеть», наилучшим, с нашей точки, образом формируется в процессе использования так называемых *кейс-заданий* (примеры кейс заданий представлены в п. 10 раздела «Случайные величины»).

7. Формированию стохастической компетенции/компетентности способствует **освоения следующего содержания образования** (см. также [3], [4], [5]).

Множества и комбинаторика. Множества, элементы множества. Подмножества. Объединение и пресечение множеств. Диаграммы Эйлера. Примеры решения комбинаторных задач: перебор вариантов, правило умножения. Основные виды выборок: размещения, перестановки, сочетания. Бином Ньютона.

Алгебра событий Достоверное и невозможное события. Случайные события. Действия над событиями. Виды случайных событий.

Статистический эксперимент. Частота события в статистическом эксперименте. Частота и вероятность.

Вероятность. Классическая модель вероятности. Элементарные и сложные события в классической модели вероятности. Вероятность сложного события. Условная вероятность. Независимые события. Геометрическая вероятность. Схема гипотез.

Теоретические распределения. Дискретные и непрерывные случайные величины. Функция и плотность распределения. Числовые характеристики. Специальные распределения (равномерное, геометрическое, биномиальное, Пуассоновское, показательное, нормальное). Закон больших чисел.

Эмпирические распределения. Представление данных в виде таблиц, диаграмм, графиков. Генеральная совокупность. Выборка и ее анализ статистическими методами. Средние результаты измерений. Понятие о статистическом выводе на основе выборки. Точечные и интервальные оценки параметров распределения.

ТЕРМИНОЛОГИЧЕСКОЕ ПОЛЕ

Элементы комбинаторики

Комбинаторика – это раздел математики, изучающий задачи нахождения количества всевозможных конечных подмножеств данного множества, если эти подмножества обладают заданной характеристикой.

Кортеж есть упорядоченное подмножество, составленное из элементов данного множества; длина кортежа есть количество составляющих его элементов.

Декартово произведение множеств $G_1,...,G_n$ есть множество всех G кортежей вида $g=(g_1,...,g_n)$, где $g_k \in G_k$, $1 \le k \le n$.

Размещения с повторениями – это кортежи вида $g=(g_1,...,g_n) \in G \times ... \times G$.

Размещения без повторений (из *m* элементов по *k*) – это кортежи длины *k* из элементов одного и того же *m*-элементного множества G так, чтобы элементы в кортеже не повторялись.

Перестановки – это размещения (без повторений) из *m* элементов по *m.*

Сочетания из m по k элементов – это неупорядоченные подмножества по *k* элементов, взятых из некоторого *m*-элементного множества.

Разложение степени ***бинома Ньютона*** есть разложение вида

$$(\alpha+\beta)^m = \sum_{k=0}^{m} C_m^k \alpha^{m-k} \beta^k .$$

Биномиальные коэффициенты – это коэффициенты разложения степени бинома, т.е. числа вида C_m^k.

Треугольник Паскаля – это треугольник, образованный биномиальными коэффициентами разложений нулевой, первой, второй,… степеней бинома.

Теория вероятностей и математическая статистика

Теория вероятностей – это математическая наука, предметом которой является изучение закономерностей массовых случайных явлений.

Классическая вероятность события A – это отношение числа m исходов, благоприятствующих A, к общему числу n всевозможных элементарных исходов опыта.

Суммой (***произведением***) ***событий*** называется событие, состоящее в появлении хотя бы одного (совместно всех) данных событий. Обозначение:

Противоположными называют события A и $\overline{A}$, если они несовместны и образуют полную группу.

Схема гипотез предполагает ситуацию, когда событию А предшествует появление одного и только одного из полной группы попарно несовместных событий (гипотез), но заранее неизвестно, какая именно из гипотез наступит.

Схема Бернулли предполагает наличие n однотипных опытов (испытаний), в каждом из которых вероятность появления события A является постоянной (равной некоторому p).

Случайная величина X есть числовая величина, которая в каждом опыте принимает одно и только одно значение, наперед неизвестное и зависящее от случайных причин.

Дискретной называется такая ***случайная величина*** X, все возможные значения которой можно записать в виде числовой последовательности (конечной или бесконечной).

Непрерывная случайная величина X принимает сплошь все значения из некоторого числового промежутка.

Закон распределения дискретной случайной величины X есть соответствие между ее возможными значениями и вероятностями этих значений.

Ряд распределения есть таблица, с помощью которой задается закон распределения.

Функция распределения (интегральная функция) соотносит каждому $x \in (-\infty;+\infty)$ вероятность события, состоящая в принятии величиной X значения левее точки x.

Плотность распределения (дифференциальная функция) непрерывной случайной величины есть производная функции распределения.

Числовые характеристики случайной величины X – это ***математическое ожидание*** $M(X)$ («среднее значение») ***дисперсия*** $D(X)$

(степень рассеяния значений X относительно математического ожидания) и ***среднее квадратическое отклонение*** $\sigma(X)=\sqrt{D(X)}$**.**

Выборка **(**выборочная совокупность) есть множество объектов случайно отобранных из большего множества (называемого генеральной совокупностью).

Объем выборки есть количество отобранных объектов.

Варианты **–** это наблюдаемые значения количественного признака (значения случайной величины, которыми характеризуются объекты выборочной совокупности); ***частота варианты*** **–** число, показывающее, сколько раз наблюдалось в выборочной совокупности данная варианта.

Статистическое распределение выборки есть соответствие между вариантами и соответствующими им значениями частот.

Вариационный ряд есть таблица, с помощью которой задается статистическое распределение.

Относительная частота наблюдаемого значения (варианты) есть отношения соответствующей частоты к объему выборки.

Мода есть варианта, имеющую наибольшую частоту.

Медиана **–** это значение, соответствующее середине вариационного ряда.

Размахом варьирования R называют разность между наибольшей и наименьшей вариантой.

Выборочная средняя есть среднее арифметическое всех наблюдаемых (в выборке) значений.

Выборочная дисперсия есть характеристика рассеяния наблюдаемых значений относительно их выборочной средней.

Точечная оценка параметра θ есть его приближенное значение θ^*, которая определяется одним числом; и***нтервальная оценка*** определяется двумя числами – концами интервала.

Доверительный интервал есть интервал, который покрывает (содержит) параметр θ с заданной надежностью (доверительной вероятностью) γ.

ЭЛЕМЕНТЫ КОМБИНАТОРИКИ

ОБУЧАЮЩИЙ МОДУЛЬ

Перечислительной ***комбинаторикой*** называется раздел математики, изучающий задачи нахождения количества всевозможных конечных подмножеств данного множества, если эти подмножества обладают заданной характеристикой.

1. Кортежи. Прямые произведения

1.1. Пусть даны множества $G_1,...,G_n$. ***Кортежем длины n,*** составленным из элементов этих множеств, называется любая последовательность вида $g=(g_1,...,g_n)$, где $g_k \in G_k$, $1 \le k \le n$.

Кортежи длины 2, т.е. кортежи вида (g_1, g_2), называются ***упорядоченными парами***; кортежи длины 3 - упорядоченными тройками и т.д.

Декартовым произведением $G_1 \times ... \times G_n$ ***множеств*** $G_1,...,G_n$ называется множество всех G кортежей вида $g=(g_1,...,g_n)$.

Например, декартово произведение множеств $G_1 \times G_2$ состоит из всевозможных упорядоченых пар (g_1, g_2), где $g_1 \in G_1, g_2 \in G_2$.

1.2. **Пример**. Брошены два игральных кубика. Найти количество всевозможных вариантов (всевозможных пар) очков на выпавших гранях.

Решение. Следует найти количество всех упорядоченных пар (g_1, g_2), где g_k - число очков, выпавших на k-ой игрального кубика, $k=1,2$. Значение g_1 - любое из чисел $1,2,...,6$. В паре с фиксированным (выбранным) g_1 может оказаться любое $g_2 \in \{1,2,...,6\}$, т.е. таких пар будет 6, соответственно количеству возможных значений g_2. Учитывая, что выбор g_1 возможен тоже шестью способами, имеем все упорядоченные пары в количестве $6 \cdot 6 = 36$.

1.3. Обозначим через $\nu(M)$ количество элементов конечного множества M .

Теорема 1 ("принцип умножения"). Имеет место равенство

$$\nu(G_1 \times G_2) = \nu(G_1) \cdot \nu(G_2).$$

Для доказательства надо повторить рассуждения, приведенные при решении примера п. 1.2. Используя результат теоремы 1 и метод математической индукции, получаем следующее утверждение:

$$\nu(G_1 \times \ldots \times G_n) = \nu(G_1) \cdot \ldots \cdot \nu(G_n).$$

1.4. В частности, если имеется k *экземпляров одного и того же множества* G и требуется найти число кортежей длины k в которые входит *по одному* (и только одному) элементу из каждого экземпляра, то такие кортежи называются ***размещениями с повторениями***; другими словами размещения с повторениями это кортежи вида $g = (g_1, \ldots, g_n) \in G \times \ldots \times G$. Число таких кортежей, согласно принципу умножения, очевидно, есть

$$\nu(G^k) = m^k .$$

1.5. **Пример.** Какое максимальное число шестизначных телефонных номеров может быть в городской телефонной сети ?

Решение. Имеем кортежи длины k=6 , в которых каждый элемент принадлежит множеству G={0,1,2,...,9}, содержащему m=10 элементов. Следовательно, число таких кортежей $\nu(G^6) = 10^6$. Однако телефонного номера, состоящего сплошь из нулей, обычно не бывает. Следовательно, искомое число номеров равно $10^6 - 1$=999999.

2. Размещения, перестановки, сочетания

2.1. Если строить *кортежи длины k из элементов одного и того же множества* G так, чтобы *элементы в кортеже не повторялись*, то такие кортежи называются ***размещениями без повторений.***

Рассуждая подобно п. 1.2, получаем следующее утверждение.

Теорема. Количество размещений (из m элементов по k элементов) есть число

$$A_m^k = m(m-1)\ldots(m-k+1).$$

Доказательство теоремы вытекает из принципа умножения (п. 1.3).

Если использовать обозначение $m! = 1 \cdot 2 \cdot \ldots \cdot (m-1)m$, считая 0!=1!=1, то легко проверить, что

$$A_m^k = \frac{m!}{(m-k)!}.$$

2.2. В частности, размещения из m по k элементов называют ***перестановками***; число всевозможных перестановок из m элементов есть

$$P_m = m!\ .$$

2.3. **Пример.** Сколькими способами можно разложить в ряд на витрине магазина пять DVD-дисков?

Решение. Имеем кортежи длиной в 5 элементов, составленные из элементов множества G, для которого $\nu(G) = 5$. Следовательно, ищем количество перестановок из 5 элементов

$$P_5 = 5! = 120.$$

2.4. **Пример.** Сколькими способами можно расставить на полке 6 книг, если

а) две определенные книги должны всегда стоять рядом,

б) эти две книги не должны стоять рядом?

Решение. а) Пару книг, которые должны стоять рядом, условимся пока рассматривать как одну книгу. Тогда нужно расставить 5 книг по пяти местам, что можно сделать $P_5 = 5!$ способами. Учитывая теперь порядок расположения тех двух книг, которые мы посчитали за одну, имеем P_2 перестановок между ними Согласно принципу умножения, получаем окончательно число способов $P_5 \times P_2 = 120 \times 2 = 240$.

б) Способов переставить 6 книг существует $P_6 = 720$, но из них, как установлено в п.а) существует 240 способов поставить определенные книги вместе. Следовательно, число способов поставить книги так, чтобы 2 заданные книги не стояли, равно разности 720-240=480.

2.5. **Пример.** На заседании Думы из 6 возможных кандидатов выбирают председателя комитета , его первого и второго заместителя. Сколько существует способов формирования руководящего состава этого комитета?

Решение. Имеем кортежи длиной в 3 элемента (упорядоченные тройки), составленные из элементов множества G, для которого $\nu(G) = 6$. Следовательно, ищем количество размещений A_6^3:

$$A_6^3 = \frac{6!}{(6-3)!} = \frac{3! \cdot 4 \cdot 5 \cdot 6}{3!} = 120$$

2.6. Если теперь строить из элементов множества G, для которого

$\nu(G) = m$, *неупорядоченные* подмножества по k элементов, то такие подмножества называют ***сочетаниями из** m **по** k **элементов.***

Теорема . Количество всевозможных сочетаний из m по k элементов может быть вычислено по формуле

$$C_m^k = \frac{A_m^k}{P_k}. \tag{2.1}$$

Доказательство. Чтобы получить всевозможные *размещения* из m по k элементов, можно сначала «организовать» всевозможные сочетания (неупорядоченные подмножества из m по k) в количестве C_m^k, а затем в каждом таком подмножестве произвести всевозможные упорядочения, т.е. произвести всевозможные перестановки в количестве P_k. Рассуждая как в п. 1.2, получаем тогда $A_m^k = C_m^k P_k$, откуда и вытекает утверждение теоремы.

2.7. Имеют место равенства

$$C_m^k = \frac{m!}{k!(m-k)!}, \tag{2.2}$$

$$C_m^k = C_m^{m-k}, \tag{2.3}$$

$$C_m^k = C_{m-1}^{k-1} + C_{m-1}^k. \tag{2.4}$$

Соотношение (2.2) вытекает из (2.1) и формулы для вычисления числа размещений; соотношения (2.3) и (2.4) нетрудно проверить, если воспользоваться представлением (2.2).

2.8. **Пример.** Имеется 10 различных игрушек, из которых формируют комплекты подарков по три игрушки в каждом. Сколько таких различных комплектов можно сформировать?

Решение. Имеем всевозможные неупорядоченные подмножества по 3 элемента из 10. Согласно п.п. 2.4, 2.5 ищем количество сочетаний, т.е.

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

Итак, можно сформировать 120 различных комплектов подарков.

2.9. **Пример.** Найти из уравнения $C_{n+3}^n - C_{n+2}^{n-1} = 15n + C_{15}^1$.

Решение. Воспользовавшись формулой (2.2), запишем уравнение в виде

$$\frac{(n+3)!}{n!(n+3-n)!)}-\frac{(n+2)!}{(n-1)!(n+2-n+1)!)}=15n+\frac{15!}{1!14!}$$

или

$$\frac{n!(n+1)(n+2)(n+3)}{n!3!}-\frac{(n-1)!n(n+1)(n+2)}{(n-1)!3!}=15(n+2)$$

Сокращая дроби и сокращая обе части уравнения на $n+2$ (что можно делать, т.к. $n>0$), имеем

$(n+1)(n+3)-n(n+1)=15\cdot 3!$, откуда $3n=30$, т.е. $n=10$.

2.10. Размещения, перестановки, сочетания можно интерпретировать как всевозможные *выборки* заданного количества элементов из некоторой совокупности, содержащей m элементов. При этом выборки-размещения различаются как самими элементами, так и их порядком; выборки-сочетания разнятся только самими элементами (порядок их следования несущественен), выборки перестановки из m элементов друг от друга могут отличаться лишь порядком следования элементов. Именно такие признаки обычно используются для распознавания вида выборок в теории вероятностей.

2.10. **Пример.** В микроавтобусе 11 мест. Сколькими способами 11 человек могут расположиться в этой машине, если занять место водителя могут только трое из них, имеющие доверенность на право управления этой машиной?

Решение. На место водителя можно посадить только одного из трех человек, т.е. существуют 3 способа занять первое место. Остальные 10 мест могут быт заняты любым из оставшихся десяти пассажиров числом способов $P_{10}=10!$ (эти выборки различаются лишь порядком следования элементов, т.е. являются перестановками из 10 элементов). Используя принцип умножения, получаем произведение $3\times 10!=3628800$.

3. Число элементов в объединении множеств

3.1. Напомним, что ***объединением множеств*** A и B называется множество $A\cup B$, каждый элемент которого содержится хотя бы в одном из множеств A или B, т.е. содержится или в A, или в B, или в обоих этих множествах. Общая же часть $A\cap B$ этих множеств называется их ***пересечением***.

3.2. **Теорема 1**. Если конечные множества A и B имеют пустое пересечение, то количество элементов

$$n(A \cup B) = n(A) + n(B).$$

Теорема 2. Если конечные множества A и B имеют непустое пересечение, то

$$n(A \cup B) = n(A) + n(B) - n(A \cap B).$$

Доказательства теорем 1 и 2 легко вытекают из геометрических интерпретаций объединения непересекающихся и пересекающихся множеств (диаграмм Вьена).

3.3. **Пример**. Все акционеры ЗАО переводят дивиденды на счета в банках. При этом известно, что 4 акционера имеют счета в Сбербанке, 6 человек – в Россельхозбанке, но из них двое имеют также счета и в Сбербанке. Сколько всего акционеров в данном ЗАО?

Решение. Пусть A – множество акционеров, имеющих счета в Сбербанке, B – множество акционеров, имеющих счета в Россельхозбанке. Тогда требуется найти $n(A \cup B)$. Имеем

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) = 4 + 6 - 2 = 8.$$

4. Бином Ньютона

4.1. Известные формулы квадрата, куба суммы могут быть обобщены следующим образом.

Теорема. Для любого натурального m имеет место соотношение

$$(\alpha + \beta)^m = \sum_{k=0}^{m} C_m^k \alpha^{m-k} \beta^k. \qquad (4.1)$$

Следует обратить внимание на то, что в слагаемых правой части разложения (4.1) степени α последовательно убывают от m к нулю, а степени β возрастают от нуля к m; степень каждого одночлена $\alpha^{m-k}\beta^k$ (от двух переменных) равна показателю m степени бинома; коэффициенты членов, равноудаленных от первого и последнего слагаемых разложения – одинаковы, поскольку имеет место свойство (2.3).

Числа C_m^k здесь называют биномиальными коэффициентами.

