МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего профессионального образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Г.В. Вавилова, Н.С. Гнездилова

РКИ: ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-методическое пособие по русскому языку как иностранному

*Рекомендовано в качестве учебного пособия
Редакционно-издательским советом
Томского политехнического университета*

Издательство
Томского политехнического университета
2014

УДК 811.161.1`243: 519.2 (075.8)
ББК Ш141.2 96 141.2-96
В121

Вавилова Г.В.
В121 РКИ: ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА/ Г.В. Вавилова, Н.С. Гнездилова; Томский политехнический университет. − Томск: Изд-во Томского политехнического университета, 2014. – 101 с.

Данное учебное пособие предназначено для иностранных учащихся старших курсов, обучающихся по направлению 200100 «Приборостроение».

Представленные в пособии учебные материалы созданы на основе оригинальных текстов по дисциплине «Математические основы измерений» с учебными заданиями, направленными на изучение особенностей языка науки и формирование навыков работы с научным текстом (чтения, аудирования, понимания, репродуцирования и продуцирования).

УДК 811.161.1`243: 519.2 (075.8)
ББК Ш141.2 96 141.2-96

Рецензенты

Кандидат технических наук,
директор ООО «Потенциал»
И.А. Лежнина

Кандидат технических наук,
доцент кафедры ИИТ ИНК
Д.В. Миляев

Кандидат педагогических наук, доцент кафедры теории языка и методики обучения русскому языку и литературе ТГПУ
А.В. Гузеева

Содержание

От авторов

Учебно-методическое пособие предназначено для иностранцев, обучающихся по программе «2+2», имеющих подготовку по русскому языку в объеме I сертификационного уровня.

Данное пособие поможет улучшить навыки чтения и понимания текстов по специальности, а также развить навыки по аудированию и письму.

В учебном пособии представлены задания по чтению, говорению, письму и аудированию, все тексты и задания ориентированы на понимание узкоспециальных текстов, а также на понимание лекций по специальности. Также в данном пособии представлены оригинальные тексты для самостоятельного чтения, после которых следуют вопросы для самопроверки. В конце учебно-методического пособия представлен словник, содержащий основные термины данной дисциплины, термины и определения переведены на китайский язык.

Тексты и задания данного пособия способствуют развитию умений работы с текстовым материалом, формированию навыков написания текстов на основе прочитанного, составлению плана текста, умению воспроизвести текст по плану. Кроме того, вопросы, приведённые к заданиям в данном пособии, способствуют развитию навыков говорения.

Тема 1 Основные понятия теории вероятности

Задание 1. Составьте все возможные словосочетания, соединив слова из двух колонок. Запишите их. Объясните значения получившихся словосочетаний.

Благоприятный	*Герб*
Теория	*Случайные*
События	*Проводимый*
Опыт (эксперимент)	*Исход*
Случайное	*Цифра*
Появление	*Вероятностей*
Достоверное	*Событие*
Невозможное	

Работа с текстами лекций по специальности

Обратите внимание!

Тезисы – это кратко сформулированные главные мысли статьи, доклада, курсовой или дипломной работы и т. д.

Задание 2. Прочитайте текст. Ответьте на вопросы. Запишите текст в виде тезисов.

Текст

Элементарные события

Теория вероятностей – математическая наука, позволяющая по вероятности одних случайных событий находить вероятности других случайных событий, связанных каким–либо образом между собой.

Любой проводимый опыт (эксперимент) имеет результат. В теории вероятностей результат опыта называют *исходом опыта*.

При бросании монеты возможны два исхода: появление герба – один исход, появление цифры – другой исход. При бросании игральной кости возможны шесть исходов, а именно, появление цифр 1, 2, 3, 4, 5 и 6. При вынимании шаров из урны, в которой находится три красных и два синих шара, возможны пять исходов: два исхода, когда вынимаются синие шары, и три исхода, когда вынимаются красные шары.

Исход или группа исходов, удовлетворяющих определённым требованиям, называются *событием*.

Пример событий: появление при бросании монеты только герба; появление двух гербов при двукратном бросании монеты; появления дамы при вынимании карты из колоды игральных карт; поражение мишени при первом выстреле. Событие при проведении опыта может произойти, а может и не произойти. Если считаем событием появление герба при бросании монеты, а появилась цифра, то событие не произошло. Исходы опытов, соответствующие установленному событию, называются *благоприятными исходами.*

Исходя из вышесказанного, различают события: достоверные, невозможные и случайные.

Достоверное событие – это такое событие, которое всегда происходит в рассматриваемом эксперименте.

Невозможное событие – это такое событие, которое никогда не может наступить в рассматриваемом эксперименте.

Событие, которое при воспроизведении опыта может наступить, а может и не наступить, называется *случайным событием*.

<u>Тезисы</u>

____
____
____
____
____

Вопросы:
1. Что такое *теория вероятностей?*
2. Что значит *исход опыта?*
3. Что называется *событием?*
4. Какие *исходы* называются *благоприятными?*
5. Скажите, какие различают события?
6. Что значит *достоверное событие?*
7. Что значит *невозможное событие?*
8. Какое событие называют *случайным?*

Задание 3. По модели ЧТО (И.п.) НАЗЫВАЕТСЯ ЧЕМ (Тв.п.) дайте определения следующим терминам:

Теория вероятностей ____

____

____

____

____

Исход опыта ____

____

____

____

____

Событие ____

____

____

____

____

Благоприятный исход ____

____

____

____

____

Достоверное событие ____

____

____

____

____

Невозможное событие ____

____

____

____

____

Случайное событие ____

____

____

____

____

ОБРАТИТЕ ВНИМАНИЕ!

«При бросании (Пр.п.) монеты (Р.п.) возможны два исхода: появление герба – один исход, появление цифры – другой исход. При бросании игральной кости возможны шесть исходов, а именно, появление цифр 1, 2, 3, 4, 5 и 6. При вынимании шаров из урны (Р.п), в которой находится три красных и два синих шара, возможны пять исходов: два исхода, когда вынимаются синие шары, и три исхода, когда вынимаются красные шары».

Задание 4. По модели **ПРИ +сущ. (Пр.п.) + сущ. (Р.п.)** составьте как можно больше предложений. Запишите их.

____

____

____

____

____

____

____

____

____

Задание 5. Прочитайте новые слова. Уточните их значение в словаре. Составьте с ними словосочетания или предложения. Запишите.

Урна	*Шар*	*Билет*
Эксперимент	*Вынимание*	*Оба*
Попарно (от «пара»)	*Выигрыш*	*Выпасть/выпадать*

Задание 6. Прочитайте текст. Ответьте на вопросы.

Текст

В урне находится два белых шара, три черных шара. Вынимается один шар. Достоверным событием для данного эксперимента будет вынимание шара любого цвета; невозможным событие – вынимание зеленого шара; случайным событием – вынимание белого шара.

Несколько возможных событий образуют *полную группу событий*, если в результате опыта обязательно должно произойти хотя бы одно из них.

События, образующих полную группу: появление герба или появление цифры при бросании монеты; попадание или промах при выстреле в цель; появление четного или появление нечётного числа при бросании игральной кости.

Пример 1. Приобретены два билета денежно–вещевой лотереи. Обязательно произойдет одно и только одно из следующих событий: «выигрыш выпал на первый билет и не выпал на второй», «выигрыш не выпал на первый билет и выпал на второй», «выигрыш выпал на оба билета», «на оба билета выигрыш не выпал». Эти события образуют полную группу попарно несовместных событий.

Несколько событий в опыте называются *несовместными*, если они не могут произойти одновременно. Несовместными событиями являются попадание в цель и промах при одном выстреле в цель.

В частности, *если события, образующие полную группу, попарно несовместны, то в результате испытания появится одно и только одно из этих событий.*

Равновозможными событиями называются такие события в опыте, каждое из которых объективно не является более возможным, чем другое.

Вопросы и задания:

1. Что образует *полную группу событий?*
2. Какие события в опыте называют *несовместимыми?*
3. Какие события называют *равновозможными*?
4. Перескажите *пример 1* своими словами.

Задание 7. Прочитайте новые слова. Уточните их значение в словаре. Вставьте пропущенные слова в текст.

Монета	*Чеканка*	*Выпадение (чего?)*
Цилиндрический	*Отклонение*	*Рассуждение*
Материал	*Влияющие (на что?)*	*Появление (чего?)*

Текст

Имеется ____________________ правильной ______________________ формы, изготовленная из однородного ________________________; ______________________ произведена без ______________________, ______________________ на ______________________ той или иной стороны монеты. Данные ______________________ приводят к выводу: ____________________ герба и появление цифры при бросании монеты – события равновозможные.

Задание 8. Прочитайте новые слова. Уточните их значение в словаре. Вставьте пропущенные слова в текст. Ответьте на вопросы.

<u>События</u>	*<u>Менее</u>*	*Вероятность*
<u>Возможно</u>	*Сравнивать (что? с чем?)*	*Отношение (чего? к чему?)*

Более	*Степень*	*Утверждать/утвердить*

Текст

Вероятность события. Относительная частота события

Можно утверждать, что одни ________________ ______________ возможны, а другие _________________ ____________________. Появление герба при одном бросании монеты более возможно, чем появление двух гербов при двух последовательных бросаниях монеты.

Чтобы количественно ___________________ между собой события по ________________________ их ______________________, нужно с каждым событием связать некоторое число, которое тем больше, чем более возможно событие. Такое число называют вероятностью события. ______________________ *события* есть численная мера степени объективной возможности этого события.

______________________ *события А* называют отношение благоприятного числа исходов опыта к общему числу всех *равновозможных несовместных элементарных* исходов, образующих *полную группу*:

$$P(A) = \frac{m}{n}, \tag{1.1}$$

где m – благоприятное число исходов опыта, n – общее число исходов опыта.

Выражение (1.1) называется *классической формулой* для вычисления вероятности события.

Вопросы и задания:

1. Какое условие должно выполняться, чтобы количественно сравнивать между собой события по степени их возможности?
2. Что значит *вероятность события?*
3. Прочитайте *классическую формулу* для вычисления вероятности события.

Задание 9. Составьте предложения с подчеркнутыми словами из задания 8.

____

____

____

Задание 10. Прочитайте текст. Ответьте на вопросы.

Текст

Пример 2. В ящике находятся три зеленых и два черных шара. Из урны вынимается один шар. Какова вероятность вынуть при этом черный шар.

Решение.

Общее число исходов в данном опыте равно пяти (по числу шаров в ящике) $n=5$. Число благоприятных исходов равно числу черных шаров в ящике $m=2$. Вероятность события, вынуть белый шар, определяем по классической формуле для вычисления вероятности (1.1)

$$P(A)=\frac{m}{n}=\frac{2}{5}.$$

Из определения вероятности вытекают следующие ее свойства:

Свойство 1. Вероятность достоверного события равна единице.

Если событие достоверно, то каждый элементарный исход испытания благоприятствует событию. В этом случае $m=n$, следовательно,

$$P(A)=\frac{m}{n}=\frac{n}{n}=1.$$

Свойство 2. Вероятность невозможного события равна нулю.

Если событие невозможно, то ни один из элементарных исходов испытания не благоприятствует событию. В этом случае $m=0$, следовательно,

$$P(A) = \frac{m}{n} = \frac{0}{n} = 0.$$

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

Случайному событию благоприятствует лишь часть из общего числа элементарных исходов испытания. В этом случае $0 < m < n$, значит, $0 < \frac{m}{n} < 1$, следовательно,

$$0 < P(A) < 1.$$

Формула (1.1) применима только для определения вероятности события в опытах с *равновероятными* исходами, поэтому имеет ограниченное применение. Если игральная кость не имеет правильную кубическую форму или сделана из неоднородного материала, то нельзя считать появления чисел 1, 2, 3, 4, 5, 6 при бросании кости равновероятными исходами. В этом случае нельзя утверждать, что при бросании кости вероятность появления какой-либо цифры равна 1\6.

Относительная частота наряду с вероятностью принадлежит к основным понятиям теории вероятностей.

Относительной частотой события называют отношение числа испытаний, в которых событие появилось, к общему числу фактически произведенных испытаний. Таким образом, относительная частота события *A* определяется формулой

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

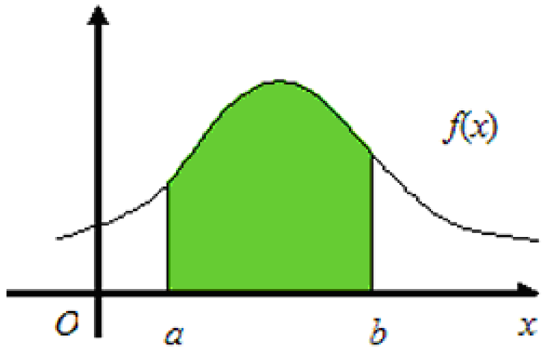
где m – число появлений события, n – общее число испытаний.

Сопоставляя определения *вероятности* и *относительной частоты*, заключаем: определение вероятности не требует, чтобы испытания производились в действительности; определение же относительной частоты предполагает, что испытания были произведены фактически. Другими словами, *вероятность вычисляют до опыта, а относительную частоту – после опыта.*

Вопросы и задания:

1. Перечислите свойства *вероятности события.*
2. Что называют *относительной частотой?*
3. Какая формула применима для определения вероятности события в опытах с *равновероятными* исходами?
4. Перескажите решение *примера 2.*
5. Чем отличаются понятия *вероятность события* и *относительная частота* события?

Тема 2. Случайные величины



Задание 1. Прочитайте слова и словосочетания. Знакомые слова объясните без помощи словаря. Определите значения незнакомых слов по словарю. Скажите, какие слова могут употребляться не только в научном стиле речи? Приведите примеры. Запишите их.

Понятие	*Орудие*	*Центральный*
Теория вероятностей	*Прицел*	*Коэффициент*
Случайная величина (СВ)	*Дискретный*	*Асимметрия*
Погрешность	*Множество*	*Скошенность*
Измерение	*Ряд*	*Эксцесс*
Снаряд	*Многоугольник*	*Отклонение*
Новорожденный	*Функция*	*Дискретный*
Промежуток	*Распределение*	*Устройство*
Непрерывный	*Расход*	*Координаты*
Диапазон	*Электроэнергия*	*Ожидание*
Расплывчатый	*Неопределённый*	*Выстрел*
Мода	*Медиана*	*Абсцисса*

ЗАПОМНИТЕ!

В русском языке для называния понятия, предмета, явления есть 3 грамматические модели:

Что – это что
Что называется чем (чем называется что)
Что является чем

Например:

\# *Что называется чем (чем называется что)*.

Случайной величиной (СВ) называется такая величина, которая в результате опыта может принимать то или иное значение, причём до завершения опыта неизвестно какое.

Что – это что.

Закон распределения случайной величины – это любое правило, устанавливающее соответствие между значениями случайной величины и вероятностями ее наступления.

Что является чем.

Случайная величина (СВ) является такой величиной, которая в результате опыта может принимать то или иное значение, причём до завершения опыта неизвестно какое.

Задание 2. Напишите определения понятий, правильно расставив слова. Используйте модель ЧТО – ЭТО ЧТО, ЧТО НАЗЫВАЕТСЯ ЧЕМ, ЧЕМ НАЗЫАЕТСЯ ЧТО, ЧТО ЯВЛЯЕТСЯ ЧЕМ.

1. *Случайная величина,* в результате опыта может принимать то или иное значение, величина, причём до завершения опыта неизвестно какое.

____

____

____

____

2. *Дискретная величина,* может принимать отдельные, изолированные значения с определенными вероятностями, случайная величина.

____

____

____

____

3. *Закон распределения* случайной величины, устанавливающее соответствие между значениями случайной величины и вероятностями ее наступления, любое правило.

____

____

____

____

Работа с текстами лекций по специальности

Задание 3. Прочитайте текст. Ответьте на вопросы.

Текст
Случайная величина

Одним из важнейших понятий в теории вероятностей является понятие случайной величины. *Случайной величиной* (СВ) называется такая величина, которая в результате опыта может принимать то или иное значение, причём до завершения опыта неизвестно какое. Примером случайной величины может служить погрешность измерения.