Общий член суммы (4.1), т.е. слагаемое стоящее на k+1 месте, считают k-ым членом разложения и обозначают T_k, т.е.

$$T_k = C_m^k \alpha^{m-k} \beta^k, \; k = 0,1,...m.$$

4.2. Один из возможных способов доказательства формулы (4.1) бинома Ньютона состоит в следующем. Сначала определяем коэффициенты a_k $(k = 0,1,...,m)$ в разложении

$$(1+x)^m = a_0 + a_1 x + a_2 x^2 + ... + a_{m-1} x^{m-1} + a_m x^m. \qquad (4.2)$$

Для этого полагаем в соотношении (4.2) $x = 0$; тогда $a_0 = 1 = C_m^0$. Далее дифференцируем обе части (4.2):

$$m(1+x)^{m-1} = a_1 + 2a_2 x + 3a_3 x^2 + ... + ma_m x^{m-1}.$$

Если выбрать $x = 0$, то получим $a_1 = m = C_m^1$. После второго дифференцирования имеем

$$m(m-1)(1+x)^{m-2} = 2!a_2 + 3 \cdot 2x + ... + m(m-1)a_m x^{m-2}.$$

В случае $x = 0$ получаем $m(m-1) = 2!a_2$, откуда

$$a_2 = \frac{A_m^2}{P_2} = C_m^2.$$

Теперь процесс ясен: в общем случае

$$a_k = C_m^k, k = 0,1,...,m.$$

Следовательно, соотношение (4.2) приобретает вид

$$(1+x)^m = 1 + C_m^1 x + C_m^2 x^2 + ... + C_m^{m-1} x^{m-1} + C_m^m x^m. \qquad (4.3)$$

Переход к утверждению (4.1) осуществляется, если положить теперь $x = \dfrac{\beta}{\alpha}$ в равенстве (4.3).

4.3. При x=1 из соотношения (4.3) вытекает следующий интересный факт:

$$C_m^0 + C_m^1 + C_m^2 + ... + C_m^{m-1} + C_m^m = 2^m.$$

Если вспомнить, что C_m^k есть количество всех неупорядоченных k-элементных подмножеств данного множества (содержащего всего m элементов), то указанная сумма выражает собою количество вообще всех различных подмножеств данного множества. Итак, ***число всех подмножеств множества, содержащего m элементов, равно*** 2^m.

Пример. В аудитории 8 люминесцентных ламп. Сколько существует различных способов освещения данной аудитории?

Решение. Любая комбинация из включенных в данный момент ламп есть некоторое подмножество данного множества, содержащего 8 элементов. Число всевозможных комбинаций (подмножеств) есть тогда 2^8, т.е. 512. Исключая случай нуль-элементного подмножества (случай, когда ни одна лампа не горит), получаем в ответе 511 способов освещения.

4.4. **Треугольник Паскаля**. Расположив коэффициенты разложения бинома при m=0,1,2,… в строки, получаем следующий треугольник, образованный такими строками:

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & . & & & . & & & . & &
\end{array}
$$

В силу свойства (2.4) любой коэффициент (кроме стоящих «по краям» единиц), получаются сложением двух стоящих над ним чисел в предыдущей строке. Так, в последней из выписанных строк 5=1+4, 10=4+6.

4.5 **Пример**. Записать разложение для $(a+b)^5$.

Решение. Записывая строку треугольника, соответствующую возведению в шестую степень, получаем следующую последовательность коэффициентов: 1,6, 15,20,15, 6,1. Теперь, согласно описанному в п. 4.1 принципу образования слагаемых разложения, будем иметь

$$(a+b)^6 = a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + b^6.$$

4.6. **Пример**. Найти член разложения степени бинома $(\frac{1}{\sqrt{x}} + x)^{12}$, не зависящий от *x*.

Решение. Имеем общий член разложения в виде

$$T_k = C_{12}^k (\frac{1}{\sqrt{x}})^{12-k} x^k = C_{12}^k x^{\frac{1}{2}k-6+k}.$$

Этот член не будет зависеть от *x*, если показатель степени *x* равен нулю: $\frac{3}{2}k - 6 = 0$, откуда получаем $k = 4$. Следовательно,

$$T_4 = C_{12}^4 = \frac{12!}{4!\cdot 8!} = \frac{9\cdot 10\cdot 11\cdot 12}{1\cdot 2\cdot 3\cdot 4} = 495\,.$$

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

ОБУЧАЮЩИЙ МОДУЛЬ

Теория вероятностей – это математическая наука, предметом которой является изучение закономерностей массовых случайных явлений.

Основными понятиями теории вероятностей являются понятия события и его вероятности. События можно разбить на три категории: достоверные (наверняка происходящие при выполнении данного комплекса условий; достоверные события обозначаем символом E), невозможные (наверняка не происходящие; невозможные события обозначаем символом $\varnothing$) и случайные (могут как произойти, так и не произойти при выполнении данного комплекса условий; обозначения: $A, B, C, \ldots$).

Вероятность понимается как некоторая численная мера степени объективной возможности появления данного события, т.е. каждому событию A сопоставляется (единственным образом) некоторое число $P = P(A)$.

СЛУЧАЙНЫЕ СОБЫТИЯ

1. Алгебра событий

1.1 . Во множестве событий вводятся следующие действия.

Сложение событий. Событие $A = A_1 + A_2 + \ldots + A_n$ называется суммой конечного количества событий $A_1, A_2, \ldots, A_n$, если событие А состоит в наступлении хотя бы одного из указанных $A_k, k = 1,\ldots,n.$

Событие $A = A_1 + A_2 + \ldots + A_n + \ldots$ называется суммой бесконечного количества событий, если событие А состоит в наступлении хотя бы одного из указанных $A_k, k = 1,2,\ldots$

Умножение событий. Событие $B = A_1 \cdot A_2 \cdot \ldots \cdot A_n$ называется произведением конечного количества событий $A_1, A_2, \ldots, A_n$, если событие А, состоит в совместном наступлении всех указанных $A_k, k = 1,\ldots,n.$

Замечание. Знаки равенства в этих определениях употреблены в смысле обозначений. Когда же мы будем *утверждать*, что события A и B равны, то означать это будет следующее: если происходит событие A, то наступает и B (т.е. событие B следует из A), и наоборот, событие A следует из B.

1.2. На основании введенных определений суммы и произведения событий легко проверить

а) ***свойства коммутативности*** операций сложения и умножения:

$$A_1 + A_2 = A_2 + A_1, \qquad A_1 \cdot A_2 = A_2 \cdot A_1,$$

б) ***свойства ассоциативности***:

$$A_1 + A_2 + A_3 = (A_1 + A_2) + A_3 = A_1 + (A_2 + A_3),$$

$$A_1 \cdot A_2 \cdot A_3 = (A_1 \cdot A_2) \cdot A_3 = A_1 \cdot (A_2 \cdot A_3)$$

б)***дистрибутивное свойство***

$$(A_1 + A_2)\, A_3 = A_1 A_3 + A_2 A_3.$$

1.3. События A_1, A_2 называются ***несовместными***, если $A_1 \cdot A_2 = \varnothing$.

Другими словами, два события несовместны, если в результате опыта наступление одного из них исключает наступление другого.

События $A_1, A_2, ..., A_n$ называются ***попарно несовместными***, если любые два из них несовместны.

1.4. Говорят, что события $A_1, A_2, ..., A_n$ образуют ***полную группу***, если

$$A_1 + A_2 + ...A_n = E.$$

Другими словами, события $A_1, A_2, ..., A_n$ образуют полную группу, если в результате опыта наступление хотя бы одного из них является достоверным.

1.5. События A и $\overline{A}$ называются ***противоположными***, если они несовместны и образуют полную группу: $A \cdot \overline{A} = \varnothing$, $A + \overline{A} = E$.

1.6. Пример. Блок в электрической схеме содержит два параллельно соединенных элемента; событие A_1 есть исправность первого элемента, A_2 – второго. Выразить через A_1 и A_2 следующие события : A блок пропускает ток; B блок не пропускает тока.

Решение. Поскольку наступление события A означает исправность хотя бы одного из элементов, то $A = A_1 + A_2$. Событие B означает неисправность каждого элемента, т.е. совместное наступление $\overline{A_1}$ и $\overline{A_2}$, а значит, $B = \overline{A_1} \cdot \overline{A_2}$

1.7. ***Алгеброй событий*** называется всякое множество событий U, в котором

-введены операции сложения и умножения, результаты выполнения которых также содержатся в *U*,

-которое содержит достоверные события,

-и которое вместе с каждым событием A содержит ему противоположное $\overline{A}$.

Алгебра событий, содержащая также всевозможные бесконечные суммы, называется σ -алгеброй (борелевской алгеброй).

2.Классическая вероятность

2.1. Будем рассматривать события *A*, *B*, *C*, ... как исходы некоторого опыта; число исходов считаем конечным. Каждому опыту сопоставим множество всех его ***элементарных*** (простейших, «неразложимых») ***исходов*** $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$, которое дает полную информацию о предполагаемых результатах этого опыта.

Элементарными называются такие исходы ω_i опыта, которые удовлетворяют следующим условиям:

а) ***группа исходов полна***, т.е. обязательно произойдет хотя бы один из ω_i;

б) исходы попарно ***несовместны***;

в)все ω_i – ***равновозможны***, т.е. объективно ни один из исходов не является более возможным, чем любой другой.

2.2. Среди элементов множества Ω имеются исходы, ***благоприятствующие*** событию *A*, то есть те, в результате которых событие *A* наступает.

Классической вероятностью события *A* называется отношение числа $m = m_A$ элементарных исходов, благоприятствующих *A*, к общему числу *n* всевозможных элементарных исходов опыта:

$$P(A) = \frac{m}{n}. \tag{2.1}$$

Для вычисления количества всевозможных и благоприятных исходов опыта часто пользуются формулами комбинаторики.

2.3. Очевидны следующие свойства классической вероятности:

$$P(E) = 1, \quad P(\varnothing)=0, \tag{2.2}$$

$$0 < P(A) < 1 \text{ для всякого случайного события } A. \tag{2.3}$$

Действительно, в первом случае все исходы благоприятны, т.е. $m_E = n$, а во втором $m_\varnothing = 0$, так что соотношения (2.2) – в силу определения (2.1) – очевидны; наконец, $0 < m_A < n$ для всякого случайного события A, а поэтому оценка (2.3) имеет место.

3. Относительная частота и статистическая вероятность

3.1. Это понятие вводится в результате анализа проведенных опытов. Если в результате n опытов событие А появилось $m = m_A$ раз, то относительной частотой события *A* называют число

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

3.2. Практика показывает, что с ростом числа однотипных опытов относительная частота приобретает свойство устойчивости, колеблясь относительно некоторого числа $P = P(A)$, которое называется статистической вероятностью события А.

3.3. Очевидно, что, относительная частота W(A) удовлетворяет свойствам, аналогичным (2.2), (2.3); в частности, $0 \le W(A) \le 1$.

4.Геометрическая вероятность

4.1. Рассмотрим следующий опыт: в некоторую ограниченную область *E* бросается точка, причем попадания ее в любые части, имеющие одинаковую меру (например, площадь в случае плоских областей) считаются равновозможными. Пусть событие *A* – ее попадание в область $A \subset E$. Если $S(A)$ и $S(E)$, $S(E) \ne 0$ – соответственно меры *A* и *E*, то геометрической вероятностью события *A* будем называть число

$$P(A) = \frac{S(A)}{S(E)}.$$

4.2. Для геометрической вероятности сохраняются, очевидно, свойства (2.2), (2.3) и оценка $0 \le P(A) \le 1$ для всякого случайного события A. Заметим, что здесь равенства $P(A) = 1$, $P(A) = 0$ оказываются возможными и для некоторых *случайных* событий *A*. Так например, если *E* –квадрат, а случайное событие A – попадание точки на любую его диагональ, то $P(A)=0$, поскольку диагональ имеет нулевую площадь .

5. Понятие об аксиомах вероятности

5.1. Вышеприведенные модели вероятности события (классическая, статистическая, геометрическая) обладают некоторыми общими свойствами, которые можно положить в основу аксиоматического определения понятия вероятности.

В первых, вероятность задается как некоторая функция P на алгебре событий U (или даже σ-алгебре событий). Во вторых, общими для рассмотренных определений являются свойства

А1 (неотрицательности) : $P(A) \geq 0$ для всякого $A \in U$;

А2 (нормированности): $P(E)=1$;

А3 (аддитивности**):** $P(A_1 + A_2 + \ldots + A_n) = \sum\limits_{k=1}^{n} P(A_k)$,

если события $A_1, A_2, \ldots, A_n$ попарно несовместны;

А4 («расширенной» аддитивности): $(A_1 + A_2 + \ldots + \mathrm{A}_n + \ldots) = \sum\limits_{k=1}^{\infty} P(A_k)$,

если события $A_1, A_2, \ldots, A_n, \ldots$ попарно несовместны (числовой ряд предполагается сходящимся).

5.2. Если для вводимой модели вероятности выполнены свойства А1 – А4, то имеют место следующие утверждения.

Теорема 5.1 Для всякой пары противоположных событий $A, \overline{A}$ справедливо равенство

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

Действительно, в силу соотношения $A + \overline{A} = E$ и нормированности вероятности (см. А1) имеем

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

Далее, в силу несовместности A, $\overline{A}$ и А3 получаем

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

что и утверждалось.

В частности, так как достоверное и невозможное события противоположны, то

$$P(\varnothing)=1-P(E)=1-1=0.$$

Теорема 5.2. Всякое событие A имеет вероятность $P(A)\le 1$.

Для доказательства воспользуемся результатом теоремы 5.1 и свойством неотрицательности (свойство А1), примененным к $\overline{A}$:

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

5.3. В параграфе 2 была введена классическая вероятность и установлено, в частности, что она удовлетворяет свойствам А1 и А2. Проверим для нее теперь справедливость А3. Начнем с рассмотрения двух несовместных событий и установим, что

$$P(A_1+A_2)=P(A_1)+P(A_2).$$

Пусть опыт имеет n элементарных исходов, из которых m_1 благоприятствует A_1, а m_2 исходов $-A_2$. Ввиду несовместности событий A_1 и A_2 , среди m_1 и m_2 нет общих исходов. В силу этого событию A_1+A_2 благоприятствуют ровно m_1+m_2 исходов, а тогда

$$P(A_1+A_2)=\frac{m_1+m_2}{n}=\frac{m_1}{n}+\frac{m_2}{n}=P(A_1)+P(A_2).$$

5.3. Утверждение п.5.2 можно распространить теперь на произвольное конечное количество событий, пользуясь ассоциативным свойством сложения и действуя по индукции. Так, например, если три события A_1, A_2, A_3 попарно несовместны, то события (A_1+A_2), и A_3 также, очевидно, будут несовместными. Теперь, на основании утверждения п. 6.1 (установленного для двух событий) получаем

$$P(A_1+A_2+A_3)=P((A_1+A_2)+A_3))=$$

$$=P(A_1+A_2)+P(A_3)=P(A_1)+P(A_2)+P(A_3),$$

так что свойство А3 оказывается верным и в случае трех слагаемых.

6. Вероятность произведения событий

6.1. Рассмотрим способ вычисления классической вероятности произведения событий. Обозначим $P_A(B)$ означают вероятность события B, вычисленную при условии, что A произошло и будем называть ее условной вероятностью события B.

Теорема 6.1 Имеет место соотношение

$$P(AB) = P(A) \cdot P_A(B). \qquad (6.1)$$

Доказательство. Предположим, что среди n элементарных исходов опыта ровно m благоприятствуют произведению $A \cdot B$, и m_A
исходов благоприятствуют событию A. Тогда

$$P(A \cdot B) = \frac{m}{n} = \frac{m_A}{n} \cdot \frac{m}{m_A} = P(A) \cdot \frac{m}{m_A}. \qquad (6.2)$$

Остается установить, что

$$P_A(B) = \frac{m}{m_A}. \qquad (6.3)$$

Действительно, по наступлении A лишь те m_A исходов, которые ему благоприятствовали, оказываются возможными для дальнейшего рассмотрения, а событие B имеет тогда благоприятными исходы, благоприятные одновременно и для A. Таким образом, по наступлении А оказывается $m_B = m$, и вычисление $P_A(B)$ приводит к равенству (6.3).
Соединив теперь равенства (6.2) и (6.3), получаем утверждение (6.1).

6.2. В случае произведения трех и большего числа событий имеет место результат, аналогичный теореме 6.1. Так, например,

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

Доказательство проводится по индукции с использованием свойства ассоциативности умножения событий.

Замечание. Соотношение

$$P_A(B) = \frac{P(AB)}{P(A)},$$

вытекающее (для классической вероятности) при $P(A) \neq 0$ из (6.1), в общем случае может быть положено в основу *определения* условной вероятности события B.

6.3. Вычисление условной вероятности $P_A(B)$ предполагает, что вероятность события В зависит от того, наступило или не наступило событие А. Так, например, если среди N предметов будет ровно M ($0<M<N$) меченых (окрашенных, бракованных и т.п.) и производится безвозвратная выборка, то

вероятность извлечения второго предмета меченым (событие B) зависит от того, меченым (событие A) или немеченым (событие $\overline{A}$) был извлечен первый предмет:

$$P_A(B) = \frac{M-1}{N-1}$$

(так как по наступлении A среди $N-1$ предметов осталось $M-1$ меченых) и

$$P_{\overline{A}}(B) = \frac{M}{N-1}$$

(по наступлении $\overline{A}$, т.е. когда первый предмет оказался немеченым, осталось M меченых предметов среди $N-1$) .

Однако, если бы первый из извлеченных предметов был возвращен в исходную совокупность, то вероятность извлечения второго предмета меченым была бы одной и той же (независимо от того, каким был первый предметов):

$$P\ (B) = P_A(B) = P_{\overline{A}}(B) = \frac{M}{N}.$$

Подобные ситуации отражаются в понятии независимых событий: два события A и B называются независимыми, если вероятность каждого из них не зависит от того, наступило ли другое событие, т.е. абсолютно постоянна в условиях данных опытов. Аналогично, события $A_1, A_2, ..., A_n$ называются независимыми в совокупности, если каждая из вероятностей $P_j = P(A_j)$ остается абсолютно постоянной в условиях данных опытов.