Пример 1. Число родившихся мальчиков среди ста новорожденных есть случайная величина, которая имеет следующие возможные значения: 0, 1, 2, ..., 100.

Пример 2. Расстояние, которое пролетит снаряд при выстреле из орудия, есть случайная величина. Действительно, расстояние зависит не только от установки прицела, но и от многих других причин (силы и направления ветра, температуры и т.д.), которые не могут быть полностью учтены. Возможные значения этой величины принадлежат некоторому промежутку (a, b).

Случайные величины делят на непрерывные и дискретные. Дискретная случайная величина в заданном диапазоне может принимать конечное число значений. Например, число попаданий при 10 выстрелах в цель является дискретной случайной величиной с возможными значениями от нуля до десяти. Непрерывная случайная величина в заданном диапазоне значений может принимать любое значение и число таких значений бесконечно. Границы непрерывной случайной величины могут быть точно определены, например координаты пасущейся на огороженном участке коровы, а чаще всего являются неопределёнными, расплывчатыми, например погрешность измерения.

Итак, случайные величины могут быть дискретными или непрерывными, а область возможных исходов может быть представлена конечным множеством, счетным или бесконечным.

Вопросы и задания:

1. Какое понятие в теории вероятностей является важным?
2. Что такое *случайная величина*?
3. Что может служить примером случайной величины?
4. Перескажите примеры 1 и 2 другими словами.
5. На какие величины делят случайные величины?
6. Какая величина может принимать конкретное число значений?
7. Какая величина может принимать любое значение (число этих значений бесконечно)?

Обратите внимание!

План текста – это «каркас» текста, который последовательно, абзац за абзацем, отражает изложение материала.
План текста может быть *называным, вопросным, цитатный, план – опорная схема.*

Задание 4. Составьте план текста из задания 3. Запишите. Восстановите текст, опираясь на план (устно или письменно).

План

____
____
____
____
____
____
____
____
____
____
____
____
____
____
____
____
____

Задание 5. Прочитайте текст. Ответьте на вопросы.

Текст
Ряд, многоугольник и функция распределения дискретной случайной величины

Дискретной называют случайную величину, которая может принимать отдельные, изолированные значения с определенными вероятностями (ДСВ).

Закон распределения случайной величины – любое правило, устанавливающее соответствие между значениями случайной величины и вероятностями ее наступления.

Дискретная случайная величина X может быть задана рядом распределения или функцией распределения (интегральным законом распределения).

Рядом распределения называется совокупность всех возможных значений x_i и соответствующих им вероятностей $p_i = P(X = x_i)$, он может быть задан в виде таблицы 2.1:

Таблица 2.1

x_1	x_2	…	x_n
p_1	p_2	…	p_n

При этом вероятности p_i удовлетворяют условию

$$p_1 + p_2 + ... + p_n = 1,$$

так как события $\{X = x_i\}$, ..., $\{X = x_n\}$ несовместны и образуют полную группу.

Графическое изображение ряда распределения называется *многоугольником распределения.* Для его построения возможные значения случайной величины (x_i) откладываются по оси абсцисс, а вероятности p_i – по оси ординат; точки A_i с координатами (x_i, p_i) соединяются ломаными линиями.

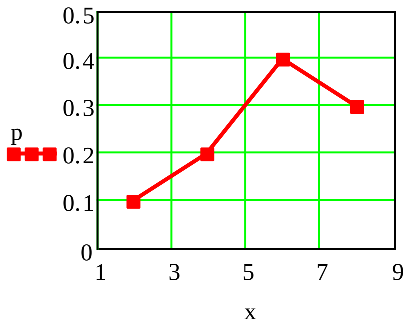


Рисунок 2.1 – Многоугольник распределения

Функцией распределения случайной величины x называется функция *F(x),* значение которой в точке *x* равно вероятности того, что случайная величина *X* будет меньше этого значения *x*, то есть

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

Функция $F(x)$ для дискретной случайной величины вычисляется по формуле

$$F(x) = \sum_{x_i < x} p_i,$$

где суммирование ведется по всем значениям i, для которых $x_i < x$.

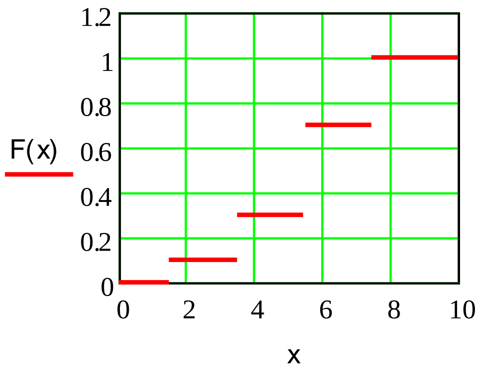


Рисунок 2.2 – Функция распределения

Функция распределения обладает следующими свойствами:

1. $F(-\infty) = 0,\ \ F(+\infty) = 1,$
2. $F(x_1) \le F(x_2),$ при $x_1 < x_2,$
3. $P(a \le X < b) = F(b) - F(a).$

Вопросы и задания:

1. Дайте определение понятию «*дискретная величина»*.
2. Что такое *закон распределения* случайной величины?
3. Какая случайная величина может быть задана рядом распределения?
4. Что такое *ряд распределения*?
5. Что значит *многоугольник распределения*?
6. Опишите рисунок 2.1.
7. Какими свойствами обладает *функция распределения*?

Задание 6. Прочитайте текст. Выпишите из текста все определения, записав их по моделям ЧТО - ЭТО ЧТО, ЧТО ЯВЛЯЕТСЯ ЧЕМ, ЧТО НАЗЫВАЕТСЯ ЧЕМ (ЧЕМ НАЗЫВАЕТСЯ ЧТО) (измените модели, по которым они написаны в тексте). Ответьте на вопросы.

Текст
Функция распределения и плотность вероятности непрерывной случайной величины

Непрерывной называют случайную величину, возможные значения которой непрерывно заполняют некоторые конечные или бесконечные промежутки (НСВ).

Случайная величина называется *непрерывной*, если существует неотрицательная функция *f(x),* удовлетворяющая при любых х равенству:

$$F(x) = \int_{-\infty}^{x} f(x)dx. \tag{2.1}$$

Плотностью распределения (или плотностью вероятности) непрерывной случайной величины X в точке x называется производная ее функции распределения в этой точке, обозначаемая *f(x)*. Функция *f(x)* определяется соотношением

$$f(x) = \lim_{\Delta x \to 0} \frac{P(x < X < x + \Delta x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = \frac{dF(x)}{dx}. \tag{2.2}$$

График плотности распределения называется *кривой распределения*.

Непрерывная случайная величина задается либо *функцией распределения* $F(x)$ (интегральным законом распределения), либо *плотностью вероятности* $f(x)$ (дифференциальным законом распределения).

В геометрической интерпретации $F(x)$ равна площади, ограниченной сверху кривой плотности распределения $f(x)$ и лежащей левее точки x.

Основные свойства плотности распределения:

1. Плотность распределения неотрицательна: $f(x) \geq 0$.
2. Условие нормировки: $\int_{-\infty}^{\infty} f(x)dx = 1$.

<u>Определения:</u>

____
____
____
____
____
____
____
____
____
____
____
____
____

Вопросы и задания:

1. Какую случайную величину называют *непрерывной*?
2. При каком условии случайная величина считается непрерывной?
3. Прочитайте формулу для функции распределения (формула 2.1).
4. Прочитайте формулу для плотности распределения (формула 2.2).
5. Приведите синоним для термина «кривая распределения».
6. Назовите основные свойства плотности распределения.

Обратите внимание!

Аннотация – это краткое, обобщенное описание книги, статьи.

Образцы клише в аннотации:
В тексте исследуется
(что?)………………………………………………………….
Большое место в работе занимает исследование
(чего?)…………………….
В книге анализируется
(что?)…………………………………………………………
В тексте на основании (чего) показано
(что?)………………………………….

Задание 7. Прочитайте текст. Приготовьтесь отвечать на вопросы. Напишите аннотацию к тексту.

Текст

Числовые характеристики случайной величины

Для решения практических задач бывает достаточно знать лишь некоторые числовые параметры, которые позволяют представить основные особенности СВ в сжатой форме. Их называют *числовыми характеристиками СВ.*

Одна из основных характеристик СВ - *математическое ожидание*:

$$m_x = M[x] = \begin{cases} \sum_i x_i \cdot P\{X = x_i\} & \text{для} \quad \text{ДСВ}, \\ \int_{-\infty}^{\infty} x \cdot f(x)dx & \text{для} \quad \text{НСВ}. \end{cases}$$

Математическое ожидание характеризует среднее значение СВ и обладает следующими свойствами:

1. $M[c] = c,$
2. $M[c \cdot X] = c \cdot M[X],$
3. $M[X + c] = M[X] + c,$
4. $M[X_1 + X_2] = M[X_1] + M[X_1].$

Модой случайной величины называется ее наиболее вероятное значение, т.е. то значение, для которого вероятность p_i (для дискретной

СВ) или *f(x)* (для непрерывных СВ) достигает максимума. Обозначения: *Mx, Mo*.

Медианой случайной величины *X* называется такое ее значение, для которого выполняется условие $P\{X < Me\} = P\{X \geq Me\}$. Медиана, как правило, существует только для непрерывных случайных величин.

Геометрически *медиана* – это абсцисса тоски, в которой площадь, ограниченная кривой плотности распределения *f(x)*, делится пополам.

Начальный момент s-го порядка СВ *X* есть математическое ожидание *s*-й степени этой случайной величины:

$$\alpha_S = M[X^S].$$

Центрированной случайной величиной называется отклонение СВ от математического ожидания:

$$\overset{0}{X} = X - m_x.$$

Центральный момент s-го порядка СВ *X* есть математическое ожидание *s*-й степени центрированной случайной величины:

$$\mu_s = M\left[\overset{0}{X}{}^s\right] = M\left[\left(X - m_x\right)^s\right].$$

Для любой случайной величины центральный момент первого порядка равен 0.

Коэффициентом асимметрии (или скошенности) распределения называется величина

$$A = \frac{\mu_3}{\sigma^3}.$$

Для симметричной СВ $\mu_3 = 0$

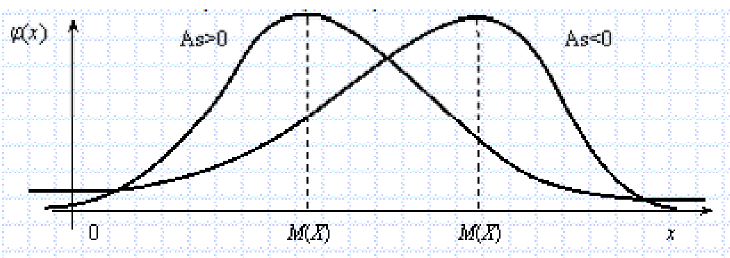


Рисунок 2.3 – Коэффициент асимметрии

Коэффициентом эксцесса (или островершинности) распределения называется величина

$$E = \frac{\mu_4}{\sigma^4} - 3.$$

Коэффициент эксцесса является обобщенной характеристикой плосковершинности плотности распределения.

Для нормального распределения эксцесс $E = 0$, для более островершинных законов распределения $E > 0$, а для плосковершинных $E < 0$.

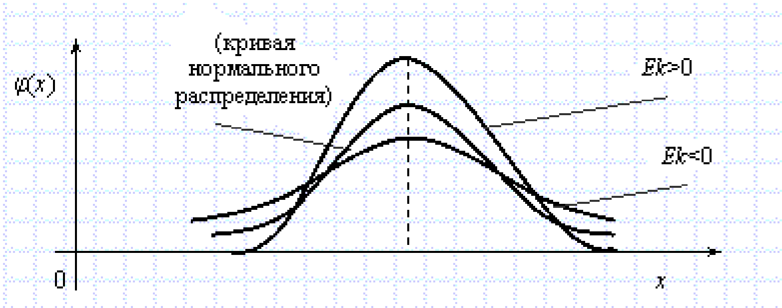


Рисунок 2.4 – Коэффициент эксцесса

Дисперсия случайной величины есть математическое ожидание квадрата соответствующей центрированной случайной величины.

Расчетные формулы:

$$D[x] = \begin{cases} \sum\limits_i (x_i - m_x)^2 \cdot P\{X = x_i\} = \sum\limits_i x_i^{\,2} \cdot P\{X = x_i\} - (m_x)^2 & \text{для} \quad ДСВ, \\ \int\limits_{-\infty}^{\infty} (x - m_x)^2 \cdot f(x)dx = \int\limits_{-\infty}^{\infty} x^2 \cdot f(x)dx - (m_x)^2 & \text{для} \quad НСВ. \end{cases}$$

Вычислить дисперсию можно и через второй начальный момент:

$$D[X] = M\left[X^2\right] - m_x^2.$$

Дисперсия случайной величины характеризует степень рассеивания (разброса) значений случайной величины относительно ее математического ожидания и обладает следующими свойствами:

1. $D[c] = 0,$
2. $D[X + c] = D[c],$
3. $D[c \cdot X] = c^2 \cdot D[X],$

4. $\sigma[c \cdot X] = c \cdot \sigma[X]$.

Средним квадратическим отклонением (СКО) СВ X называется характеристика

$$\sigma_x = \sigma[X] = \sqrt{D[X]}.$$

СКО измеряется в тех же физических единицах, что и СВ, и характеризует ширину диапазона значений СВ.

Вопросы и задания:

1. Выберите одну из основных характеристик СВ (мат. ожидание, дисперсия и т. д.). Прочитайте формулы для вычисления выбранной характеристики СВ.
2. Прочитайте четыре свойства, которыми обладает математическое ожидание.
3. Что называется модой случайной величины?
4. Что называется медианой случайной величины?
5. Что такое медиана с точки зрения геометрии?
6. Какая величина называется коэффициентом асимметрии (прочитайте)?
7. Прочитайте рисунок 2.3.
8. Какая величина называется коэффициентом эксцесса (прочитайте)?
9. Прочитайте рисунок 2.4.
10. Прочитайте свойства дисперсии СВ.
11. Прочитайте характеристику среднего квадратического отклонения (СКО) СВ X.

$$\sigma[X] = \sqrt{D[X]} = \sqrt{2}.$$

<u>Аннотация</u>

____

____

____

____

____

____

____

____
____
____
____
____
____
____

Обратите внимание!

Аббревиатура – это слово, 1) образованное из первых букв словосочетания (РФ, КНР, США, вуз); 2) образованное из начальных частей слов (колхоз); 3) образованное путем сложения начала одного слова с другим словом (спецодежда).

Задание 9. Расшифруйте аббревиатуру. Дайте определение терминам.

СВ - ____
____

НСВ - ____
____

ДСВ- ____
____

СКО - ____
____

Тема 3 Элементы математической статистики

Задание 1. Прочитайте слова. Значение знакомых слов объясните без помощи словаря. Определите значения незнакомых слов по словарю.

Статистика	*Интерпретация*
Математическая	*Опираться/опереться (на что)*
Раздел	*Процедура*
Сбор	*Статистический*
Систематизация	*Данные*
Обработка	*Сплошной*
Обследование	*Признак*
Уничтожение	*Однородный*
Выборка	*Стандартность (чего?)*
Генеральный	*Облегчение (чего?)*
Совокупность	*Сказываться (на чём?)*
Репрезентативность (представительность)	*Повтор*

Работа с текстами лекций по специальности

Задание 2. Прочитайте текст. Напишите аннотацию к тексту. Ответьте на вопросы.

Текст

Математическая статистика – это раздел математики, посвященный математическим методам сбора, систематизации, обработки и интерпретации статистических данных, а также использование их для научных или практических выводов. Правила и процедуры математической статистики опираются на теорию вероятностей, позволяющую оценить точность и надежность выводов, получаемых в каждой задаче на основании имеющегося статистического материала. *Статистическими данными* называются сведения о числе объектов в какой-либо более или менее обширной совокупности, обладающих теми или иными признаками.