Из теоремы 6.1 примененной к двум независимым событиям вытекает, что

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

Распространяя результат на n событий, независимых в совокупности, приходим к соотношению

$$P(A_1 \cdot A_2 \cdot ... \cdot A_n) = P(A_1) \cdot P(A_2) \cdot ... \cdot P(A_n).$$

7. Вероятность суммы совместных событий

7.1. В параграфе 5 была рассмотрена вероятность суммы попарно несовместных событий. Следующий результат в случае суммы любых двух событий A и B является более общим.

Теорема 6.1 $P\left(A_1+A_2\right)=P\left(A_1\right)+P\left(A_2\right)-P\left(A_1A_2\right).$

В частности, для *несовместных* A_1 и A_2 имеем $P\left(A_1A_2\right)=P(\varnothing)=0$, откуда следует уже известное равенство

$$P\left(A_1+A_2\right)=P\left(A_1\right)+P\left(A_2\right).$$

7.2. В случае суммы n событий имеет место

Теорема 7.1 $P(A_1+A_2+\ldots+A_n)=1-P(\overline{A_1}\cdot\overline{A_2}\cdot\ldots\cdot\overline{A_n})$.

Доказательство вытекает из теоремы 5.1, поскольку события A и $\overline{A}$ появления хотя бы одного из указанных событий

$$A=A_1+A_2+\ldots+A_n$$

и ненаступления ни одного из них

$$\overline{A}=\overline{A_1}\cdot\overline{A_2}\cdot\ldots\cdot\overline{A_n}$$

являются, очевидно, противоположными.

7.2. Пусть теперь события $A_1, A_2, \ldots, A_n$ *независимы в совокупности*, $p_j=P(A_j)$ и $q_j=P(\overline{A_j})=1-P(A_j)$, $j=1,2,\ldots,n.$ Из теоремы 7.1 вытекает тогда соотношение

$$P(A_1+A_2+\ldots+A_n)=1-q_1\cdot q_2\cdot\ldots\cdot q_n.$$

В частности, *для двух независимых событий* A_1 и A_2 получаем

$$P(A_1+A_2)=1-q_1q_2=1-(1-p_1)(1-p_2)=p_1+p_2-p_1p_2,$$

и мы получили в этом случае утверждение п. 7.1.

8. Формула полной вероятности и формулы Бейеса

8.1. Предположим, что событие А есть результат наступления хотя бы одного из событий H_1, H_2, ..., H_n попарно несовместных и образующих полную группу, но при этом неизвестно, какое именно из H_i наступит. В этом случае события $H_1, H_2, \ldots, H_n$ называют гипотезами по отношению к A.

8.2 В условиях п. 8.1 справедливо соотношение (называемое формулой *полной вероятности*)

$$P(A) = \sum_{k=1}^{n} P(H_k)P_{H_k}(A). \qquad (8.1)$$

Доказательство. Ввиду полноты группы гипотез имеем

$$E = H_1 + H_2 + \ldots + H_n.$$

Учитывая очевидное представление $A = A \cdot E$ и свойства п.1.2 операций сложения и умножения , получаем тогда

$$A = A \cdot E = H_1 \cdot A + H_2 \cdot A + \ldots + H_n \cdot A.$$

События $H_k \cdot A,\ k = 1,2,\ldots n$ оказываются попарно несовместными (предположив противное, мы получили бы совместность гипотез), а поэтому

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

Чтобы установить (8.1), осталось применить к каждому слагаемому последней суммы теорему умножения 6.1.

В частном случае двух гипотез формула полной вероятности принимает вид

$$P(A) = P(H_1)P_{H_1}(A) + P(H_2)P_{H_2}(A).$$

8.3. Пусть выполнены условия п.8.1 и при этом событие *A наступило.* Тогда вероятность того, что событие *A оказалось следствием гипотезы именно H_k*, определяется в виде

$$P_A(H_k) = \frac{P(H_k)P_{H_k}(A)}{P(A)}, \ k=1,2,\ldots,n, \qquad (8.2)$$

где $P(A) \neq 0$ –полная вероятность (8.1). Соотношения (8.2) называют формулами Бейеса.

Доказательство (8.2) вытекает из равенства

$$A \cdot H_к = H_к \cdot A$$

(коммутативное свойство умножения), применив к которому теорему умножения, получим

$$P(A)P_A(H_k) = P(H_k)P_{H_k}(A), \quad k = 1,2,\ldots n,$$

что и равносильно (8.2).

9. Классическая вероятность: типовые задачи

Задача 9.1. Подбрасываются два игральных кубика. Найти вероятность события A, состоящего в том, что суммарное число выпавших очков не будет превосходить 4.

Р е ш е н и е. Элементарные исходы опыта можно интерпретировать как пары чисел (1, 1), (1, 2),…; попарная несовместность, полнота группы и равновозможность исходов очевидны. Найдем общее число исходов опыта n. Согласно примеру 1.2 обучающего модуля по комбинаторике, имеем n=36. Число благоприятствующих событию A исходов
исходов найдем простым их перебором: 1+1, 1+2, 1+3, 2+1,
2+2, 3+1, т.е. m_A=6. Следовательно, на основании определения классической вероятности

$$P(A) = \frac{6}{36}, \quad \text{т.е. } P(A) = \frac{1}{6}.$$

Задача 9.2. В кооперативном доме-новостройке имеется 20 квартир, среди которых 8 расположены на крайних этажах. Квартиры распределяются по жребию. Найти вероятность того, что эти 8 квартир достанутся данным восьми семьям-переселенцам.

Р е ш е н и е . Поскольку всего участников жеребьевки – 20, то элементарные исходы опыта – это упорядоченные выборки из 20 по 20, т.е. перестановки. Количество всевозможных исходов, таким образом есть n=20!

Если событие A состоит в том, что 8 квартир на крайних этажах достанутся восьми семьям-переселенцам, то число благоприятных исходов m_A можно вычислить следующим образом. Среди данных восьми семей возможно 8! способов распределения квартир на крайних этажах, и каждый такой способ сочетается с остальными 12! способами распределения квартир среди остальных 12 участников жеребьевки. Согласно принципу умножения $m_A = 8! \cdot 12!$ Итак,

$$P(A) = \frac{m_A}{m} = \frac{8!12!}{20!} = \frac{1}{125970}.$$

Задача 9.3. Имеются 10 предметов, среди которых ровно 5 меченых. Случайным образом извлекают 5 предметов. Какова вероятность, что все они меченые?

Р е ш е н и е. Элементарными исходами опыта являются всевозможные выборки – сочетания из 10 предметов по 5. Найдем общее число элементарных исходов:

$$n = C_{10}^{5} = \frac{10!}{5!\cdot(10-5)!} = \frac{5!\cdot 6\cdot 7\cdot 8\cdot 9\cdot 10}{5!\cdot 5!} = 252.$$

Благоприятным для события A – извлечения пяти меченых предметов – будет единственный исход :

$$m = m_A = C_5^5 = 1.$$

Итак, $P(A) = \dfrac{1}{252}$.

Задача 9.4. В урне 6 белых и 4 черных шара. Наудачу извлекли два шара. Найти вероятность следующих событий:

а) оба шара белых;

б) только один шар белый;

в) хотя бы один шар белый.

Р е ш е н и е . а) Событие A – оба извлеченных шара – белые. Элементарные исходы опыта – выборки.

1-й способ (с использованием (2.1) и комбинаторных формул). Так как набор из двух шаров неупорядочен, то число возможных исходов

$$n = C_{10}^{2} = \frac{10!}{8!2!} = 45\,.$$

Число благоприятных исходов

$$m = C_6^2 = \frac{6!}{4!2!} = 15$$

(выбор двух шаров из шести белых).Следовательно,

$$P(A) = \frac{C_6^2}{C_{10}^2} = \frac{15}{45} = \frac{1}{3}.$$

2-й способ (с использованием теоремы 6.1). Рассмотрим события: A_1 - первый извлеченный шар - белый, A_2 - второй шар - белый. Тогда

$$P(A_1)=\frac{6}{10}=\frac{3}{5}, P_{A_1}(A_2)=\frac{5}{9}.$$

При этом $A=A_1A_2$, то есть

$$P(A_1A_2)=P(A_1)P_{A_1}(A_2)=\frac{3}{5}\cdot\frac{5}{9}=\frac{1}{3}.$$

б) Событие B – только один извлеченный шар – белый.

1-й способ (с использованием (2.1) и комбинаторных формул). Число всевозможных элементарных исходов $n=C_{10}^2=45$. Число благоприятных исходов $m=6\cdot 4=24$; действительно, среди извлеченных белый шар берется из шести, так что $m_1=C_6^1=6$; черный шар берется из четырех: $m_2=C_4^1=4$; в силу принципа умножения теперь $m=m_1m_2=6\cdot 4=24$. Итак,

$$P(B)=\frac{24}{45}=\frac{8}{15}.$$

2-й способ (с использованием формул сложения и умножения). Рассмотрим события: A_1 – извлечен первым белый шар, A_2 – второй шар – белый. Поскольку надо извлечь первым белый, а вторым – черный шар (событие $A_1\overline{A}_2$) или наоборот (событие $\overline{A}_1A_2$), то $B=A_1\overline{A}_2+\overline{A}_1A_2$. Теперь

$$P(A_1)=\frac{6}{10},\ P(\overline{A}_1)=\frac{4}{10},\ P_{A_1}(\overline{A}_2)=\frac{4}{9},\ P_{\overline{A}_1}(A_2)=\frac{6}{9}.$$

События $A_1\overline{A}_2$ и $\overline{A}_1A_2$ несовместны. Следовательно,

$$P(B)=P(A_1\overline{A}_2+\overline{A}_1A_2)=P(A_1\overline{A}_2)+P(\overline{A}_1A_2)=$$

$$=P(A_1)P_{A_1}(\overline{A}_2)\ +P(\overline{A}_1)P_{\overline{A}_1}(A_2)=\frac{6}{10}\cdot\frac{4}{9}+\frac{4}{10}\cdot\frac{6}{9}=\frac{8}{15}.$$

в) C - хотя бы один шар белый.

1-й способ (с использованием (1.2) и комбинаторных формул). Число всевозможных элементарных исходов $n=C_{10}^2=45$, число благоприятных исходов $m=C_6^2+C_6^1\cdot C_4^1=39$. Действительно, событие C означает, что извлечены оба белых шара или один белый и один черный; оба белых шара можно взять $m_1=C_6^2=15$ различными способами; один белый и один черный шар можно взять $m_2=C_6^1C_4^1=24$ различными способами. Итак,

$$P(C)=\frac{39}{45}=\frac{13}{15}.$$

2-й способ (с использованием формул сложения и умножения). Рассмотрим события: A_1 - первый извлеченный шар белый, A_2 - второй шар - тоже белый. Тогда $C=A_1A_2+A_1\overline{A}_2+\overline{A}_1A_2$. События A_1A_2, $A_1\overline{A}_2$, $\overline{A}_1A_2$ - попарно несовместны. Значит

$$P(C)=P(A_1A_2+A_1\overline{A}_2+\overline{A}_1A_2)=P(A_1A_2)+P(A_1\overline{A}_2)+(\overline{A}_1A_2)=$$

$$=P(A_1)P_{A_1}(A_2)+P(A_1)P_{A_1}(\overline{A}_2)+P(\overline{A}_1)P_{\overline{A}_1}(A_2)=\frac{6}{10}\cdot\frac{5}{9}+\frac{6}{10}\frac{4}{9}+\frac{4}{10}\cdot\frac{6}{9}=\frac{13}{15}.$$

3-й способ (использование теоремы 7.1):

$$P(C)=1-P(\overline{A}_1)P_{\overline{A}_1}(\overline{A}_2)=1-\frac{4}{10}\cdot\frac{3}{9}=\frac{13}{15}.$$

Задача 9.5. На склад поступают детали заводов № 1 и № 2. Первый завод производит 80% стандартных изделий, завод № 2 – 60%. Наудачу взяли по одной детали каждого завода. Найти вероятности следующих событий: а) обе детали стандартны; б) только одна деталь стандартна; в) хотя бы одна деталь стандартна.

Р е ш е н и е. а) Событие A – обе детали стандартны. Введем события A_1 – произведенная заводом № 1 деталь стандартна, A_2 – деталь завода № 2 стандартна; следовательно, $A=A_1A_2$. События A_1 и A_2 , очевидно, независимы; поэтому

$$P(A)=P(A_1)P(A_2)=0{,}8\cdot 0{,}6=0{,}48.$$

б) Событие B – только одна деталь стандартна; $B=A_1\overline{A}_2+\overline{A}_1A_2$. События A_1 и $\overline{A}_2$, $\overline{A}_1$ и A_2 – независимы, $A_1\overline{A}_2$ и $\overline{A}_1A_2$ несовместны. Тогда

$$P(B)=P(A_1)P(\overline{A}_2)+P(\overline{A}_1)P(A_2)=0{,}8\cdot 0{,}4+0{,}2\cdot 0{,}6=0{,}44.$$

в) C – хотя бы одна деталь стандартна; очевидно, что $C=A_1+A_2$, причем события A_1 и A_2 – совместны; следовательно

$$P(C)=P(A_1)+P(A_2)-P(A_1)P(A_2)=0{,}8+0{,}6-0{,}48=0{,}92.$$

Задача 9.6 По самолету выпущена ракета. Самолет может уничтожить ее с вероятностью 0,6. Если этого сделать не удается, то ракета поражает самолет с вероятностью 0,8. Какова вероятность поражения самолета ракетой?

Р е ш е н и е . Поражение самолета – сложное событие С, состоящее в совместном наступлении события А – неуничтожении ракеты и события В – поражения в этом случае самолета ракетой. Следовательно, С=АВ. Теперь

$$P(C) = P(A)P_A(B).$$

Перейдем к нахождению вероятностей. Событие А противоположно событию уничтожения ракеты, следовательно,

$$P(A) = 1 - 0{,}6 = 0{,}4; \quad \text{при этом} \quad P_A(B){=}0{,}8.$$

Следовательно,

$$P(C) = 0{,}4\cdot),8 = 0{,}32$$

Задача 9.7. На конвейер поступают 20 % изделий цеха № 1 и 80 % – цеха № 2. Вероятность брака для изделия из первого цеха P_1 = 0,1; для второго цеха – P_2 = 0,05.

l) Какова вероятность, что наугад взятое изделие - бракованное ?

2) Случайно выбранное изделие оказалось с браком. Какова вероятность, что его произвел цех № 1 ?

Р е ш е н и е . Пусть A – событие, состоящее в обнаружении брака в случайно отобранном изделии. Возможны гипотезы: H_1 – выбранное изделие произведено цехом №1, H_2 – цехом № 2. Очевидно, что выполнены условия несовместности и полноты группы H_1 и H_2. Здесь $P(H_1)$ = 0,2; $P(H_2)$ = 0,8; $P_{H_1}(A) = 0{,}1$; $P_{H_2}(A) = 0{,}05$. Тогда:

1)по формуле полной вероятности

$$P(A) = P(H_1)P_{H_1}(A) + P(H_2)P_{H_2}(A) = 0{,}2\cdot 0{,}1 + 0{,}8\cdot 0{,}05 = 0{,}06;$$

2)по формулам Бейеса для первой из гипотез

$$P_A(H_1) = \frac{P(H_1)P_{H_1}(A)}{P(A)} = \frac{0{,}2\cdot 0{,}1}{0{,}06} = \frac{1}{3}.$$

Задача 9.8 Компьютерной диагностике подвергается группа участников диспансеризации, среди которых 10% страдают некоторыми заболеваниями.

В результате диагностики болезнь выявляется с вероятностью

0,95, и с вероятностью, равной 0,03 здоровый участник признается больным. У произвольно выбранного протестированного участника компьютер выявил заболевание. Какова вероятность, что произошла ошибка?

Р е ш е н и е . Пусть событие A состоит в том, что протестированный участник признан больным. Возможны предположения (гипотезы):

H_1 – тестируется участник, страдающий заболеванием;

H_2 – тестируется здоровый участник.

Требуется найти вероятность гипотезы H_2, при условии, что наступило событие A; следовательно применима формула Бейеса

$$P_A(H_2) = \frac{P(H_2)P_{H_2}(A)}{P(A)}.$$

По условию задачи гипотезы имеют вероятности $P(H_1)$= 0,1 и $P(H_2)$= 0,9. Соответствующие условные вероятности события имеют вид

$$P_{H_1}(A) = 0{,}95;\ P_{H_2}(A) = 0{,}03.$$

Вероятность события А находим по формуле полной вероятности

$$P(A) = P(H_1)P_{H_1}(A) + P(H_2)P_{H_2}(A) = 0{,}1 \cdot 0{,}95 + 0{,}9 \cdot 0{,}03 = 0{,}122;$$

теперь

$$P_A(H_2) = \frac{0{,}9 \cdot 0{,}03}{0{,}122} = \frac{27}{122}.$$

10.Повторение опытов. Формула Бернулли

10.1. Пусть один и тот же опыт повторяется n раз, и при этом вероятность наступления события A в каждом таком опыте остается неизменной, равной некоторому p; пусть $q = 1 - p$. Описанная ситуация независимых опытов называется *схемой Бернулли*.

Обозначим через $P_n(k)$ вероятность того, что ***событие А появится ровно k раз в n опытах.*** В этом случае говорят также о k «успехах» в n опытах. Имеет место следующая

***Формула Бернулли*:**

$$P_n(k) = C_n^k p^k q^{n-k}. \tag{10.1}$$

Доказательство. Фиксируем какие-либо k опытов. Наступление события A ровно k раз (т.е. в этих в k опытах) происходит совместно с его ненастулением (наступлением $\overline{A}$) $n-k$ раз. Любая такая фиксированная комбинация $A_{n,k}$ имеет, очевидно, вероятность

$$P(A_{n,k}) = \underbrace{p\cdot\ldots\cdot p}_{k}\cdot\underbrace{q\cdot\ldots\cdot q}_{n-k} = p^k\cdot q^{n-k}.$$

Событие $B_{n,k}$, состоящее в том, что A появится ровно k раз равносильно наступлению какого-либо из описанных (попарно несовместных) $A_{n,k}$, т.е. их сумме. Тогда

$$P_n(k) = P(B_{n,k}) = p^k\cdot q^{n-k} + p^k\cdot q^{n-k} + \ldots p^k\cdot q^{n-k},$$

причем слагаемых в этой сумме столько, сколько существует способов сформировать различные комбинации типа $A_{n,k}$. Это количество равно C_n^k – числу (неупорядоченных) выборок из n по k. Таким образом, слагаемое $p^k\cdot q^{n-k}$ повторяется C_n^k раз и мы приходим к (10.1).

10.2 . В условиях пункта 10.1 ***вероятность наступления события*** A ***в*** n ***опытах от*** k_1 ***до*** k_2 ***раз***, есть

$$P_n(k_1 \le k \le k_2) = \sum_{k=k_1}^{k_2} C_n^k p^k q^{n-k} \qquad (10.2)$$

Для доказательства (10.2) заметим, что $P_n(k_1 \le k \le k_2)$ есть вероятность наступления хотя бы одного из попарно несовместных событий B_{n,k_1}, $B_{n,k_1+1},\ldots,$ B_{n,k_2} где $B_{n,j}$ – событие, состоящее в наличии j « успехов» (в n опытах). Таким образом, $P_n(k_1 \le k \le k_2)$ будет равна сумме соответствующих вероятностей Бернулли, а это и означает справедливость (10.2).

10.3. Число m_o появления события называется *наивероятнейшим*, если вероятность появления события m_o раз в n испытаниях превышает (или, по крайней мере, не меньше) вероятности остальных возможных исходов испытаний. Наивероятнейшее число m_o (или наивероятнейшие числа, поскольку их может быть и два) может быть определено из двойного неравенства

$$(n+1)p - 1 \le m_o \le (n+1)p \qquad (10.3)$$

10.4. Пусть теперь от опыта к опыту вероятность наступления события A может меняться, так что в первом опыте она равна p_1, во втором – p_2, в n-ом – p_n; обозначим соответствующие вероятности непоявления A

через $q_1 = 1 - p_1,\ q_2 = 1 - p_2, \ldots, q_n = 1 - p_n$. Тогда вероятность наступления события A в этих n опытах ровно k раз оказывается равной коэффициенту при z^k в многочлене

$$\varphi_n(z) = \prod_{j=1}^{n} (q_j + p_j z),$$

называемом производящей функцией.

Если же все $p_j = p$ одинаковы и, соответственно, одинаковы все $q_j = 1 - p_j = q$, то

$$\varphi_n(z) = (q + pz)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k} z^k$$

согласно формуле разложения бинома Ньютона. Коэффициент при z^k здесь совпадает с правой частью (10.1). Таким образом, результат настоящего пункта служит обобщением формулы Бернулли.

11. Предельные теоремы в схеме Бернулли

11.1. С ростом n ($n \to \infty$) использование формулы Бернулли и ее следствий становится затруднительным с точки зрения громоздкости вычислений; нужны, следовательно, приближенные формулы, которые обходят эту трудность.

При больших значениях n и $0 < p < 1$ ***значение*** вероятности Бернулли $P_n(k)$ можно ***приближенно вычислить по "локальной" формуле Лапласа***

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x_{n,k}),$$

где

$$x_{n,k} = \frac{k - np}{\sqrt{npq}},\ \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Значения функции $\varphi(x)$ находятся по табл. 1 (см. приложения). При этом используется свойство четности функции φ : $\varphi(-x) = \varphi(x)$.

Указанная приближенная формула при $n \to \infty$ $|x_{n,k}| \le const$ вытекает из представления (асимптотической формулы)

$$P_n(k)=\frac{1}{\sqrt{npq}}\varphi(x_{n,k})\cdot(1+\varepsilon_{n,k}) \quad , \text{где} \quad |\,\varepsilon_{n,k}\,|<\frac{const}{\sqrt{n}}. \qquad (11.1)$$

11.2. Если количество n опытов велико и $0<p<1$, то вероятность $P_n(k_1\le k\le k_2)$ того, что ***событие A произойдет не менее k_1 и не более k_2 раз, может быть найдена по приближенной интегральной формуле Лапласа***

$$P_n\left(k_1\le k\le k_2\right)\approx\Phi\left(x_{k_2}\right)-\Phi\left(x_{k_1}\right),$$

где

$$x_{k_1}=\frac{k_1-np}{\sqrt{npq}},\quad x_{k_2}=\frac{k_2-np}{\sqrt{npq}},\quad \Phi(x)=\int\limits_0^x\varphi(t)dt.$$

Значения $\Phi(x)$ находят по табл. 1 (см. приложение). При этом используется свойство нечетности $\Phi(x)$: $\Phi(-x) = -\Phi(x)$.

Приближенная интегральная формула Лапласа вытекает из асимптотического представления (11.1) вероятности Бернулли и следствия (10.2) формулы Бернулли.

11.3. Установим, что вероятность отклонения относительной частоты $w(A)$ события A от его вероятности $p(A)$ может быть найдена по приближенной формуле

$$P(|\frac{m}{n}-p|<\varepsilon)\approx 2\Phi(\varepsilon\sqrt{\frac{n}{pq}}), \qquad (11.2)$$

где Φ - как и выше - интегральная функция Лапласа.

Для доказательства (11.2) запишем неравенство

$$|\frac{m}{n}-p|<\varepsilon$$

в равносильном виде

$$n(p-\varepsilon)<m<n(p+\varepsilon)$$

и заметим, что m есть число наступлений события A в n опытах, а поэтому можно применить приближенную интегральную формулу Лапласа с

$$k_1=n(p-\varepsilon),\ k_2=n(p+\varepsilon).$$

Имеем тогда

$$x_{k_1}=\frac{k_1-np}{\sqrt{npq}}=\frac{-n\varepsilon}{\sqrt{npq}}=-\varepsilon\sqrt{\frac{n}{pq}} \quad \text{и} \quad x_{k_2}=\frac{n\varepsilon}{npq}=\varepsilon\sqrt{\frac{n}{pq}}.$$

Следовательно,

$$P(|\frac{m}{n}-p|<\varepsilon)=\Phi(\varepsilon\sqrt{\frac{n}{pq}})-\Phi(-\varepsilon\sqrt{\frac{n}{pq}}),$$

откуда, в силу нечетности интегральной функции Лапласа, и будет вытекать соотношение (11.2).

11.4. Из результатов п.11.3 может быть получен так называемый закон больших чисел Бернулли.

Теорема. Пусть $w_n(A)=\frac{m_A}{n}$ – относительная частота события A в n опытах и $p(A)$ – его вероятность. Тогда для любого $\varepsilon>0$ имеет место соотношение

$$\lim_{n\to\infty}P(|\frac{m_A}{n}-p(A)|<\varepsilon)=1.$$

Утверждение следует из (11.2) и того факта, что при $n\to\infty$ его правая часть стремится к значению $2\Phi(\infty)=2\cdot 0{,}5=1$.

Закон Бернулли есть математически строгая форма свойства устойчивости относительной частоты, описанного выше. Говорят также, что ***относительная частота события*** A ***сходится по вероятности*** (при $n\to\infty$) ***к вероятности*** $p(A)$.

11. 5. Пусть в каждом опыте вероятность события A постоянна, но от серии к серии опытов (с ростом их количества n) значение произведения $\lambda=np$. При сформулированных условиях имеет место соотношение

$$\lim_{n\to\infty}P_n(k)=\frac{\lambda^k\cdot e^{-\lambda}}{k!}\quad. \tag{11.3}$$

Для его доказательства представим вероятность Бернулли в виде

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

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

$$= \frac{\lambda^k}{k!} \cdot \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \right) \cdot \left(\left(1 - \frac{\lambda}{n} \right)^{-\frac{n}{\lambda}} \right)^{-\frac{\lambda}{n} \cdot (n-k)} . \qquad (11.4)$$

При вычислении предела при $n \to \infty$ заметим, что отношение $\dfrac{\lambda^k}{k!}$ остается постоянным, произведение же в скобках содержит k множителей, каждый из которых стремится к единице. Последний множитель содержит в скобках выражение вида

$$(1+t)^{1/t}, \qquad \text{где} \quad t = -\frac{\lambda}{n}, \text{ так что} \quad t \to 0, \text{ если } n \to \infty .$$

Тогда в пределе этого выражения получаем (второй замечательный предел) число e. Показатель же степени

$$-\frac{\lambda}{n} \cdot (n-k) = -\lambda + \frac{k}{n}$$

при этом стремится к значению $-\lambda$.

Таким образом произведение в правой части (11.4) стремится к значению

$$\frac{\lambda^k}{k!} \cdot 1^k \cdot e^{-\lambda},$$

чем и доказано утверждение (11.3).

Установленный результат означает, что если вероятность p появления события A в каждом из n опытов мала, а при этом n - велико и $\lambda = np$, то имеет место приближенная формула

$$P_n(k) \approx \frac{\lambda^k e^{-\lambda}}{k!}.$$

Вероятность же того, что при достаточно большом числе опытов событие произойдет не менее m_1 и не более m_2 раз можно найти (ср.п.10.2) по приближенной формуле

$$P_n\left(m_1 \le m \le m_2\right) \approx \sum_{k=m_1}^{m_2} \frac{\lambda^k}{k!} e^{-\lambda}.$$

12. Схема Бернулли: типовые задачи

Задача 12.1. Найти вероятность того, что событие A появится в пяти независимых опытах: а) два раза; б) менее двух раз, если вероятность появления события A в одном опыте $p = 0{,}4$.

Р е ш е н и е . а) Пусть событие B состоит в появлении A ровно два раза в пяти опытах. Тогда по формуле Бернулли

$$P(B) = P_5(2) = C_5^2(0{,}4)^2 \cdot (0{,}6)^2 = 0{,}3456.$$

б) Если событие C означает появление A менее двух раз, то есть или ни разу ($k = 0$) или один раз ($k = 1$), то согласно п.3.2

$$P(C) = P_5(0) + P_5(1) = C_5^0 \cdot 0{,}4^0 \cdot 0{,}6^5 + C_5^1 \cdot 0{,}4 \cdot 0{,}6^4 =$$
$$= 0{,}14256 + 0{,}4752 = 0{,}61776.$$

Задача 12.2. Что вероятнее: выиграть у равносильного соперника две шахматные партии из четырех или три из шести?

Р е ш е н и е . Имеем схему Бернулли, в которой вероятность события в единичном опыте (выигрыша одной партии) $p = \frac{1}{2}$, $q = 1 - p = \frac{1}{2}$. Теперь надо сравнить вероятности $P_4(2)$ и $P_6(3)$. Имеем

$$P_4(2) = C_4^2(\frac{1}{2})^2(\frac{1}{2})^2 = \frac{6}{16}, \quad P_6(3) = C_6^3(\frac{1}{2})^3(\frac{1}{2})^3 = \frac{5}{16}.$$

Следовательно, вероятность выиграть 2 партии из 4 – выше.

Задача 12.3. Тестируется каждый из 15 блоков некоторого электронного прибора. Вероятность пройти тест для каждого блока составляет 0,9. Найти наивероятнейшее число выдержавших тестирование блоков и его вероятность.

Р е ш е н и е. Так как n =15, p = 0,9, то по формуле (10.3) имеем

$$16 \cdot 0{,}9 - 1 \le m_o \le 16 \cdot 0{,}9 \quad \text{т.е.} \quad m_o = 14.$$

Теперь осталось найти

$$P_{15}(14) = C_{15}^{14}(09)^{14}(0{,}1)^1 \approx 0{,}343.$$

Задача 12.4. Вероятность того, что на странице рукописи имеется опечатка, равна 0,2. Найти вероятность того, что на 400 страницах будет ровно 104 опечатки.

Для решения задачи можно использовать формулу Лапласа п. 3.3, где

$n = 400$, $p = 0{,}2$, $q = 0{,}8$; тогда

$$x_{n,k} = \frac{104 - 80}{8} = 3.$$

По табл. 1 Прил. находим $\varphi(3) = 0{,}0044$. Имеем:

$$P_{400}(104) \approx \frac{1}{\sqrt{400 \cdot 0{,}2 \cdot 0{,}8}} \varphi(3) = \frac{0{,}0044}{8} = 0{,}00055.$$

Задача 12.5. Вероятность того, что на странице рукописи есть опечатка равна 0,2. Найти вероятность того, что на 400 страницах будет не менее 72 и не более 104 опечаток.

Р е ш е н и е . Используем интегральную формулу Лапласа, в которой $n = 400$, $p = 0{,}2$, $q = 0{,}8$; тогда т

$$x_{n,k_1} = \frac{72 - 80}{8} = -1, \quad x_{n,k_2} = \frac{104 - 80}{8} = 3.$$

По табл. 1 Прил. находим $\Phi(-1) = -\Phi(1) = -0{,}3413$, $\Phi(3) = 0{,}4986$. Следовательно,

$$P_{400}(72 \le k \le 104) = \Phi(3) - \Phi(-1) = \Phi(3) + \Phi(1) = 0{,}8399.$$

Задача 12.6 Сколько раз надо бросить монету, чтобы с вероятностью р=0,9544 утверждать, что относительная частота выпадения герба отклонилась от 0,5 не более, чем на 0,05?

Р е ш е н и е . Согласно приближенной формуле (11.2) имеем

$$0{,}9544 = P\left(\left|\frac{m}{n} - 0{,}5\right| < 0{,}05\right) \approx 2\Phi\left(0{,}05\sqrt{\frac{n}{0{,}5(1-0{,}5)}}\right),$$

где m – число появлений герба при n подбрасываниях монеты. Отсюда

$$0{,}9544 \approx 2\Phi(0{,}05\sqrt{4n}) \quad \text{или} \quad \Phi(0{,}1\sqrt{n}) \approx 0{,}4772.$$

По таблице значений интегральной функции Лапласа (см. таблицу 1 приложений) находим значение аргумента интегральной функции Лапласа: $0{,}1\sqrt{n} \approx 2$ откуда $n \approx 400$. Итак, монету надо бросить 400 раз.

Задача 12.7. Вероятность того, что абонент неправильно наберет телефонный номер, принимается для всех абонентов равной 0,001. Определить вероятность того, что среди 500 произведенных независимо один от другого вызовов, ровно один абонент наберет неправильно телефонный номер.

Р е ш е н и е . Имеем: $n = 500$, $p = 0{,}001$, $\lambda = np = 0{,}5$. По табл. 2 Прил. находим $e^{-0{,}5} \approx 0{,}6065$, тогда по формуле Пуассона

$$P_{500}(1) \approx \frac{0,5 \cdot 0,6065}{1!} = 0,30325\,.$$

Задача 12.8. В течение года из аэропорта города N отправляется 1200 авиарейсов. Вероятность задержки каждого вылета по метеоусловиям равна 0,005. Какова вероятность задержки по метеоусловиям в течение года не менее 2 рейсов?

Р е ш е н и е. По условию задачи, вероятность наступления события в единичном опыте (задержки рейса по метеоусловиям) мала: p=0,005, тогда как число опытов (число рейсов) велико: n=1500. Следовательно, в расчетах возможно использование формулы Пуассона. Событие задержки не менее 2 рейсов противоположно событию задержки $k \le 1$ рейсов. Имеем $\lambda = np = 1200 \times 0,005 = 6$ и

$$p_{1500}(0 \le k \le 1) = p_{1500}(0) + p_{1500}(1) = \frac{6^0 e^{-6}}{0!} + \frac{6^1 e^{-6}}{1!} = 7e^{-6},$$

т.е. $p_{1500}(k \ge 2) = 1 - 7e^{-6} \approx 0,97.$

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

1 . Ряд распределения дискретной случайной величины. Числовые характеристики

1.1. ***Случайной величиной*** называется числовая величина X , которая в каждом опыте принимает одно и только одно значение, заранее неизвестное и зависящее от случайных причин. Если все возможные значения величины X можно записать в виде числовой последовательности $\{x_n\}$, n = 1, 2, ... (конечной или бесконечной), то X называется дискретной; если же возможные значения X заполняют целиком некоторый числовой интервал, то величина X называется непрерывно распределенной на этом интервале.

Примером дискретной случайной величины является ежедневно фиксируемый рублевый курс доллара, непрерывной – время t загрузки файла, скачиваемого из Интернета: $t \in (0,\infty)$.

1.2. ***Законом распределения дискретной случайной величины*** X называется соответствие между ее возможными значениями x_k и вероятностями $p_k = P(X = x_k)$ события, состоящего в принятии величиной X значения именно x_k .

Обычный способ задания такого закона – ряд (таблица) распределения, который в случае конечного числа n значений величины X (записанных в порядке возрастания) имеет вид

X	x_1	x_2	...	x_k	...	x_n
P	p_1	p_2	...	p_k	...	p_n

Заметим, что

$$\sum_{k=1}^{n} p_k = 1. \tag{1.1}$$

Действительно, все события вида $A_k = \{X = x_k\}$ образуют полную группу, так что

$$A_1 + A_2 + \ldots + A_n = E,$$

и, по свойству А1,

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

Благодаря попарной несовместности событий A_k и свойству А3, получаем тогда из последнего равенства утверждение (1.1).

Возможно также рассмотрение ряда распределения с бесконечным набором $\{x_1, x_2, \ldots x_n, \ldots\}$ значений X. В этом случае для соответствующих вероятностей имеет место соотношение

$$\sum_{k=1}^{\infty} p_k = 1.$$

При этом мы рассматриваем распределения, для которых записанный числовой ряд является сходящимся.

1.3. Наряду с законом (рядом) распределения часто бывает удобно пользоваться числами, которые описывают случайную величину суммарно; такие числа называются ***числовыми характеристиками*** случайной величины. Они помогают «в сжатой форме» выразить наиболее существенные черты распределения. Основными числовыми характеристиками случайной величины являются математическое ожидание, характеризующее среднее значение случайной величины, и дисперсия, характеризующая степень рассеяния случайной величины относительно ее математического ожидания.

Ограничимся рассмотрением дискретной случайной величины с конечным набором возможных значений.