Пусть требуется изучить совокупность однородных объектов относительно некоторого качественного или количественного признака, характеризующего эти объекты. Например, если имеется партии деталей, то качественным признаком может служить стандартность детали, а количественным – контролируемый размер детали.

Иногда проводят сплошное обследование, т.е. обследуют каждый из объектов совокупности относительно признака, которым интересуются. На практике, однако, сплошное обследование применяют сравнительно редко. Например, если совокупность содержит очень большое число объектов, то провести сплошное обследование физически невозможно. Если обследование объекта связано с его уничтожением или требует больших материальных затрат, то проводить сплошное обследование практически не имеет смысла. В таких случаях случайно отбирают из всей совокупности ограниченное число объектов и подвергают их изучению.

Выборочной совокупностью или просто *выборкой* называют совокупность случайно отобранных объектов.

Генеральной совокупностью называют совокупность объектов, из которых производится выборка.

Объемом совокупности (выборочной или генеральной) называют число объектов этой совокупности.

Если из 1000 деталей отобрано для обследования 100 деталей, то объем генеральной совокупности $N = 1000$, а объем выборки $n = 100$.

Часто генеральная совокупность содержит конечное число объектов. Однако если это число достаточно велико, то иногда в целях упрощения вычислений, или для облегчения теоретических выводов, допускают, что генеральная совокупность состоит из бесчисленного множества объектов. Такое допущение оправдывается тем, что увеличение объема генеральной совокупности (достаточно большого объема) практически не сказывается на результатах обработки данных выборки.

При составлении выборки можно поступать двумя способами: после того как объект отобран и над ним произведено наблюдение, он может быть возвращен либо не возвращен в генеральную совокупность. В соответствии со сказанным выборки подразделяют на повторные и бесповторные.

Повторной называют выборку, при которой отобранный объект (перед отбором следующего) возвращается в генеральную совокупность.

Бесповторной называют выборку, при которой отобранный объект в генеральную совокупность не возвращается. На практике обычно пользуются бесповторным случайным отбором.

Для того чтобы по данным выборки можно было достаточно уверенно судить об интересующем признаке генеральной совокупности, необходимо при формировании выборки обеспечить:

- Достаточный объем выборки;
- Репрезентативность (представительность).

Необходимый объем выборки зависит от решаемой задачи и от заданной точности и надежности получаемых результатов, в каждом конкретном случае определяется по-своему.

Выборка называется *репрезентативной (представительной),* если она правильно представляет генеральную совокупность. В силу закона больших чисел можно утверждать, что выборка будет репрезентативной, если ее осуществить случайно: каждый объект выборки отобран случайно из генеральной совокупности, если все объекты имеют одинаковую вероятность попасть в выборку.

Если объем генеральной совокупности достаточно велик, а выборка составляет лишь незначительную часть этой совокупности, то различие между повторной и бесповторной выборками стирается. В предельном случае, когда рассматривается бесконечная генеральная совокупность, а выборка имеет конечный объем, это различие исчезает.

Аннотация

Вопросы:

1. Что такое *выборка*?
2. При каких условиях выборка будет репрезентативной?
3. Что такое *повторная выборка?*
4. Чем отличается повторная выборка *от бесповторной?*
5. Что называется *объемом* выборки?

Задание 3. Составьте план текста из задании 2. Запишите его. Восстановите текст по плану.

План

Задание 4. Напишите определения приведённым ниже терминам по модели ЧТО НАЗЫВАЕТСЯ ЧЕМ (ЧЕМ НАЗЫВАЕТСЯ ЧТО):

Математическая статистика, статистические данные, выборочная совокупность, генеральная совокупность, объем совокупности, повторная выборка, бесповторная выборка, репрезентативная выборка.

Тема 4. Основы теории оценивания

Задание 1. Прочитайте слова и словосочетания. Значение знакомых слов объясните без помощи словаря. Уточните значения незнакомых слов по словарю.

Свойства (чего?)	*Оценка*
Обширный	*Многократный*
Класс	*Детальное исследование*
Гипотеза	*Ограниченное число (чего?)*
Зависимость	*Наблюдение*
Внутренний	*Конечный (-ая, -ое)*
Содержание	

Задание 2. Внимательно прочитайте части текста. Расположите их в правильном порядке.

<u>1</u>	Основная цель статистического анализа состоит в исследовании свойств случайных величин.
	Задачи статистического анализа делят на три обширных класса: 1. Статистическое оценивание; 2. Проверка статистических гипотез; 3. Построение статистических зависимостей.
	Эти задачи существенно различаются как по внутреннему содержанию, так и по методам их решения.
	При этом различают правила оценивания, называемые *методами получения оценок.*
	Экспериментальной основой таких исследований являются результаты многократных наблюдений случайной величины.
	В этом разделе уделяется внимание первому классу задач: статистическому оцениванию. Задачи оценивания направлены на вычисление количественных характеристик случайных величин по конечной выборке.
	При статистическом анализе в каждом конкретном случае возможно решение ряда задач, каждая из которых в общем случае включает два этапа: обработку выборки и принятие решения. Поскольку существуют условия неопределенности из-за конечного объема и случайного содержания выборки, принимаемое решение носит вероятностный характер.
	Выборка подвергается детальному исследованию, по результатам которого требуется определить свойства (характеристики) генеральной совокупности.
	Применение методов математической статистики для обработки результатов измерений базируется на том, что многократные измерения соответствуют *выборочному методу*.
	Сущность *выборочного метода* заключается в следующем: имеется некоторая большая совокупность объектов, называемая *генеральной совокупностью*. Из этой совокупности случайным образом извлекается ограниченное число объектов, образующих *выборку*.

Работа с текстами лекций по специальности

Задание 3. Прослушайте текст. Проверьте себя, правильно ли вы расставили цифры в задании 2. Ответьте на вопросы.

Вопросы:

1. В чем заключается основная цель статистического анализа?
2. На какие классы делятся задачи статистического анализа?
3. На вычисление чего направлены задачи оценивания?
4. В чем заключается сущность *выборочного метода?*

Задание 4. Прослушайте продолжение текста. Вставьте пропущенные слова.

Текст

В распоряжении ________________________ имеются лишь данные _______________________, например значения ________________________ признака x_1, x_2, ..., x_n, полученные в ______________________ n наблюдений. Через эти данные и выражают оцениваемый ________________________ Θ. Поскольку истинное значение искомой _____________________ получить принципиально невозможно, то в качестве его принимают некоторое _________________ по выборке значение $\tilde{\Theta}$. Полученное по результатам ____________________ приближенное значение $\tilde{\Theta}$ называют *оценкой истинного значения* параметра Θ, которую определяют как функцию результатов наблюдений объема n:

$$\tilde{\Theta} = f(x_1, x_2, ..., x_n).$$

Чтобы оценки в большинстве опытов были _____________________ близки к __________________ значению измеряемой величины, к ним предъявляют следующие_____________________.

1. *Несмещенность.* Оценку называют несмещенной, если при любом объеме выборки n _____________________________ ожидание оценки равно истинному значению ___________________________ параметра:

$$M\left[\overline{\Theta}\right] = \Theta.$$

Во многих случаях смещение убывает с возрастанием ________________________ выборки.

2. *Эффективность.* ____________________________ называю статистическую оценку, которая при _______________________ объеме выборки n имеет наименьшую возможную дисперсию.

3. *Состоятельность.* Оценка является __________________________, если при $n \to \infty$ __________________ по вероятности к оцениваемому параметру. Например, если ____________________ несмещенной оценки при $n \to \infty$ стремится к нулю, то такая оценка оказывается и состоятельной.

Состоятельность оценки означает, что чем больше объем _________________________ n, тем больше ___________________ того, что ошибка не превысит сколь угодно малого ______________________ числа ε.

$$\lim_{n\to\infty} P\left(\left|\tilde{\Theta}_n - \Theta\right| < \varepsilon\right) = 1.$$

4. *Устойчивость* ______________________ чувствительность оценки к аномальным результатам (__________________ и выбросам).

Задание 5. Прочитайте слова и словосочетания. Значение знакомых слов объясните без помощи словаря. Уточните значения незнакомых слов по словарю. Составьте с подчёркнутыми словами предложения или словосочетания.

<u>Точечный</u>	*<u>Взвешенный</u> (-ая, -ое)*	*<u>Нахождение</u> (чего?)*
<u>Предполагать</u> (что?)	*Выборочная средняя*	*<u>Велик</u> (-а, -о)*
<u>Непригодный</u>	*<u>Исправить/исправлять</u> (что?)*	*Квадратический*

____
____
____
____
____
____

Обратите внимание!

Абзац – в начале первой строки текста или части текста. С помощью абзацев текст делится на смысловые отрезки.

Задание 6. Прочитайте текст. Разделите текст на абзацы. Обоснуйте выделение абзацев.

Текст

Точечная оценка параметров распределения

Точечная оценка предполагает нахождение единственной числовой величины, которая и принимается за значение параметра. Такую оценку целесообразно определять в тех случаях, когда объем экспериментальных данных достаточно велик. Причем не существует единого понятия о достаточном объеме экспериментальных данных (ЭД), его значение зависит от вида оцениваемого параметра. При малом объеме ЭД точечные оценки могут значительно отличаться от истинных значений параметров, что делает их непригодными для использования. *Генеральной средней* $\bar{x}_г$ называют среднее арифметическое значений признака генеральной совокупности. Если рассматривать обследуемый признак X генеральной совокупности как случайную величину, то математическое ожидание признака равно генеральной средней этого признака: $M(X) = \bar{x}_г$. В результате независимых наблюдений над количественным признаком X из генеральной совокупности извлечена повторная выборка объема n со значениями признака $x_1, x_2, \ldots, x_n$. Генеральная средняя $\bar{x}_г$ неизвестна и требуется оценить ее по данным выборки. В качестве оценки генеральной средней $\bar{x}_г$ принимают выборочную среднюю $\bar{x}_в$. *Выборочной средней* $\bar{x}_в$ называют среднее арифметическое значение признака выборочной совокупности. Для того чтобы охарактеризовать рассеяние значений количественного признака X генеральной совокупности вокруг своего среднего значения, вводят сводную характеристику – генеральную дисперсию. *Генеральной дисперсией* $D_г$ называют среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения $\bar{x}_г$:

$$D_г = \frac{\sum_{i=1}^{N}(x_i - \bar{x}_г)^2}{N}$$

Для того чтобы охарактеризовать рассеянность наблюдаемых значений количественного признака выборки вокруг своего среднего значения $\bar{x}_в$, вводят сводную характеристику – выборочную дисперсию. *Выборочной дисперсией* $D_в$ называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения $\bar{x}_в$:

$$D_в = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_в)^2}{n}$$

Если в качестве оценки генеральной дисперсии принять выборочную дисперсию, то эта оценка будет приводить к систематическим ошибкам, давая заниженное значение генеральной дисперсии. Объясняется это тем, что выборочная дисперсия является смещенной оценкой $D_г$. Другими словами, математическое ожидание выборочной дисперсии не равно оцениваемой генеральной дисперсии, а равно

$$M[D_в] = \frac{n-1}{n} D_г.$$

Легко «исправить» выборочную дисперсию так, чтоб ее математическое ожидание было равно генеральной дисперсии. Достаточно для этого умножить $D_в$ на дробь $\frac{n}{n-1}$. Сделав это, получим *исправленную дисперсию,* которую обычно обозначают через S^2:

$$S^2 = \frac{n}{n-1} D_в = \frac{n}{n-1} \frac{\sum_{i=1}^{n}(x_i - \bar{x}_в)^2}{n} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}_в)^2}{n-1}$$

При достаточно больших значениях n объема выборки выборочная и исправленная дисперсии различаются мало. На практике пользуются исправленной дисперсией, если примерно $n < 30$.

Задание 7. Составьте 10-15 вопросов к тексту из задания 6.

____
____
____
____
____
____
____
____
____
____
____
____
____
____
____
____

Задание 8. Прослушайте текст. Вставьте пропущенные слова.

Текст

Методы получения точечных оценок

На __________________ при оценке ____________________ не всегда удается _________________________ одновременно требование состоятельности, несмещенности и эффективности. __________________ выбору оценки всегда __________________ предшествовать критическое _______________ ее со всех _____________ ________________

Количественные значения оценок искомых ___________________ в общем случае являются __________________ элементов выборки.

Для нахождения ______________________ вычисления оценок наиболее широкое ___________________________________ получили следующие методы: ________________________ правдоподобия, моментов, порядковых ____________________________ и др.

Задание 9. Прочитайте новые слова и словосочетания. Значение знакомых слов объясните без помощи словаря. Уточните значения незнакомых слов в словаре. Составьте с подчеркнутыми словами предложения или словосочетания.

<u>*Максимальный*</u>	<u>*Произведение*</u>
Правдоподобие	*Априорное (доопытное) распределение*
Обобщенный	<u>*Наиболее*</u>

____

____

____

____

____

____

____

Задание 10. Прочитайте текст. Ответьте на вопросы. Выпишите все определения терминам, данным в тексте, по модели ЧТО ЯВЛЯЕТСЯ ЧЕМ.

Текст
Метод максимального правдоподобия

Метод максимального правдоподобия, предложенный Гауссом в 1848 г. и обобщенный Фишером, с теоретической точки зрения является наиболее разработанным методом получения оценок. Идея метода достаточно проста и наглядна: он основан на априорном (доопытном) распределении выборки. Если $x = (x_1, x_2, ..., x_n)$ - результаты наблюдений при измерении параметра Θ имеют плотность распределения $f(x,\ \Theta)$, то наиболее вероятным значением оцениваемых величин будут такие, при которых совместная плотность распределения выборки, называемая *функцией правдоподобия,* достигает максимума. При независимых результатах наблюдений функция правдоподобия равна произведению плотностей распределения результатов наблюдений:

$$L = \prod_{i=1}^{n} f(x_i, \tilde{\Theta}).$$

Для нахождения оценок необходимо исследовать функцию правдоподобия на экстремум. С целью упрощения вычислительной процедуры обычно максимизируют не саму функцию правдоподобия, а ее логарифм $\ln L(x, \tilde{\Theta})$. Если функция правдоподобия достаточно гладкая и дифференцируемая, то нахождение оценок сводится к решению нелинейного уравнения (уравнений) относительного искомого параметра (параметров) вида

$$\frac{d \ln L(x, \tilde{\Theta})}{d\tilde{\Theta}} = 0.$$

Определения

Вопросы:

1. В чём заключается *метод максимального правдоподобия?*
2. Кем и когда был предложен метод максимального правдоподобия?
3. Чему равна *функция правдоподобия* при независимых результатах наблюдения?
4. Зачем необходимо исследовать функцию правдоподобия на экстремум?

Задание 11. Прочитайте слова. Значение знакомых слов объясните без помощи словаря. Уточните значения незнакомых слов в словаре. С подчёркнутыми словами составьте предложения. Запишите их.

<u>Момент</u>	*Следуя (+Д.п.)*	*Эмпирический*
<u>Предложен (кем?)</u>	*Состоятелен*	*<u>Порядок (чего?)</u>*
Приравнивание	*<u>Целесообразно</u>*	*Соответственно*

Задание 12. Прочитайте текст. Составьте план к тексту. По плану восстановите текст.

Метод моментов

Метод моментов предложен К. Пирсоном в 1894г. Сущность метода состоит в приравнивании теоретических моментов рассматриваемого распределения соответствующим эмпирическим моментам того же порядка и решении полученной системы уравнений относительно неизвестных параметров распределения. Выбирается столько эмпирических моментов, сколько требуется оценить неизвестных параметров распределения. Желательно применять моменты младших порядков, так как погрешности вычисления оценок резко возрастают с увеличением порядка момента.