Математическое ожидание $M(X)$ дискретной величины X определяется в виде

$$M(X) = \sum_{k=1}^{n} x_k p_k \, , \tag{1.2}$$

дисперсия $D(X)$ – в виде

$$D(X) = \sum_{k=1}^{n} \left(x_k - M(X)\right)^2 p_k \tag{1.3}$$

Из свойства А1 неотрицательности вероятностей следует, что всегда $D(X) \geq 0$, а тогда можно рассмотреть характеристику рассеяния вида

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

называемую ***средним квадратическим отклонением*** . Она предпочтительней дисперсии тогда, когда желательно чтобы оценка рассеяния имела размерность случайной величины (заметим, что дисперсия $D(X)$ имеет размерность квадрата случайной величины, что не всегда удобно).

1.4. Дисперсию можно вычислить также по формуле

$$D(X) = \sum_{k=1}^{n} (x_k)^2 p_k - (M(X))^2 , \tag{1.4}$$

для доказательства которой выполним в (1.3) следующие преобразования: раскроем скобки, возводя разность в квадрат, перегруппируем слагаемые, записав три соответствующие суммы и воспользуемся определением (1.2) и свойством вероятностей (1.1). Тогда будем иметь

$$D(X) = \sum_{k=1}^{n} (x_k^2 p_k - 2M(X) \cdot x_k p_k + (M(X))^2 \cdot p_k) =$$

$$= \sum_{k=1}^{n} x_k^2 p_k - 2M(X) \cdot \sum_{k=1}^{n} x_k p_k + (M(X))^2) \cdot \sum_{k=1}^{n} p_k =$$

$$= \sum_{k=1}^{n} x_k^2 p_k - 2(M(X))^2 + (M(X))^2 \cdot 1 = \sum_{k=1}^{n} x_k^2 p_k - (M(X))^2 ,$$

чем и установлено (1.4).

2. Функция распределения

2.1. Универсальным способом описания всякой случайной величины X является ***функция распределения*** (синонимы: интегральный закон распределения, интегральная функция), имеющая вид

$$F(x) = P(X < x). \qquad (2.1)$$

Она соотносит каждому $x \in (-\infty; +\infty)$ вероятность события, состоящая в принятии величиной X значения левее точки x (рис. 2.1)

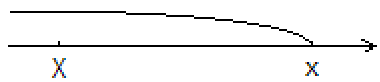

Рис. 2.1. Интервал значений случайной величины

2.2. Рассмотрим функцию распределения дискретной случайной величины, обладающей конечным перечнем значений $\{x_1, x_2, \ldots x_n\}$. Утверждается, что

$$F(x) = \sum_{k | x_k < x} p_k \ ; \qquad (2.2)$$

суммирование в (2.2) проводится по тем k, для которых соответствующие x_k оказываются меньшими x. Чтобы установить соотношение (2.2), отметим расположим на числовой оси значения $x_1, x_2, \ldots, x_n$. Если аргумент x удовлетворяет соотношению $x \le x_1$, то событие $X < x$ является невозможным, и следовательно, $F(x) = 0$. В случае $x_1 \le x < x_2$ событие $X < x$ равносильно событию $X = x_1$, а тогда $F(x) = p_1$. Если $x_2 \le x < x_3$, то $F(x) = p_1 + p_2$.

Рассуждая таким образом и далее, мы и получаем (2.2). В частности, при $x_n < x$, имеем

$$F(x) = p_1 + p_2 + \ldots + p_n = 1.$$

В «развернутой» форме (2.2) принимает вид

$$F(x)=\begin{cases} 0, \quad x\le x_1 \\ p_1, \quad x_1<x\le x_2 \\ p_1+p_2, \; x_2<x\le x_3 \\ \ldots \\ p_1+p_2+\ldots p_{n-1}, \; x_{n-1}<x\le x_n \\ 1, \qquad x>x_n \end{cases}.$$

Мы получаем неубывающую функцию, значения которой расположены в промежутке $[0,\ 1]$, непрерывную в каждой точке слева. Ее предел при на $-\infty$ равен 0, а при стремлении к $+\infty$ равен 1. Эти свойства, как мы увидим ниже, распространяются на случай функции распределения произвольной случайной величины.

2.3. Итак, установим что в общем случае функции распределения $F(x)=P(X<x)$ обладает свойствами:

а) $0\le F(x)\le 1$;

б) $P(a\le X<b)=F(b)-F(a)$;

в) $F(x)$ - неубывающая функция.

Доказательство. Свойство а) очевидно, так как $F(x)$ – это вероятность. Для доказательства б) представим значение $F(b)$ виде

$$F(b)=P(X<b)=P(\{X<a\} \text{ или } \{a\le X<b\}).$$

Ввиду несовместности соответствующих событий и свойству вероятностей А3 получим

$$F(b)\ =P(X<a)+P(a\le x<b).$$

Таким образом,

$$F(b)=F(a)+P(a\le x<b),$$

а это и равносильно свойству б). Наконец, из последнего соотношения и оценки $F(a)\ge 0$ получаем, что $F(b)\ge F(a)$ при $b>a$, а это и означает неубывание функции распределения.

2.4. Следующие свойства относятся к случаю непрерывной $F(x)$. В дальнейшем, говоря о непрерывной случайной величине, будем предполагать, что $F(x)$ также и кусочно-дифференцируема.

а)Соотношение

$$P(X = x) = 0$$

имеет место для любого действительного *x*. В частности,

$$\begin{aligned} &P(a \le X \le b) = P(a < X \le b) = P(a < X < b) = \\ &= P(a \le X < b) = F(b) - F(a) \end{aligned}, \qquad (2.3)$$

б) $F(+\infty) = 1$ и $F(-\infty) = 0$, где, по определению,

$$F(\pm\infty) = \lim_{x \to \pm\infty} F(x).$$

Оба свойства основаны на возможности предельного перехода под знаком функции $F(x)$ (что вытекает из ее непрерывности) . Так, например, свойство а) есть результат предельного перехода (при $\Delta x \to 0$) в равенстве (см. б) в п. 2.3)

$$P(x \le X < x + \Delta x) = F(x + \Delta x) - F(x);$$

соотношения (2.3) следуют из того же свойства б) п. 2.3, поскольку вероятность принятия непрерывной случайной величиной каждого конкретного значения равна нулю.

3. Плотность распределения

3.1. ***Плотностью распределения*** (плотностью вероятности или дифференциальной функцией) назовем функцию вида

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

Согласно принятому выше соглашению для всякой непрерывной случайной величины *X* плотность распределения существуют (хотя бы в виде односторонней производной) в каждой точке *x*.

Использование термина «*плотность*» оправдано следующими соображениями. По определению производной имеем (при выборе $\Delta x > 0$)

$$f(x) = F'(x) = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x}. \qquad (3.1)$$

Отношение, записанное под знаком предела (3.1) напоминает среднюю линейную плотность массы, сосредоточенной в промежутке $(x, x + \Delta x)$, и называется поэтому средней плотностью вероятности (соответствующей значениям случайной величины, заключенным в указанном промежутке); само же значение предела является как бы «плотностью вероятности в точке «*x*».

В соответствии с (3.1) при малых Δx справедливо приближенное равенство

$$P\big(x < X < x + \Delta x\big) \approx f(x) \cdot \Delta x .$$

3.2. Имеют место следующие свойства плотности распределения:

а) $f(x) \geq 0$ для всех x;

б) $P(a < X < b) = \int\limits_{a}^{b} f(x)dx;$

в) функция распределения $F(x)$ восстанавливается (по известной дифференциальной функции) в виде

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

г) $\int\limits_{-\infty}^{+\infty} f(x)dx = 1$ (свойство нормированности плотности).

Свойство а) очевидно, т.к. $f(x)$ есть производная неубывающей функции.

Докажем соотношение б). Согласно определению 3.1, функция распределения $F(x)$ является одной из первообразных для $f(x)$. Тогда применима формула Ньютона-Лейбница

$$\int\limits_{a}^{b} f(x)dx = F(b) - F(a),$$

в правой части которой мы и узнаем значение $P\big(a < X < b\big)$, если воспользоваться соотношениями (2.3).

Равенства в) и г) вытекают из свойства б) и определения несобственных интегралов; заметим, в частности, что интеграл по «промежутку» $(-\infty, +\infty)$ здесь понимается в смысле так называемого главного значения:

$$\int\limits_{-\infty}^{+\infty} f(x)dx = \lim_{T \to +\infty} \int\limits_{-T}^{T} f(x)dx .$$

3.3. Числовыми характеристиками непрерывных случайных величин являются ее ***математическое ожидание***

$$M(X) = \int\limits_{-\infty}^{\infty} x f(x)dx$$

и ***дисперсия***

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

при условии сходимости указанных несобственных интегралов.
Из определения $D(X)$ и свойства неотрицательности $f(x)$ вытекает оценка $D(X)\geq 0.$

Среднее квадратическое отклонение непрерывной случайной величины определяется в виде

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

3.4. Для вычисления дисперсии можно воспользоваться также формулой:

$$D(X)=\int_{-\infty}^{\infty}x^2 f(x)dx-\left(M(X)\right)^2,$$

доказательство которой проводится с помощью рассуждений, аналогичных п. 1.4, если воспользоваться свойством нормированности 3.2 г) плотности распределения.

В случае, если $f(x)$ принимает нулевые значения вне интервала $[a, b]$, соответствующие формулы принимают вид

$$M(X)=\int_a^b x f(x)dx \text{ и } D(X)=\int_a^b x^2 f(x)dx-\left(M(X)\right)^2.$$

4. Случайные величины: типовые задачи

Задача 4.1. Плотность распределения задана в виде

$$f(x)=\begin{cases}0, & x<0\\ c\cdot x, & 0\leq x<1\\ 0, & x\geq 1\end{cases},$$

где c – постоянная величина. Найти значение c, функцию распределения $F(x)$ и вероятность $P(\frac{1}{2}<X<\frac{3}{2})$.

Р е ш е н и е . По свойству нормированности

$$1=\int_{-\infty}^{+\infty}f(x)dx=\int_{-\infty}^{0}0\cdot dx+\int_0^1 c\cdot x\cdot dx+\int_0^{+\infty}0\cdot dx=$$

$$=0+\frac{1}{2}\cdot c\cdot x^2\Big|_0^1+0=\frac{1}{2}\cdot c,$$

откуда $c = 2$. Теперь

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

Найдем $F(x)$ по пункту б). Рассмотрим все возможные случаи: $x \le 0$, $0 < x < 1$, $x \ge 1$. Имеем

1) при $x \le 0$:

$$F(x) = \int_{-\infty}^{x} 0 \cdot dx = 0;$$

2)при $0 < x < 1$:

$$F(x) = \int_{-\infty}^{0} 0 \cdot dx + \int_{-\infty}^{x} 2x \cdot dx = x^2;$$

3) при $x \ge 1$:

$$F(x) = \int_{-\infty}^{0} 0 \cdot dx + \int_{-\infty}^{1} 2x \cdot dx + + \int_{1}^{x} 0 \cdot dx = 1.$$

Таким образом,

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

Так как $\frac{3}{2} \in [1, +\infty)$, $\frac{1}{2} \in [0, 1)$, то

$$P(\frac{1}{2} < X < \frac{3}{2}) = F(\frac{3}{2}) - F(\frac{1}{2}) = 1 - (\frac{1}{2})^2 = \frac{3}{4}.$$

Задача 4.2. Ряд распределения дискретной случайной величины X имеет вид

X	1	2	4
p	0,5	p_2	0,2

Определить p_2, $M(X)$ и $D(X)$.

Р е ш е н и е . Согласно свойству вероятностей (см. п. 5.2)

$$p_1 + p_2 + p_3 = 1,$$

имеем: $p_2 = 1 - 0{,}5 - 0{,}25$, то есть $p_2 = 0{,}25$.

Далее, согласно формулам п.3.3 для вычисления математического ожидания и дисперсии получаем в нашем случае:

$$M(X) = x_1 p_1 + x_2 p_2 + x_3 p_3,$$

$$D(X) = (x_1)^2 p_1 + (x_2)^2 p_2 + (x_3)^2 p_3 - (M(X))^2,$$

то есть

$$M(X) = 1 \cdot 0{,}5 + 2 \cdot 0{,}25 + 4 \cdot 0{,}25 = 2;$$

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

Задача 4.3. Непрерывная случайная величина X задана функцией распределения

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

Найти плотность распределения $f(x)$ и числовые характеристики $M(X)$ и $D(X)$.

Р е ш е н и е . По определению п.3.1 имеем:

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

Графики $y = F(x)$ и $y = f(x)$ изображены на рисунке:

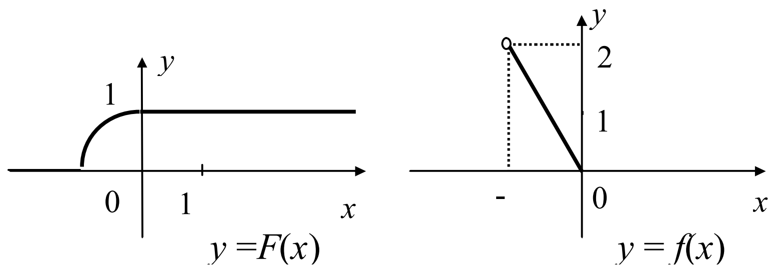

Рис .4.1. Графики функции и плотности распределения

Далее, согласно формулам п. 3.3 имеем:

$$M(X)=\int_{-1}^{0}x(-2x)dx=-\frac{2}{3}x^3\Big|_{-1}^{0}=-\frac{2}{3};$$

$$D(X)=\int_{-1}^{0}x^2(-2x)dx-\left(-\frac{2}{3}\right)^2=-\frac{2}{4}x^4\Big|_{-1}^{0}-\frac{4}{9}=\frac{1}{18}.$$

5. Специальные виды распределений

5.1. Пусть дискретная случайная величина X принимает значения, равные количеству появлений события A в n испытаниях, при условии, что в каждом вероятность $P(A)$ одна и та же, равная некоторому p; $q=1-p$ (схема Бернулли). Ее закон распределения называется биномиальным; соответствующий ряд распределения имеет вид

X	0	1	...	K	...	n
P	q^n	$C_n^1 p^1 q^{n-1}$	...	$C_n^k p^k q^{n-k}$		p^n

Название распределения объясняется тем, что сумма всех вероятностей представляет собой сумму всех членов разложения бинома Ньютона

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

Как оказывается, что ***математическое ожидание биномиальной случайной величины X равно произведению числа испытаний на вероятность появления события в одном испытании***

$$M(X)=np,$$

а её дисперсия равна произведению числа испытаний на вероятности появления и непоявления события в одном испытании

$$D(X)=npq.$$

Установим, например, первое из соотношений. По определению математического ожидания дискретной случайной величины и согласно виду ряда распределения биномиальной величины имеем ее математическое ожидание

$$M(X) = \sum_{k=0}^{n} kC_n^k p^k q^{n-k}.$$

Заметим, что суммирование фактически производится по k=1,..., n , так как слагаемое, соответствующее значению k=0, обращается в ноль.

Преобразуем полученную сумму следующим образом:

$$M(X) = \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} =$$

$$\sum_{k=1}^{n} k \frac{n(n-1)!}{k(k-1)!((n-1)-(k-1))!} p \cdot p^{k-1} q^{(n-1)-(k-1)} =$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!} p^{k-1} q^{(n-1)-(k-1)}$$

(общий множитель np вынесен за знак суммы). Перенумеровав члены (то есть заменив индекс суммирования $k-1$ на k), получаем

$$M(X) = np \sum_{k=1}^{n} C_{n-1}^{k-1} p^{k-1} q^{(n-1)-(k-1)} = np \sum_{k=1}^{n-1} C_{n-1}^{k} p^k q^{(n-1)-k} =$$

$$= np(p+q)^{n-1} = np \cdot 1 = np,$$

что и утверждалось

5.2. Дискретная случайная величина X распределена по закону Пуассона с параметром λ, если она принимает значения $\{0,1,2,...,m,...\}$ с вероятностями

$$p_m = P(X = m) = \frac{\lambda^m e^{-\lambda}}{m!} \quad (m = 0,1,2,...).$$

Закон Пуассона можно понимать как предельный случай (при $n \to \infty$, $\lambda = np = const$) биномиального закона. Так как вероятность появления события A в каждом испытании мала, то закон Пуассона называют законом редких явлений. По закону Пуассона распределены многие случайные величины, например, число сбоев на автоматической линии, число бракованных изделий в большой партии товара и т.д.

Проверим, что сумма вероятностей в ряде распределения Пуассона

X	0	1	2	…	m	…
P	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2}$	…	$\frac{\lambda^m e^{-\lambda}}{m!}$	…

равна единице, т.е. вычислим сумму ряда

$$\sum_{m=0}^{\infty} \frac{\lambda^m e^{-\lambda}}{m!} \quad .$$

Использовав разложение Маклорена экспоненциальной функции, получаем сумму этого ряда в виде

$$e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^m}{m!} = e^{-\lambda} e^{\lambda} = 1,$$

что и требовалось установить.

Оперируя с суммами степенных рядов и обращаясь все к тому же разложению Маклорена, можно доказать, что математическое ожидание и дисперсия случайной величины, распределенной по закону Пуассона, совпадают между собою и и равны параметру λ этого распределения:

$$M(X) = \lambda, \; D(X) = \lambda .$$

5.3. Непрерывная случайная величина X распределена по равномерному закону распределения на отрезке $[a, b]$, если ее плотность вероятности постоянна на этом отрезке. Если воспользоваться свойством нормированности плотности, то нетрудно проверить, что

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

Действительно,

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{a}^{b} \frac{1}{b-a} dx = \frac{1}{b-a} x \Big|_a^b = \frac{b-a}{b-a} = 1.$$

Найдем теперь функцию распределения $F(x)$ равномерно распределенной случайной величины. Для этого воспользуемся приведенной выше формулой

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

Имеем, очевидно, при $x \le a$ значения функция распределения $F(x)$=0. При $a < x \le b$ получим:

$$F(x) = \int_{-\infty}^{a} 0dt + \int_{a}^{x} \frac{1}{b-a} dt = \frac{x-a}{b-a}.$$

Наконец, при $x > b$

$$F(x) = \int_{-\infty}^{a} 0dt + \int_{a}^{b} \frac{1}{b-a} dt + \int_{a}^{x} 0dt = \frac{b-a}{b-a} = 1.$$

Итак,

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

Установим, что математическое ожидание и дисперсия равномерно распределенной на отрезке $[a, b]$ случайной величины вычисляются, соответственно, по формулам

$$M(X) = \frac{a+b}{2} \text{ и } D(X) = \frac{(b-a)^2}{12}.$$

Имеем

$$M(X) = \int_{-\infty}^{+\infty} xf(x)dx = \int_{a}^{b} \frac{x}{b-a} dx = \frac{1}{b-a} \left.\frac{x^2}{2}\right|_{a}^{b} = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

Второе соотношение (формула для дисперсии) доказывается аналогично:

$$D(X) = \int_{-\infty}^{+\infty} x^2 f(x)dx - [M(X)]^2 = \frac{a^2 + ab + b^2}{3} - \frac{(a+b)^2}{4} =$$

$$= \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}.$$

Равномерные распределения довольно часто встречаются на практике. Например, равномерно распределение на отрезке $[0; t_0]$ время ожидания транспорта, если предположить, что пассажир приходит на остановку в случайный момент времени, а транспорт ходит регулярно с интервалом t_0 мин. При компьютерном моделировании случайных явлений используется так называемый «генератор случайных чисел», который генерирует значения случайной величины, распределенной равномерно на отрезке $[0;1]$.