Можно доказать, что начальные и центральные эмпирические моменты являются состоятельными оценки соответственно начальных и центральных теоретических моментов того же порядка.

$$\tilde{\alpha}_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i\right)^k,\ \tilde{\mu}_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^k.$$

Теоретическими моментами называются моменты, вычисленные аналитическим путем через плотность *f(x)* или функцию распределения *F(x)*. Теоретические моменты являются функциями параметров распределения:

$$\alpha_k = \alpha_k\left(\Theta_1, \Theta_2, ..., \Theta_n\right),\ \mu_k = \mu_k\left(\Theta_1, \Theta_2, ..., \Theta_n\right).$$

Рассмотрим случай, когда метод моментов используется для нахождения оценки одного параметра. Положим, что плотность распределения $f(x;\Theta)$ случайной величины X зависит только от одного параметра Θ, и необходимо найти оценку параметра Θ. Для нахождения оценки одного параметра достаточно иметь одно уравнение относительно этого параметра, используя, например, на основании выборки $x_1, x_2, ..., x_n$ первый начальный момент.

Следуя методу моментов, приравняем, например, начальный теоретический момент первого порядка начальному эмпирическому $\alpha_1 = M_1$. Учитывая, что $\alpha_1 = M(X)$ и $M_1 = \overline{x}_в$, получим $M(X) = \overline{x}_в$.

Математическое ожидание $M(X)$, как видно из соотношения $M(X) = \int\limits_{-\infty}^{\infty} x f(x;\Theta)dx = \varphi(\Theta),$ есть функция от Θ. Решив это уравнение относительно параметра Θ, найдем его точечную оценку $\bar{\Theta}$, которая является функцией от выборочной средней, следовательно, и от вариант выборки.

Метод моментов позволяет получить состоятельные, достаточные оценки, они при довольно общих условиях распределены асимптотически нормально. Смещение удается устранить введением поправок. Эффективность оценок невысокая, т.е. даже при больших объемах выборок дисперсия оценок относительно велика (за исключением нормального распределения, для которого метод моментов дает эффективные оценки). В реализации метод моментов проще метода максимального правдоподобия. Напомним, что метод целесообразно применять для оценки не более чем четырех параметров, так как точность выборочных моментов резко падает с увеличением их порядка.

<u>План</u>

____
____
____
____
____
____
____
____
____
____
____
____

Обратите внимание!

Синонимы – это слова, разные по звучанию и написанию, но имеющие одинаковое или очень близкое значение.

Задание 13. Прочитайте текст. Найдите подчеркнутые слова и выражения. Замените их на синонимичные, скажите другими словами.

Текст
Интервальное оценивание

Так как оценки являются случайными величинами, то их отклонения от оцениваемых параметров также случайны. Истинное значение искомого параметра может быть представлено точечной оценкой $\Theta = \overline{\Theta} \pm \sigma_{\overline{\Theta}}$, но данная оценка не всегда объективна.

Чтобы <u>дать представление о</u> точности и надежности оценки в математической статистике, пользуются так называемыми доверительными интервалами и доверительными вероятностями.

Доверительным интервалом называется интервал $[\Theta_{i1};\Theta_{i2}]$, между границами которого с заданной вероятностью $P = 1 - \alpha$ будет заключено истинное значение x_i, т.е.

$$P\left[\Theta_{i1} < \Theta_i < \Theta_{i2}\right] = 1 - \alpha.$$

Вероятность $P = 1 - \alpha$ <u>принято называть</u> *доверительной вероятностью,* а α *– уровнем значимости.*

Чем уже интервал $[\Theta_{i1};\Theta_{i2}]$, тем точнее оценка неизвестного параметра Θ_i. Задача состоит в нахождении по заданной доверительной вероятности границ этого интервала. В измерительной практике доверительная вероятность обычно принимается равной 0,90; 0,95; 0,98; 0,99 и реже (в особо ответственных случаях) 0,999 в зависит от конкретных условий

С вероятностью P справедливо неравенство: $\bar{\Theta}_{p1} \le \Theta \le \bar{\Theta}_{p2}$. Представив отклонения квантилей от центра распределения через среднее квадратичное отклонение оценки $\Theta = \bar{\Theta} \pm t_p \sigma_{\bar{\Theta}}$, имеем интегральную оценку вида

$$\bar{\Theta} - t_p \sigma_{\bar{\Theta}} \le \Theta \le \bar{\Theta} + t_p \sigma_{\bar{\Theta}}.$$

Данное неравенство справедливо для симметричного распределения оценки. В противном случае имеют место два коэффициента t_1 *и* t_2.

Полученный доверительный интервал с вероятностью P заключает в себе (накрывает) истинное значение: истинное значение параметра постоянно, а доверительный интервал случайный по положению и ширине.

Определение законов распределения оценок - сложная и трудоемкая задача. Поэтому желательно найти такую функцию $f(\bar{\Theta}, \Theta)$, точное распределение которой известно и не зависит от Θ. Например, при интервальном оценивании параметров нормального распределения такими функциями являются переменные Стьюдента $t = (\bar{x} - m) \Big/ \sqrt{\frac{S^2}{n}}$ и Пирсона $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$, которые, как известно, распределены соответственно по законам Стьюдента и хи-квадрат с $n - 1$ степенями свободы.

Задание 14. Выпишите все определения по модели ЧТО - ЭТО ЧТО из текста (задание 13). Составьте 10-12 вопросов к тексту.

____

____

____

____

____

Тема 5. Проверка статистических гипотез

Задание 1. Прочитайте слова. К подчёркнутым словам подберите синонимы. Составьте с ними предложения или словосочетания. Запишите их.

Статистическая	*Выдвигаемая*
Проверка	*<u>Противоречить</u>*
<u>Позволять</u>	*Конкурировать*
Предположения	*Согласоваться (с чем?)*
<u>Корреляция</u>	*Выдвинуть (что?)*

Работа с текстами лекций по специальности

Задание 2. Прочитайте текст. Разделите текст на абзацы.

Текст
Статистические гипотезы. Ошибки первого и второго рода

Второй важнейшей задачей математической статистики после статистического оценивания параметров распределения является *статистическая проверка гипотез*. Методы математической статистики позволяют проверить предположения о законе распределения некоторой случайной величины (генеральной совокупности), о значениях параметров этого закона (например, математического ожидания или дисперсии), о наличии корреляционной зависимости между случайными величинами, определенными на множестве объектов одной и той же генеральной совокупности. Пусть по некоторым данным имеются основания выдвинуть предположения о законе распределения или о параметре закона распределения случайной величины. Задача заключается в том, чтобы подтвердить или опровергнуть это предположение, используя выборочные (экспериментальные) данные. *Статистическая гипотеза* – это любое предположение о виде неизвестного закона распределения или о параметрах известных распределений. Гипотезы о значениях параметров распределений или о сравнительной величине параметров двух распределений называются *параметрическими* гипотезами. Гипотезы о виде распределения называются *непараметрическими* гипотезами. *Проверить статистическую гипотезу* – значит проверить, согласуются ли выборочные данные с выдвинутой гипотезой. Примеры статистических гипотез:

1) генеральная совокупность подчиняется нормальному закону распределения;

2) математические ожидания двух нормальных совокупностей равны между собой. Первая гипотеза является непараметрической, а вторая - параметрической. Вместе с выдвигаемой основной гипотезой рассматривают и противоречащую ей гипотезу. В том случае, если выдвинутая гипотеза будет отвергнута, противоречащая гипотеза окажется справедливой. *Нулевая* (*основная или проверяемая*) гипотеза – это выдвинутая гипотеза, которая обозначается H_0. Конкурирующая (альтернативная) гипотеза – это гипотеза H_1, которая противоречит нулевой гипотезе.

Задание 3. Составьте 8-10 вопросов к тексту из задания 2.

____
____
____
____
____
____
____
____
____
____
____
____
____
____
____
____
____

Задание 4. Прослушайте текст. Вставьте пропущенные слова.

Текст

Гипотезы о значениях числовых характеристик

Гипотезы о ________________________ среднего значения m и ____________________ σ^2 определенным числам m_0 и σ_0^2. Они возникают, например, при ________________ качества функционирования ______________________ устройств. Если m_0 - номинальное значение измеряемого ____________________ и $m \neq m_0$, то это означает, что прибор дает _________________________ ошибку. Точность прибора определяется ________________________ σ_0 и, если $\sigma > \sigma_0$, то это означает, что качество прибора не отвечает ________________________ требованиям.

Задание 5. Внимательно прочитайте части текста. Расположите их в правильном порядке.