5.4. Непрерывная случайная величина X распределена по показательному закону с параметром λ>0, если ее плотность вероятности имеет вид:

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

Указанная $f(x)$ действительно может служить плотностью, т.к. она, очевидно, неотрицательна и обладает свойством нормированности :

$$\int\limits_{-\infty}^{+\infty} f(x)dx=\int\limits_{0}^{+\infty}\lambda e^{-\lambda x}dx=\lambda\lim_{A\to\infty}\int\limits_{0}^{A}e^{-\lambda x}dx=\lambda\lim_{A\to\infty}\left(-\frac{1}{\lambda}e^{-\lambda x}\Big|_0^A\right)=$$

$$=(\lim_{A\to\infty}e^{-\lambda A})-(-1)=0+1=1.$$

Найдем функцию распределения $F(x)$ случайной величины X. При $x\le 0$ имеем $F(x)=0$. При $x>0$ получим:

$$F(x)=\int\limits_{-\infty}^{0}0dt+\int\limits_{0}^{x}\lambda e^{-\lambda t}dt=0-\frac{\lambda}{\lambda}e^{-\lambda t}\Big|_0^x=-(e^{-\lambda x}-e^{0})=1-e^{-\lambda x}$$

Итак,

$$F(x)=\begin{cases}0\quad, & x\le 0,\\ 1-e^{-\lambda x}, & x>0.\end{cases}$$

Используя формулу интегрирования по частям, нетрудно проверить, что математическое ожидание и дисперсия показательно распределенной случайной величины имеют, соответственно, значения

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

6.Нормальное распределение

6.1. Непрерывная случайная величина X распределена по нормальному закону с параметрами a и σ>0, если ее плотность вероятности имеет вид:

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

Указанная $f(x)$ действительно может служить плотностью некоторого распределения, т.к. она, очевидно, неотрицательна и обладает свойством нормированности. При проверке этого свойства воспользуемся значением так называемого интеграла Эйлера-Пуассона:

$$\int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \sqrt{2\pi} \, .$$

С помощью замены переменной $z = \dfrac{x-a}{\sigma}$ получаем

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx =$$

$$= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sigma \, dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1,$$

что и утверждалось.

Можно установить, что соответствующая функция распределения $F(x)$ выражается через интегральную функцию Лапласа

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} \mathrm{e}^{-\frac{t^2}{2}} dt$$

следующим образом:

$$F(x) = 0.5 + \Phi(\frac{x-a}{\sigma}) \, .$$

График плотности нормального распределения (нормальная кривая, рис. 6.1) имеет следующий вид

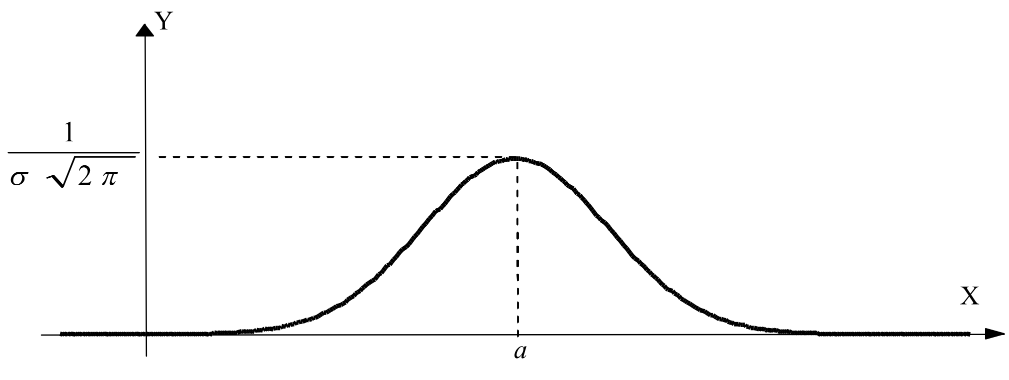

Рис. 6.1. Нормальная кривая

6.2. Вероятность попадания значений случайной величины *X*, распределенной по нормальному закону, в промежуток (α, β) может быть вычислена в виде

$$P(\alpha < X < \beta) = \Phi(\frac{\beta - a}{\sigma}) - \Phi(\frac{\alpha - a}{\sigma}).$$

Действительно, если воспользоваться видом функции распределения, установленным в п. 6.1, то получим

$$P(\alpha < X < \beta) = F(\beta) - F(\alpha) = 0.5 + \Phi(\frac{\beta - a}{\sigma}) - (0.5 + \Phi(\frac{\alpha - a}{\sigma})) =$$

$$= \Phi(\frac{\beta - a}{\sigma}) - \Phi(\frac{\alpha - a}{\sigma}),$$

что и утверждалось.

6.3. Из результата п. 6.2 вытекает следующее утверждение о вероятности малого отклонения значений нормальной величины от параметра a : для любого $\delta > 0$ имеет место равенство

$$P(| X - a |< \delta) = 2\Phi(\frac{\delta}{\sigma}).$$

Для доказательства запишем неравенство $| X - a |< \delta$ в равносильном виде $a - \delta < X < a + \delta$. Тогда, ввиду нечетности функции Лапласа, будем иметь

$$P(a - \delta < X < a + \delta) = \Phi(\frac{a + \delta - a}{\sigma}) - \Phi(\frac{a - \delta - a}{\sigma}) =$$

$$= \Phi(\frac{\delta}{\sigma}) - \Phi(\frac{-\delta}{\sigma}) = 2\Phi(\frac{\delta}{\sigma}),$$

что и требовалось доказать.

В частности при $\delta = 3\sigma$ получаем так называемое правило «трех сигм»

$$P(| X - a |< 3\sigma) = 2\Phi(\frac{3\sigma}{\sigma}) = 2\Phi(3) \approx 0{,}9973.$$

Смысл его состоит в следующем: практически достоверно, что абсолютная величина отклонения значений нормально распределенной X от параметра a меньше утроенного σ.

6.4. Вероятностный смысл параметров a и σ проясняется в следующем утверждении.

Нормально распределенная случайная величина имеет математическое ожидание

$$M(X) = a$$

и дисперсию

$$D(X)=\sigma^2.$$

Докажем, например, первое из соотношений. Используя снова замену переменной $z=\dfrac{x-a}{\sigma}$ и равенство нулю интеграла по симметричному промежутку от нечетной функции, получим

$$M(X)=\int_{-\infty}^{+\infty} x\cdot f(x)dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{+\infty} x\cdot e^{-\frac{(x-a)^2}{2\sigma^2}}dx=$$

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}(a+\sigma t)\cdot e^{\frac{-t^2}{2}}dt=$$

$$=\frac{a}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t\cdot e^{\frac{-t^2}{2}}dt+\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} t\cdot e^{\frac{-t^2}{2}}dt=a\cdot 1+0=a$$

что и утверждалось.

6.5. Нормальный закон распределения достаточно часто встречается на практике. Объясняется это тем, что суммарное поведение большого количества случайных величин, влияние каждой из которых на всю сумму ничтожно мало, есть практически поведение нормальной величины.

7. Статистическое распределение выборки

7.1. **Генеральная совокупность и выборка**. Пусть имеется множество, состоящее из конечного (но достаточно большого) числа некоторых объектов, причем каждый объект характеризуется единственным значением x_i некоторого количественного признака X. Такое множество называется ***генеральной совокупностью***; в свою очередь, совокупность из n случайно отобранных объектов называется ***выборкой объема*** n. Извлеченные выборки позволяют получить представление о распределении количественного признака X как некоторой случайной величины.

7.2. Пусть n_1 объектов из выборки характеризуются значением x_1, n_2 объектов – значением x_2, ..., n_k - объектов – значением x_k. Числа $x_1, x_2, ..., x_k$ называются ***вариантами***, а числа $n_1, n_2, ..., n_k$ – ***частотами*** соответствующих вариант. Соответствие между вариантами и соответствующими им частотами называется статистическим распределением выборки, а таблица вида

x_i	x_1	x_2	...	x_k
n_i	n_1	n_2	...	n_k

называется вариационным рядом; очевидно, что

$$n_1 + n_2 + ... + n_k = n.$$

Относительными частотами значений x_i называются соответствующие числа вида $w_i = \frac{n_i}{n}$; справедливо соотношение

$$w_1 + w_2 + ... + w_k = 1.$$

Модой M_o называют варианту, имеющую наибольшую частоту.

Медианой m_e называют варианту, которая делит находится в середине вариационного ряда, если число вариант нечетно; медиана m_e равна среднему арифметическому двух ближайших к «середине» вариант, если их число четно.

Размахом варьирования R называют разность между наибольшей и наименьшей вариантой.

7.3. ***Выборочной средней*** $\overline{x}_в$ называется среднее арифметическое всех наблюдаемых (в выборке) значений:

$$\overline{x}_в = \frac{(x_1 + ... + x_1) + (x_2 + ... + x_2) + \ldots + (x_k + ... + x_k)}{n},$$

причем в первой скобке имеется n_1 слагаемых, во второй $n_2, \ldots$, в последней - n_k слагаемых. Установим, что

$$\overline{x}_в = x_1 \frac{n_1}{n} + x_2 \frac{n_2}{n} + ... + x_k \frac{n_k}{n} \quad \text{или} \quad \overline{x}_в = \sum_{i=1}^{k} x_i w_i .$$

Действительно,

$$\overline{x_в} = \frac{x_1 n_1 + x_2 n_2 + ... + x_k n_k}{n} = \frac{x_1 n_1}{n} + \frac{x_2 n_2}{n} + ... + \frac{x_k n_k}{n} =$$

$$= x_1 w_1 + x_2 w_2 + \ldots + x_k w_k .$$

Возникает ***вопрос о количественной характеристике степени рассеивания наблюдаемых значений относительно их среднего***. Первое, что естественно сделать – это рассмотреть среднее арифметическое самих уклонений $x - \overline{x_в}$, то есть

$$\frac{((x_1 - \overline{x_в}) + ... + (x_1 - \overline{x_в})) + \ldots + ((x_k - \overline{x_в}) + ... + (x_k - \overline{x_в}))}{n}$$

или, что то же самое,

$$\frac{1}{n}\sum_{i=1}^{k}(x_i - \overline{x_в})w_i,$$

однако такая сумма равна нулю *для всякого* статистического распределения выборки (и следовательно, не может служить *характеристикой* рассеяния) :

$$\sum_{i=1}^{k}(x_i - \overline{x_в})w_i = \sum_{i=1}^{k}x_i w_i - \overline{x_в}\sum_{i=1}^{k}w_i = \overline{x_в} - \overline{x_в} = 0.$$

Очевидно, что нулевое значение вычисленной суммы получается за счёт взаимного погашения отрицательных и положительных уклонений. В этом случае естественно перейти к рассмотрению среднего арифметического *квадратов* уклонений значений x_i относительно их средней $\overline{x}_B$, называемого ***выборочной дисперсией*** D_B. Итак, выборочная дисперсия определяется в виде

$$D_B = \sum_{i=1}^{k}(x_i - x_B)^2 w_i,$$

которую удобнее вычислять по формуле (доказываемой способом, аналогичным приведенному в п. 1.4)

$$D_B = (x_1)^2\frac{n_1}{n} + (x_2)^2\frac{n_2}{n} + \ldots + (x_k)^2\frac{n_k}{n} - (\overline{x}_B)^2.$$

Выборочным средним квадратическим отклонением называется величина

$$\sigma(X) = \sqrt{D_B}.$$

7.4. Поскольку с ростом *n* относительные частоты

$$w_i = w(X = x_i)$$

становятся близкими к соответствующим вероятностям

$$p_i = P(X = x_i),$$

то значения выборочных средних $\overline{x}_B$ становятся близкими к математическому ожиданию $M = M(X)$ количественного признака X. Точно также, значения выборочных дисперсий D_B становятся (с ростом *n*) близкими к дисперсии количественного признака $\sigma^2 = D(X)$. Точные формы этих утверждений используют понятие сходимости по вероятности и выходят за рамки настоящего материала.

8. **Статистические оценки параметров распределения**

8.1. Предположим, что нас интересует неизвестное значение θ некоторого параметра, характеризующего количественный признак X генеральной совокупности. Оценкой параметра θ называют значения θ^* некоторой функции от наблюдаемых (в выборках) значений, дающие первоначальное представление о величине θ.

Точечной оценкой параметра θ называют оценку, которая определяется одним числом; так, например, точечной оценкой математического ожидания $\mu = M(X)$ количественного признака X генеральной совокупности есть значение выборочной средней.

8.2 . Пусть известен вид функции распределения $F(x,\theta)$ количественного признака X генеральной совокупности, но единственный параметр распределения θ остается неизвестным. Для его оценки достаточно иметь одно уравнение относительно этого параметра. Поскольку математическое ожидание μ случайной величины X также зависит от этого параметра, т.е. $\mu = M(X,\theta)$, то имеем (см. п. 8.1) следующее приближенное равенство для нахождения θ:

$$\bar{x}_B \approx M(X,\theta)$$

Если функция распределения определяется двумя неизвестными параметрами, то есть имеет вид $F(x,\theta_1,\theta_2)$, то для их оценки необходимы два уравнения. В этом случае следует пользоваться системой двух приближенных равенств

$$\begin{cases} \bar{x}_B \approx M(X,\theta_1,\theta_2) \\ D_B \approx D(X,\theta_1,\theta_2) \end{cases}.$$

Указанный способ нахождения неизвестных параметров распределения называется методом моментов.

Найдем, например, методом моментов оценку параметров a и b равномерного распределения равномерно распределенного количественного признака генеральной совокупности. Имеем систему

$$\begin{cases} \bar{x}_B \approx M(X,a,b) \\ D_B \approx D(X,a,b) \end{cases},$$

или, учитывая, что для равномерного распределения

$$M(X)=\frac{a+b}{2} \text{ и } D(X)=\frac{(b-a)^2}{12},$$

имеем

$$\begin{cases} \frac{1}{2}(a+b)\approx \overline{x}_B \\ \frac{1}{12}(b-a)^2 \approx D_B \end{cases}.$$

Решив эту систему, получим искомые оценки

$$a^* = \overline{x}_B - \sigma_B\sqrt{3},\ b^* = \overline{x}_B + \sigma_B\sqrt{3}, \text{ где } \sigma_B = \sqrt{D_B}.$$

8.3. ***Интервальной называют оценку***, которая определяется двумя числами –концами интервала. Надежностью (доверительной вероятностью) оценки θ^* называют вероятность γ, с которой отклонение $|\theta-\theta^*|$ оказывается заданно малым.

Наиболее часто задают надежность, равную 0,95; 0,99; 0,995. Интервал $(\theta^* - \delta, \theta^* + \delta)$ называют доверительным интервалом для оценки параметра θ

Рассмотрим так называемые ***доверительные интервалы для оценки математического ожидания*** нормального распределения при известном среднем квадратическом отклонении σ . Пусть количественный признак X генеральной совокупности распределен с плотностью

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

(нормальное распределение); здесь параметр a есть значение математического ожидания, а σ – среднее квадратическое отклонение случайной величины X. Предположим, что среднее квадратическое отклонение σ этого распределения известно, и извлечена выборка объема n. Требуется оценить неизвестное математическое ожидание a по выборочной средней с заданной надежностью γ. Если

$$P(|\overline{x}_B - a| < \delta)=\gamma,$$

то значение δ (радиус доверительного интервала) определяется в виде $\delta = t\frac{\sigma}{\sqrt{n}}$, где число t может быть найдено из соотношения $\Phi(t) = \frac{\gamma}{2}$, по табл. 1 Приложений. Следовательно, с надежностью γ доверительный интервал

$$\left(\overline{x}_B - t\frac{\sigma}{\sqrt{n}}\ ,\ \overline{x}_B + t\frac{\sigma}{\sqrt{n}}\right)$$

покрывает неизвестный параметр a; точность оценки есть $\delta = t\frac{\sigma}{\sqrt{n}}$.

9. Математическая статистика: типовые задачи

Задача 9.1. Из генеральной совокупности извлечена выборка. Известен вариационный ряд

x_i	1,1	1,2	1,3	1,4	1,5
n_i	10	25	40	15	10

Указать размах варьирования, моду, медиану вариационного ряда. Найти: а) выборочную среднюю; б) выборочную дисперсию.

Р е ш е н и е . Размах

$$R = x_{\max} - x_{\min} = 1{,}5 - 1{,}1 = 0{,}4.$$

Мода Mо=1,3, т.к. эта варианта обладает наибольшей частотой, равной 40. Медиана mе=1,3, поскольку эта варианта расположена в середине вариационного ряда.

Объем выборки

$$n = 10 + 25 + 40 + 15 + 10 = 100.$$

а) По формуле п.8.3 имеем

$$\overline{x}_B = 1{,}1\cdot\frac{10}{100} + 1{,}2\cdot\frac{25}{100} + 1{,}3\cdot\frac{40}{100} + 1{,}4\cdot\frac{15}{100} + 1{,}5\cdot\frac{10}{100} = 1{,}29.$$

б) Согласно п. 8.3

$$D_B = (1{,}1)^2\cdot 0{,}1 + (1{,}2)^2\cdot 0{,}25 + (1{,}3)^2\cdot 0{,}4 + (1{,}4)^2\cdot 0{,}15 +$$

$$+ (1{,}5)^2\cdot 0{,}1 - (1{,}29)^2 = 0{,}0319.$$

Задача 9.2. Случайная величина X имеет нормальное распределение с известным средним квадратичным отклонением σ = 4. Найти доверительный интервал для оценки неизвестного математического ожидания a по выборочной средней $\overline{x}_в$ =3,6, если объем выборки n =64 и задана надежность оценки $\gamma = 0{,}95$.

Найдем t из соотношения $\Phi(t) = \frac{0{,}95}{2} = 0{,}475$. По табл. 1 Прил. находим t = 1,96. Найдем точность оценки $\delta = \frac{1{,}96 \cdot 4}{\sqrt{64}} = 0{,}98$. Следовательно, доверительный интервал имеет вид (3,6 – 0,98; 3,6 + 0,98). Иначе говоря, с надежностью $\gamma = 0{,}95$ имеет место неравенство $2{,}62 < a < 4{,}58$.

10. Простейшие стохастические модели: кейс-задания

Задание 10.1. Передаётся 11 смс-сообщений, в каждом из которых с вероятностью p=0,2 может быть утрачена часть текста. Найти вероятность того, что:

1) из 11 сообщений ровно в двух будет утрачена часть текста;

2) все сообщения будут приняты полностью;

3) утрата части текста произойдет не менее, чем в двух сообщениях.

4) Каково среднее количество сообщений с утратой части текста?

5) Если при том же среднем количестве сообщений с утратой части текста передается 1000 сообщений, то какова будет вероятность утраты части текста в каждом сообщении? В 3-х сообщениях?

Решение. В задаче представлена *содержательная модель*. Перейдём к модели математической. Поскольку проводится n=11 опытов, в каждом из которых событие, состоящее в утрате части текста, имеет одну и ту же вероятность p=0,2, то имеем схему Бернулли. Теперь *анализируем математическую модель.*

1) В случае утраты части текста в $k = 2$ сообщениях, получаем при q=1–p= =0,8 по формуле Бернулли

$$p_{11}(2) = C_{11}^2 \cdot 0{,}2^2 \cdot 0{,}8^{11-2} = \frac{11!}{2!\cdot(11-2)!} \cdot \left(\frac{1}{5}\right)^2 \cdot \left(\frac{4}{5}\right)^9 = 0{,}275.$$

2) Вероятность принятия всех 11 полнотекстовых сообщений соответствует случаю $n = 11,\ k = 0$:

$$p_{11}(0) = 0{,}8^{11} = \left(\frac{4}{5}\right)^{11} = 0{,}084.$$

3) Событие искажения не менее двух сообщений противоположно тому, что утрата части текста произойдёт менее чем в двух сообщениях, то есть либо в одном, либо ни в одном сообщении:

$$p_{11}(k \geq 2) = 1 - p_{11}(0) - p_{11}(1)$$

или

$$p_{11}(k \geq 2) = 1 - 0{,}84 - \frac{11!}{1!\cdot(11-1)!}\cdot\left(\frac{1}{5}\right)^{1}\cdot\left(\frac{4}{5}\right)^{11-1} = 0{,}685$$

4) Среднее количество X сообщений с утратой части текста есть математическое ожидание биномиального распределения с числом опытов $n = 11$ и вероятностью $p = 0{,}2$

$$M(X) = np = 11\cdot 0{,}2 = 2{,}2 \approx 2.$$

5) Если передается 1000 сообщений, то можно пользоваться математическим ожиданием распределения Пуассона

$M(X) = np = 2{,}2,$ откуда $1000p = 2{,}2$ и искомая вероятность $p = 0{,}0022.$

При этом вероятность утраты части текста в 3-х сообщениях по формуле Пуассона ищется в виде

$$P_{1000}(3) \approx \frac{2{,}2^{3}e^{-2{,}2}}{3!} \approx 0{,}177.$$

Интерпретация модели: в условиях данного опыта ровно два сообщения будут приняты с утратой части текста в 27,5% случаев; получение всех сообщений с полным текстом возможно в 8% случаев, а наличие от 2-х до 11 искажений (утрат) текста возможно в 68,5% случаев.

Среднее количество X сообщений с утратой части текста равно 2.

Если же передается 1000 сообщений, то вероятность утраты части текста в каждом сообщении $p = 0{,}0022.$ При этом ровно в 3-х сообщениях утрата части происходит с вероятностью $\approx 0{,}177.$

Задача 10.2. Случайные отклонения диаметра шайб (используемых в конструкции данной машины) от номинала распределены нормально. Математическое ожидание диаметра 200 мм, среднее квадратическое отклонение равно 0,25 мм. Стандартными считаются шайбы, размер которых заключен между 199,5 мм и 200,5 мм.

1.Найти процент стандартных шайб.

2. Какова вероятность, что 5 случайно отобранных шайб будет стандартны?

3. Какова вероятность, что среди 5 случайным образом отобранных шайб хотя бы одна будет нестандартной?

В условии задачи уже представлена содержательная модель. Математическая же модель будет представлена соотношениями, позволяющими найти соответствующие вероятности. Если X – нормально распределённая случайная величина, представляющая диаметры, то, по условию, математическое ожидание и среднее квадратическое отклонение величины X, соответственно, будет равны:

$$a = M(X) = 200, \quad \sigma(X) = 0{,}25.$$

В п. 6.3 установлено, что вероятность малого отклонения значений нормальной величины от параметра a может быть вычислена в виде

$$P(|X - a| < \delta) = 2\Phi(\frac{\delta}{\sigma}),$$

где $\Phi(x)$ – интегральная функция Лапласа. В нашем случае, $\delta = 200{,}5 - 200 = |199{,}5 - 200| = 0{,}5$, а тогда $\frac{\delta}{\sigma}$=2. Значит,

$P(199{,}5 < X < 200{,}5) = 2\Phi(2) \approx 0{,}95$.

2. Согласно схеме Бернулли, в 5 опытах событие «все 5 шайб стандартны» имеет вероятность

$P_5(5) = 0{,}95^5 \approx 0{,}77$.

3. Наконец, нестандартность хотя бы одной шайбы есть событие, противоположное стандарности всех 5 шайб. Следовательно,

$P_5(X \geq 1) = 1 - P_5(5) \approx 1 - 0{,}77 = 0{,}23.$

Интерпретация модели. Результат решения показывает, что

1) процент стандартных шайб составляет 95%;
2) пять случайно отобранных шайб будут стандартны с вероятностью $\approx 0{,}77$;
3) хотя бы одна из них нестандартна с вероятностью $\approx 0{,}23$.

Заключение. Вероятностно-статистическая теория: история и современность

Возникновение теория вероятностей относят к середине 17 века и связывают с работами французских учёных Б. Паскаля и П. Ферма и голландского учёного Х. Гюйгенса. Стимулом явилась необходимость подсчета шанса выигрыша в той или иной азартной игре. Важный этап развития теории вероятностей связан с именем швейцарского математика Я. Бернулли, математически обосновавшего свойство устойчивости относительной частоты в виде закона больших чисел для схемы независимых испытаний с двумя исходами (1713 г.). Следующий период (18 в. и начало 19 в.) связан с именами А. Муавра (Англия), П. Лапласа (Франция), К. Гаусса (Германия) и С. Пуассона (Франция). Это - период, когда теория вероятностей уже находит ряд весьма актуальных применений в естествознании и технике (главным образом в теории ошибок наблюдений и измерений, развившейся в связи с потребностями геодезии, астрономии, теории стрельбы и др.). К этому периоду относится доказательство первых предельных теорем в схеме Бернулли, носящих теперь названия теорем Лапласа (1812) и Пуассона (1837); А. Лежандром (Франция, 1806) и Гауссом (1808) в это же время был разработан способ наименьших квадратов получения точечных оценок параметров уравнений регрессии (уравнений, описывающих экспериментальным образом выявленные зависимости функционального характера).

Третий период развития теории вероятностей (2-я половина 19 в.) связан, в основном, с выдающимися представителями русской математической школы П. Л. Чебышева, А. М. Ляпунова и А. А. Маркова. Чебышев и его ученики Ляпунов и Марков поставили и решили ряд общих задач, называемых законом больших чисел и центральной предельной теоремой; один важный случай зависимых испытаний исследованный Марковым, впоследствии получил название цепей Маркова.

Следующий этап развития теории вероятностей характеризуется чрезвычайным расширением круга её применений, созданием системы безукоризненно строгого математического обоснования, новых мощных методов, требующих иногда применения (помимо классического анализа) средств теории множеств, теории функций действительного переменного и функционального анализа. В этот период при очень большом усилении работы по теории вероятностей за рубежом (во Франции - Э. Борель, П. Леви, М. Фреше, в Германии - Р. Мизес, в США - Н. Винер, В. Феллер, Дж. Дуб, в Швеции - Г. Крамер) российская наука продолжает занимать значительное, а в ряде направлений и ведущее положение. В нашей стране новый период развития теории вероятностей открывается деятельностью С. Н. Бернштейна, значительно обобщившего классические предельные теоремы Чебышева, Ляпунова и Маркова и впервые в России широко поставившего работу по применениям теории вероятностей к естествознанию.

А. Я. Хинчин и А. Н. Колмогоров разработали применение методов теории функций действительного переменного к решению вероятностных проблем. Позднее они и их последователи заложили основы теории случайных процессов; вслед за этим вероятностные методы получают широкое применение в статистике.

Отдельно следует остановиться на персоне А.Н. Колмогорова, который внес величайший вклад в становление и развитие аппарата теории вероятностей. Его фундаментальные труды с успехом пользуются по сей день. Аксиоматика Колмогорова поставила вероятностно-статистическую теорию в один ряд с другими классическими направлениями математической науки.

На базе аппарата теории вероятностей появились , наряду с математической статистикой, теория случайных процессов, теория массового обслуживания и др. Так, в ряде естественнонаучных исследований последних десятилетий усилилась потребность, наряду с одномерными и многомерными случайными величинами, рассматривать случайные процессы, то есть процессы, для которых определена вероятность того или иного их течения. Примером случайного процесса может служить координата частицы, совершающей броуновское движение. В теории вероятностей случайный процесс рассматривают обычно как однопараметрическое семейство случайных величин $X\ (t)$; например, параметр t может является временем, протекшим с начала процесса.

Подобно тому, как случайная величина характеризуется законом распределения, случайный процесс может быть охарактеризован совокупностью совместных законов распределения для $X\ (t_1)$, $X\ (t_2)$,..., $X\ (t_n)$ для всевозможных моментов времени t_1, t_2,..., t_n при любом $n > 0$. В настоящее время наиболее интересные конкретные результаты теории случайных процессов получены в двух специальных направлениях.

Исторически первыми изучались марковские процессы(процессы без последействия). Марковские процессы являются естественным обобщением детерминированных процессов, рассматриваемых в классической физике, когда состояние системы в момент времени t_0 однозначно определяет ход процесса в будущем; в марковских процессах состояние системы в момент времени t_0 однозначно определяет распределение вероятностей хода процесса при $t > t_0$, причём никакие сведения о ходе процесса до момента времени t_0 не изменяют это распределение.

Другим направлением теории случайных процессов является теория стационарных случайных процессов. Стационарность процесса, то есть неизменность во времени его вероятностных закономерностей, налагает сильное ограничение на процесс и позволяет из одного этого допущения извлечь ряд важных следствий (например, возможность так называемого спектрального разложения). В то же время схема стационарных процессов с хорошим приближением описывает многие физические явления.

Современное состояние теории вероятностей и математической статистики характеризуется широким применением методов теории функций и функционального анализа; с другой стороны, вероятностно-статистические проблемы стимулируют развитие других ветвей математической науки.

ПРИЛОЖЕНИЯ

Таблица 1. Функции Лапласа

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{z^2}{2}} dz$$

X	0	0,5	1,0	1,5	1,645	1,96
φ(x)	0,3989	0,3521	0,2420	0,1295	0,103	0,0584
Φ(x)	0,0000	0,1915	0,3413	0,4332	0,45	0,4750
X	2,0	2,50	2,80	3,0	3,5	4
φ(x)	0,0540	0,0175	0,0078	0,0044	0,0009	0,0001
Φ(x)	0,4772	0,4938	0,4985	0,4987	0,4997	0.4999

Таблица 2. Значения экспоненциальной функции

X	1	2	3	4	5
e^{-x}	0,36788	0,13534	0,04979	0,01832	0,00674

КОНТРОЛЬНЫЙ МОДУЛЬ

КОМБИНАТОРИКА

1. ТЕОРЕТИЧЕСКИЕ УПРАЖНЕНИЯ

1. Доказать равенство $C_m^k = C_m^{m-k}$.

2. Доказать равенство $P_m = A_m^k P_{m-k}$.

3. Пользуясь результатом $n(A \cup B) = n(A) + n(B) - n(A \cap B)$, вывести формулу для подсчета количества элементов $n(A \cup B \cup C)$ в объединении трех множеств.

4. Равны ли тождественно суммы $\sum\limits_{k=0}^{m} C_m^k x^{m-k}$ и $\sum\limits_{k=0}^{m} C_m^k x^k$?

2. ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1.Логин должен начинаться с английской буквы S и состоять из четырех букв (в английском алфавите 26 букв). Сколько можно образовать таких логинов, если

а)все буквы в нем должны быть различными;

б)буквы могут повторяться.

2. Каждый из учащихся класса во время каникул побывал в походе или на экскурсии. В походе были 75% учащихся класса, а на экскурсии - 60% класса, причем некоторые учащиеся успели поучаствовать в обоих мероприятиях. Каков процент таких учащихся?

3. Сколько существует различных автомобильных номеров, которые состоят из четырех цифр, если первая из них не равна нулю?

4. Из восьми депутатов надо выбрать председателя счетной комиссии и его заместителя. Сколькими способами это можно сделать?

5. Сколько существует различных перестановок букв в слове «дорога» ?

6. Сколько существует различных диагоналей в выпуклом n-угольнике?.

7.Решить уравнение $5C_n^3 = C_{n+2}^4$.

8.Найти член разложения бинома $(3x - \frac{1}{x^2})^6$, не зависящий от x.

3. ИТОГОВЫЙ ТЕСТ

1.В отделении связи продают 10 видов конвертов и 5 видов марок. Сколькими способами можно купить конверт и марку к нему?

2.Сколько существует различных автомобильных номеров, которые состоят из пяти цифр, если номер состоит из одной буквы английского алфавита, за которой следуют любые четыре цифры?

3.Шесть меломанов спешат занять очередь к филармонической кассе. Сколькими способами может быть сформирована такая очередь?

4. В субботу в классе должно быть 4 урока. Сколько можно составить субботних расписаний, если всего имеется 10 учебных дисциплин.

5. Класс из 17 человек должен быть разделен на две подгруппы для изучения английского и французского языка, причем в «английской» группе должно быть 10 человек. Сколько существует способов формирования подгрупп, если сами учащиеся не высказывают никаких предпочтений по поводу выбора иностранного языка?

6. На прилавке цветочного магазина – восемь гвоздик, восемь роз и три тюльпана. Сколько существует способов сформировать букет из трех гвоздик, четырех роз и двух тюльпанов?

7. Решить уравнение $A_n^2 C_n^{n-1} = 48$.

8. Решить уравнение $7A_{n+1}^{n-1} + 14P_{n-1} = 30P_n$

9. Сумма всех коэффициентов разложения $(x+1)^n$ равна 64. Найти коэффициент этого разложения, записанный перед x^3.

10**.** Найти средний член разложения $(\sqrt[5]{5} + \sqrt[10]{4})^{10}$.

ОТВЕТЫ

Задание 1. 50

Задание 2. 260000

Задание 3. $6! = 720$

Задание 4. $A_{10}^4 = 5040$

Задание 5. $C_{17}^{10} \times C_7^7 = 19448$

Задание 6. 11760

Задание 7. 4

Задание 8. 7

Задание 9. 20

Задание 10. 2520

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

СЛУЧАЙНЫЕ СОБЫТИЯ

1. ТЕОРЕТИЧЕСКИЕ УПРАЖНЕНИЯ

1.Привести примеры:

а) полной группы событий;

б) двух совместных и независимых событий.

2. Если классическая вероятность некоторого события равна 1, то верно ли, что это событие –достоверное ? Почему?

3. Задача о выборке. Среди N предметов имеется M меченых. Случайным образом извлекается k предметов. A событие, состоящее в том, что в этой выборке содержится ровно l меченых предметов. Доказать, что вероятность этого события может быть найдена по формуле

$$P(A) = \frac{C_M^l \cdot C_{N-M}^{k-l}}{C_N^k}.$$

4. Доказать, что для любых двух событий A и B имеет место неравенство

$$P(A+B) \le P(A) + P(B).$$

5. Доказать, что если $p = P(A)$ и $q = P(\overline{A})$, то $pq \le \frac{1}{4}$.

6. Доказать, что если $P(B)$ не зависит от наступления (ненаступления) события A, то и $P(A)$ не зависит от B (т.е. свойство независимости – взаимное), то

Указание: воспользоваться вероятностью произведения $P(AB)$, представив ее (на основании свойства коммутативности умножения) двумя способами.

7. Доказать, что если A и B – независимые события, то и каждая пара событий A и $\overline{B}$, $\overline{A}$ и B, $\overline{A}$ и $\overline{B}$ является парой независимых событий.

8. Доказать соотношение

$$P(A_1 \cdot A_2 \cdot \ldots \cdot A_n) = P(A_1) \cdot P_{A_1}(A_2) \cdot \ldots \cdot P_{A_1 \cdot A_2 \cdot \ldots \cdot A_{n-1}}(A_n).$$

Указание: воспользоваться методом математической индукции.

9. Можно ли в случае *двух* гипотез формулу полной вероятности записать в виде

$$P(A) = P(H_1)P_{H_1}(A) + P(\overline{H}_1)P_{\overline{H}_1}(A)?$$

10. Доказать, что если вероятность p появления события A в каждом из n опытов постоянна, и $q = 1 - p$, то вероятность ρ наступления события A хотя бы один раз может быть вычислена в виде

$$\rho = 1 - q^n.$$

11. Вероятность p появления события A в каждом опыте постоянна. Опыты проводят до первого наступления события. Чему равна вероятность того, что будет проведено ровно n опытов.

12. Вероятность p появления события A в каждом из n опытов постоянна. При каких значениях p вероятность наступления события ровно n раз больше, чем вероятность непоявления события ровно n раз?

Указание: сравнить результаты применения формулы Бернулли в обоих случаях.

13. Доказать, что если вероятность p появления события A в каждом из n опытов мала (n – велико) и $\lambda = np$, то вероятность появления события A хотя бы один раз может быть найдена по приближенной формуле

$$P(k \geq 1) \approx 1 - e^{-\lambda}.$$

14. Доказать нечетность интегральной функции Лапласа $\Phi(-x) = -\Phi(x)$.

2. ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Образуют ли полную группу следующие события: A – два попадания в мишень при двух выстрелах, B – ни одного попадания при тех же двух выстрелах ?

2. Рассматриваются следующие события: A – первое из полученных электронных писем содержит навязчивую рекламу (СПАМ), B – второе письмо содержит СПАМ. Выразить с помощью операций сложения и умножения через события A и B и (или) им противоположные следующие события а) событие С – ни одно из писем не содержит СПАМ;

б) хотя бы одно письмо содержит СПАМ;

в) только одно письмо содержит СПАМ.

3. Вероятность дождливой погоды в предстоящий выходной день равна 0,7. Вероятность удачной рыбалки в дождливую погоду равна 0,8, а в ясную погоду – 0,4. Какова вероятность, что в предстоящий выходной рыбалка будет удачной.

4. Заявки работодателей на специалистов инженерных, экономических и юридические направлений поступают на биржу в отношении 6:3:1. Вероятность того, что претендент на вакансию инженера удовлетворит требованиям работодателя равна 0,8, на вакансию экономиста – 0,8, на вакансию юриста – 0,5. Найти вероятность, что

а) случайно выбранный на бирже претендент устроится на работу по своей специальности;

б)вероятность того, что устроившийся на работу специалист – экономист.

5. Имеется 10 двадцатидолларовых купюр, из которых 4 купюры фальшивые. Наугад *поочередно* извлекают две купюры и отыскивают вероятность события A, состоящего в том, что обе эти купюры окажутся фальшивыми. Можно ли применять формулу Бернулли, если а) купюра после извлечения и проверки возвращается в пачку; б)выборка безвозвратная.

Найти $P(A)$ в каждом из случаев а) и б).

6. Вероятность продать по оптимальной цене каждый из пяти пакетов акций в период их падения равна 0,25. Какова вероятность продажи по оптимальной цене большей части пакета?

7. В прямоугольник вписаны две окружности равного радиуса, касающиеся друг друга внешним образом. В прямоугольник случайным образом брошена точка. Какова вероятность, что она не попадет ни в один из кругов?
8. В семье 5 детей; вероятность рождения мальчика в данной местности равна 0,6. Найти вероятности следующих событий:
а) в семье две девочки;
б) в семье не менее двух девочек;
в) в семье мальчиков больше, чем девочек
9**.** В данной местности левши составляют 5% населения. Какова вероятность, что в на факультете, где обучаются 400 человек, окажутся не менее 3 левшей?
10. Студент одинаково плохо подготовился к каждому из трех экзаменов. С какой вероятностью он сдает каждый экзамен, если хотя бы один из них он сдаст с вероятностью 0,578125

3. ИТОГОВЫЙ ТЕСТ

Ответы к заданиям должны быть записаны в виде целого числа или десятичной дроби

1. По прогнозам экономистов, ежегодная инфляция не превысит заданного процента с вероятностью $p = 0{,}8$. Какова вероятность того, что в течение всех трех ближайших лет оправдается указанный экономический прогноз?

3. В комплекте из десяти дискет ровно две заражены вирусом. Какова вероятность того, что обе наугад взятые дискеты окажутся без вируса?

Ответ записать в виде десятичной дроби приближенно с точностью до 0,01.

3. В течение часа на сайт интернет-магазина заходит в среднем 5 человек. Вероятность того, что будет сделан заказ на товар для каждого из посетителей равна $\frac{1}{3}$. Какова вероятность того, что в течение часа не менее трех из пяти посетителей сделают заказ?

Ответ записать в виде десятичной дроби приближенно с точностью до 0,01.

4. Каждый из трех независимо работающих сигнализаторов своевременно сообщает о нарушении заданного режима работы реактора с вероятностью, соответственно, $p_1 = 0,9$; $p_2 = 0,8$; $p_3 = 0,75$. Какова вероятность того, что при нарушении заданного режима работы сигнала не поступит?

5. Среди пяти одинаковых по внешнему виду саженцев три – элитных. Наугад взяты два саженца. Какова вероятность, что ровно один из них элитный?

6. По данным социологов, в городе А данный кандидат в депутаты будет поддержан на выборах большей частью населения с вероятностью $p_1 = 0,6$; в городе В - с вероятностью $p_2 = 0,7$. Какова вероятность, что на выборах кандидат одержит победу хотя бы в одном из городов А и В?

7. Партия из пяти изделий забраковывается, если хотя бы одно из них окажется нестандартным. Каждое из производимых изделий удовлетворяет требованиям стандарта с вероятностью $p = \frac{4}{5}$. Какова вероятность, что партия будет забракована?

8. Вероятности своевременной доставки денежного перевода в города А, Б, В равны соответственно $p_1 = 0,7$, $p_2 = 0,9$, $p_3 = 0,6$. В каждый из этих городов отправлены по одному переводу. Какова вероятность того, что ровно в два города перевод будет доставлен своевременно?

9. Имеется 3 урны. В первой 2 белых и 1 черный шар. Во второй 1 белый и 1 черный шар. В третьей 3 белых и 2 черных шара. Случайным образом выбираются одна из урн и из нее извлекается 1 шар. Какова вероятность, что извлеченный шар – черный?

Ответ записать в виде десятичной дроби приближенно с точностью до 0,01.

10. Вероятность того, что каждый данный клиент Сбербанка потребует в течение календарного года закрытия своего лицевого счета, равна 0,002. Какова вероятность, что среди 1000 клиентов отделения сбербанка трое потребуют в течение года закрытия лицевых счетов?

Ответ записать в виде десятичной дроби приближенно с точностью до 0,01.

11. Сколько следует разместить люменисцентных ламп на потолке офиса, чтобы офис был освещен (хотя бы одной лампой) с вероятностью 0,99968, если вероятность перегорания каждой лампы равна 0,2?

12. На пути автомобиля 6 светофоров, каждый из которых может задержать его с вероятностью p=0,8. Какова вероятность, что в 4-х случаях автомобиль попадает под красный свет?

ОТВЕТЫ

Задание №1 *0,512*

Задание № 2 *0,62*

Задание № 3 *0,28*

Задание № 4 *0,005*

Задание № 5 *0,6*

Задание № 6 *0,88*

Задание № 7 *0,5904*

Задание № 8 *0,456*

Задание № 9 *0,41*

Задание № 10 *0,17*

Задание № 11 5

Задание № 12 0,246

СЛУЧАЙНЫЕ ВЕЛИЧИНЫ. ЭЛЕМЕНТЫ СТАТИСТИКИ

1. ТЕОРЕТИЧЕСКИЕ УПРАЖНЕНИЯ

1. Индикатор η случайного события А определяется как случайная величина со значениями р$\eta(A)=0$, если А не наступает и $\eta(A)=1$, если А наступает. Найти математическое ожидание и дисперсию индикатора.

2. Постоянная величина C понимается как величина, принимающая единственное значение c с вероятностью p=1. Доказать, что $M(C)=c$ и $D(C)=0$.

3. Возможно ли, чтобы плотность распределения была равна 1 на промежутке (-0,5, 0,7)?

4. Может ли плотность распределения принять значение, равное 1,1?

5. Может ли функция $F(x) = 1 - e^{-|x|}$ быть функцией какого- либо распределения? Может ли вообще четная функция быть функцией какого-либо распределения?

6. Если плотность распределения $f(x)$ – четна , то чему равен несобственный интеграл

$$\int_{-\infty}^{0} f(x)dx \quad ?$$

7. Может ли плотность распределения быть нечетной?

8. Пусть x_k, $k = 1,2,...,l$ – значения всех вариант, представленных в выборке и ω_k – их относительные частоты. Чему равна сумма

$$\sum_{k=1}^{l} \omega_k$$

Результат обосновать.

9. Доказать, что выборочная средняя $x_в$ принимает значения

$$x_в \in [x_{\min}, x_{\max}],$$

где $x_{\max}$ и $x_{\min}$ - соответственно наибольшее и наименьшее значения вариант x_k.

10. Количественный признак распределен по показательному закону с параметром λ ,т.е. плотность распределения имеет вид

$f(x) = \lambda e^{-\lambda x}, x > 0; \qquad f(x) = 0, \ x < 0.$

Извлечена выборка и вычислена выборочная средняя $x_в$. Используя метод моментов, найти точечную оценку параметра λ.

2. ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Устройство состоит из трех независимо работающих элементов. Вероятность отказа каждого элемента при включении равна 0,2. Составить ряд распределения числа элементов, отказавших при включении. Найти вероятность того, что откажет не более одного элемента.

2. Три стрелка стреляют по одной мишени. Вероятность попадания в мишень при одном выстреле для первого стрелка равна 0,5, для второго и для третьего – по 0,7. Пусть X - число попаданий в мишень при одном залпе. Составить ряд распределения X, найти функцию распределения $F(x)$ и построить ее график.

3. В пачке из 10 театральных билетов три билета – на премьеру. Наудачу взяты 3 билета. Составить ряд распределения случайной величины X - числа билетов на премьеру среди отобранных. Найти вероятность того, что хотя бы один из взятых билетов окажется на премьеру.

4. Вероятность того, что стрелок попадет в мишень при одном выстреле, равна 0,6. Стрелку выдаются патроны до тех пор, пока он не попадет в первый раз в мишень. Составить ряд распределения случайной величины X - числа патронов, выданных стрелку.

5. Испытывается 3 блока компьютера, причем вероятность отказа каждого не зависит от отказов остальных и составляет 0,1. Пусть X - число отказавших за время испытаний блоков. Составить ряд распределения X, найти ее математическое ожидание и дисперсию и вычислить вероятности событий: а) X = 0; б) X < 3.

6. Случайная величина X задана на всей числовой оси функцией распределения $F(x) = \dfrac{2}{\pi} arctg\, x$. Найти

а) вероятность того, что в результате испытания X примет значение, заключенное в промежутке [-1; 0];

б) плотность распределения $f(x)$.

7.Случайная величина X задана интегральной функцией (функцией распределения)

$$F(x) = \begin{cases} 0, & x \le 0 \\ \dfrac{x^2}{18} + \dfrac{x}{6}, & 0 < x \le 3 \\ 1, & x > 3 \end{cases}$$

Найти: а) дифференциальную функцию $f(x)$ (плотность распределения);

б) математическое ожидание;

в) среднее квадратическое отклонение случайной величины X;

г) вероятность попадания значений X в интервал $(0;1)$.

8. В течение 20 биржевых торгов курс доллара составил следующие значения (в рублях):

25,75; 25,8; 25,7; 25,7; 25,6; 25,65; 25,6; 25,65; 25,65; 25,7; 25,8; 25,8; 25,8; 25,7; 25,7; 25,7; 25,7; 25,6; 25,5; 25,65

Найти а) моду Mо б)медиану mе в)размах варьирования R г)средний курс доллара.

9. Из генеральной подлежащих уценке товаров сделана выборка. Известны цены (до проведения уценки) в тыс. руб. $_i$ и частоты n_i их значений в выборочной совокупности. Найти выборочную среднюю цены и ее выборочное средне-квадратическое отклонение.

x_i	2,0	2,2	2,4	2,6	2,8	3,0
n_i	26	15	12	18	16	13

10. Найти доверительный интервал для оценки математического ожидания а нормального распределения с надежностью γ=0,95, зная выборочную среднюю $\bar{x}_B$=82,51, объем выборки n=11 и среднее квадратичное отклонение σ=11.

3. ИТОГОВЫЙ ТЕСТ

1.Найти а) математическое ожидание,

б) среднее квадратическое отклонение случайной величины X, заданной законом распределения

X	-5	2	3	4
P	0,4	0,3	0,1	p_4

2. Непрерывная случайная величина X задана плотностью распределения $f(x)$ =1,5sin3x в интервале (0;π/3) и $f(x)$ = 0 вне этого интервала. Найти вероятность того, что при трех опытах X дважды попадет в интервал (π / 6; π / 4).

3 Случайная величина X задана плотностью распределения

$$f(x) = ax^2 + 4{,}5x - 6 \text{ при } x \in [2;4];\ f(x) = 0 \text{ при } x \notin [2;4].$$

Найти а)значение параметра a;

б) математическое ожидание.

4. Случайная величина X задана плотностью распределения

$$f(x) = \frac{1}{2}x \text{ при } x \in [0;\, c];\, f(x) = 0 \text{ при } x \notin [0; c].$$

Найти а)значение параметра c ;

б)вероятность $P(-100 < X < 1)$.

5. Некоторая случайная величина подчиняется закону нормального распределения с математическим ожиданием 50 и дисперсией 36. Найти вероятность того, что отдельное значение случайной величины заключено в интервале от 40 до 60. Ответ записать с точностью до 0,1.

6. Дискретная случайная величина X принимает три возможных значения: $x_1 = -2$ с вероятностью $p_1 = \frac{1}{6}$, $x_2 = 1$ с вероятностью $p_2 = \frac{1}{3}$ и некоторое значение x_3 с вероятностью p_3. При этом известно, что математическое ожидание случайной величины $M(X)$ равно 1. Найти дисперсию $D(X)$.

7. Случайная величина X задана интегральной функцией (функцией распределения)

$$F(x) = \begin{cases} 0, & x \le 2, \\ \dfrac{x^2 - 4x + 4}{4}, & 2 < x \le 4, \\ 1, & x > 4. \end{cases}$$

Найти а)математическое ожидание случайной величины X (ответ записать в виде десятичной дроби приближенно с точностью до 0,01);

б) вероятность попадания значений X в интервал $(0;3)$.

8. Из генеральной совокупности металлических шайб сделана выборка. Известны внутренние диаметры x_i и частоты n_i этих значений в выборочной совокупности (размеры даны в миллиметрах). Найти выборочную среднюю.

x_i	1,0	1,1	1,2	1,3	1,4	1,5
n_i	10	26	12	18	16	18

ОТВЕТЫ

Задание №1 а) -0,3; б) 3,9

Задание № 2 0,243.

Задание № 3 а) -0,75; б)3

Задание № 4 а)2 ; б) 0,25

Задание № 5 0,9

Задание № 6 2

Задание № 7 а)3,33; б)0,25

Задание № 8 1,258

КЕЙС-ЗАДАНИЯ

Задание 1. Шерлок Холмс расследует дело об ограблении банка. Служительница банка миссис Смит утверждает, что грабителей было трое, мисс Джонсон – что двое, а мистер Пит говорит, что у дам от страха двоилось или троилось в глазах, и что грабитель был один. Шерлок Холмс подозревает, что грабителями могли быть K,L,N (имена в целях тайны следствия он не разглашает, но уверен, что других грабителей быть не могло), каждый из которых прежде участвовал, в среднем, в каждом третьем ограблении.

Свой дедуктивный метод Холмс хочет подтвердить математическими выкладками. Каковы вероятности, что грабил только K; только L; только N; K и L; K и N; L и N; все трое? Каково наиболее вероятное число грабителей? Какова вероятность, что банк вообще ограблен не был, а Смит, Джонсон и Пит присвоили деньги?

Если последует серия подобных ограблений, то каким будет среднее число грабителей?

Задание 2. В течение двадцати дней пронаблюдайте и запишите рублевый курс доллара. Расположите варианты в порядке их возрастания и составьте соответствующий вариационный ряд. Постройте многоугольник распределения. Определите размах варьирования курса доллара, моду и медиану вариационного ряда. Найдите средне-месячный курс доллара. Определить его выборочное средне-квадратическое отклонение.

Задание 3. Транспорт ходит регулярно с интервалом $\tau = 6$ минут. Определить плотность и функцию распределения случайной величины t - времени ожидания транспорта пассажира, в случайный момент времени приходящего на остановку. Какова вероятность, что время ожидания составит не более 1,5 минуты? Какова вероятность, что все пять дней рабочей недели пассажир будет ожидать транспорт не более 1,5 минуты? Каково среднее время ежедневного ожидания данного транспорта?

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ефремов, Л. В. Вероятностное моделирование надежности машин и систем на основе логарифмически – равномерного распределения / Л.В.Ефремов //Труды четвертой международной научной школы «Современные фундаментальные проблемы и прикладные задачи теории точности и качества машин, приборов и систем», 3-7 июня 2000 г. – СПб. - 2001. С. 9-13.

2. Кучерявый, В.И., Мильков С.Н. Статистическое моделирование ресурса валов насосных магистральных агрегатов / В.И. Кучерявый, С.Н. Мильков //.Проблемы машиностроения и надежности машин. – 2010. – № 4. Режим доступа http://naukarus.com/statisticheskoe-modelirovanie-resursa-valov-nasosnyh-magistralnyh-agregatov

3. Гмурман В.Е. Теория вероятностей и математическая статистика / В.Е.Гмурман. – М.:Высшая школа, 2001. – 479 с.

4. Куликов Г.М. Элементы прикладной математики: учебное пособие / Г.М.Куликов, А.Д.Нахман, С.В.Плотникова. – Тамбов.: Изд.-во ТГТУ, 2008. – 160 с.

5. Нахман, А.Д. Элементы теории вероятностей и математической статистики: учебно-методические материалы /А.Д.Нахман. – Тамбов: ТОИПКРО, 2009. – 74 с.

ОГЛАВЛЕНИЕ