<u>1</u>	*Сравнение средних* — это проверка гипотезы о равенстве средних значений совокупностей, из которых получены выборки.
	Например, такие вопросы возникают при исследовании надежности технических систем, где результаты сравниваются с предыдущими измерениями; при контроле качества изделий, изготовленных на разных предприятиях; в финансах – при сравнении уровня доходности различных активов.
	На практике часто встречаются ситуации, когда среднее значение данных одного эксперимента отличается от среднего значения данных другого, хотя условия эксперимента являются схожими. Тогда возникает вопрос, можно ли считать это расхождение незначимым, т.е. чисто случайным, или оно вызвано существенным различием двух генеральных совокупностей.
	В зависимости от априорной информации о дисперсиях можно выделить несколько характерных случаев.

Задание 6. Составьте несколько вопросов к тексту из задания 5.

6. Звучащие материалы.

Тема 4.

Текст
Статистические оценки параметров распределения
(часть 1, задание 3)

Основная цель статистического анализа состоит в исследовании свойств случайных величин. Экспериментальной основой таких исследований являются результаты многократных наблюдений случайной величины.

При статистическом анализе в каждом конкретном случае возможно решение ряда задач, каждая из которых в общем случае включает два этапа: обработку выборки и принятие решения. Поскольку существуют условия неопределенности из-за конечного объема и случайного содержания выборки, принимаемое решение носит вероятностный характер. Задачи статистического анализа делят на три обширных класса:

1. Статистическое оценивание;
2. Проверка статистических гипотез;
3. Построение статистических зависимостей.

Эти задачи существенно различаются как по внутреннему содержанию, так и по методам их решения. В этом разделе уделяется внимание первому классу задач: статистическому оцениванию. Задачи оценивания направлены на вычисление количественных характеристик случайных величин по конечной выборке. При этом различают правила оценивания, называемые *методами получения оценок*.

Применение методов математической статистики для обработки результатов измерений базируется на том, что многократные измерения соответствуют *выборочному методу*. Сущность *выборочного метода* заключается в следующем: имеется некоторая большая совокупность объектов, называемая *генеральной совокупностью*. Из этой совокупности случайным образом извлекается ограниченное число объектов, образующих *выборку*. Выборка подвергается детальному исследованию, по результатам которого требуется определить свойства (характеристики) генеральной совокупности.

Статистические оценки параметров распределения (часть 2, задание 4)

В распоряжении исследователя имеются лишь данные выборки, например значения количественного признака $x_1, x_2, ..., x_n$, полученные в результате n наблюдений. Через эти данные и выражают оцениваемый параметр Θ. Поскольку истинное значение искомой величины получить принципиально невозможно, то в качестве его принимают некоторое найденное по выборке значение $\tilde{\Theta}$. Полученное по результатам эксперимента приближенное значение $\tilde{\Theta}$ называют *оценкой истинного значения* параметра Θ, которую определяют как функцию результатов наблюдений объема n:

$$\tilde{\Theta} = f(x_1, x_2, ..., x_n).$$

Чтобы оценки в большинстве опытов были максимально близки к истинному значению измеряемой величины, к ним предъявляют следующие требования.

1. *Несмещенность.* Оценку называют несмещенной, если при любом объеме выборки n математическое ожидание оценки равно истинному значению искомого параметра:

$$M\left[\overline{\Theta}\right] = \Theta.$$

Во многих случаях смещение убывает с возрастанием объема выборки.

2. *Эффективность.* Эффективной называю статистическую оценку, которая при постоянном объеме выборки n имеет наименьшую возможную дисперсию.

3. *Состоятельность.* Оценка является *состоятельной*, если при $n \to \infty$ стремится по вероятности к оцениваемому параметру. Например, если дисперсия несмещенной оценки при $n \to \infty$ стремится к нулю, то такая оценка оказывается и состоятельной.

Состоятельность оценки означает, что чем больше объем выборки n, тем больше вероятность того, что ошибка не превысит сколь угодно малого положительного числа ε.

$$\lim_{n\to\infty} P\left(\left|\tilde{\Theta}_n - \Theta\right| < \varepsilon\right) = 1.$$

4. *Устойчивость* характеризует чувствительность оценки к аномальным результатам (промахам и выбросам).

Методы получения точечных оценок (задание 8)

На практике при оценке параметров не всегда удается удовлетворить одновременно требование состоятельности, несмещенности и эффективности. Однако выбору оценки всегда должно предшествовать критическое рассмотрение ее со всех точек зрения.

Количественные значения оценок искомых величин в общем случае являются функциями элементов выборки.

Для нахождения алгоритмов вычисления оценок наиболее широкое распространение получили следующие методы: максимального правдоподобия, моментов, порядковых характеристик и др.

Тема 5.

Гипотезы о значениях числовых характеристик (задание 4)

Гипотезы о равенстве среднего значения m и дисперсии σ^2 определенным числам m_0 и σ_0^2. Они возникают, например, при проверке качества функционирования измерительных устройств. Если m_0 - номинальное значение измеряемого параметра и $m \neq m_0$, то это означает, что прибор дает систематическую ошибку. Точность прибора определяется значением σ_0 и, если $\sigma > \sigma_0$, то это означает, что качество прибора не отвечает стандартным требованиям.

7. Оригинальные примеры для самостоятельного чтения

Тема 1. Основные понятия теории вероятности

Текст 1. Предмет и задачи теории вероятностей

Теория вероятностей - математическая наука, позволяющая по вероятностям одних случайных событий находить вероятности других случайных событий, связанных каким-либо образом между собой.

В этом определении есть целый ряд понятий: случайное событие, вероятность случайного события, связь между случайными событиями. Все эти понятия нуждаются в определении и разъяснении. В усвоении этого круга вопросов и состоит первое знакомство с теорией вероятностей.

Теория вероятностей изучает свойства массовых случайных событий, способных многократно повторяться при воспроизведении определенного комплекса условий. Основное свойство любого случайного события независимо от его природы - вероятность его осуществления.

Предметом теории вероятностей является изучение вероятностных закономерностей массовых однородных случайных событий.

Теория вероятностей вначале развивалась как прикладная дисциплина. В связи с этим ее понятия и выводы имели окраску тех областей знаний, в которых они были получены. Лишь постепенно выкристаллизовалось то общее, что присуще вероятностным схемам независимо от области их приложения массовые случайные события, действия над ними и их вероятности, случайные величины и их числовые характеристики. Большой вклад в развитие теории вероятностей внесли русские ученые. Впервые законченную систему аксиом сформулировал в 1936г. советский математик академик А.Н. Колмогоров в своей книге «Основные понятия теории вероятностей». Практические приложения способствовали зарождению теории вероятностей, они же питают ее развитие как науки, приводя к появлению все новых ее ветвей и разделов.

Одной из важнейших сфер приложения теории вероятностей является экономика. В настоящее время трудно себе представить исследование и прогнозирование экономических явлений без использования эконометрического моделирования, регрессионного анализа, трендовых и сглаживающих моделей и других методов, опирающихся на теорию вероятностей.

Методы теории вероятностей широко применяются в различных отраслях естествознания и техники: в теории надёжности, теории массового обслуживания, в теоретической физике, геодезии, астрономии, теории стрельбы, теории ошибок наблюдений, теории автоматического управления, общей теории связи и во многих других теоретических и прикладных науках. Все это предопределяет необходимость овладения методами теории вероятностей и математической статистики как инструмента статистического анализа и прогнозирования экономических и технологических явлений и процессов. [3]

Вопросы и задания:

1. Дайте определения понятию *теория вероятностей.*
2. Что изучает теория вероятностей?
3. Кто из ученых внес вклад в развитие теории вероятностей как науки?
4. Где находят применения достижения теории вероятности?

Текст 2. Понятие вероятности события

Сравнивать случайные события (то есть события, которые могут наступить или не наступить) естественно по степени возможности их наступления. С этой целью вводится числовая характеристика этой степени возможности (случайности), называемая вероятностью события. Для события A вероятность принято обозначать $P(A)$. Существует несколько подходов, поясняющих понятие вероятности. В каждом из этих подходов указываются правила, по которым случайному событию ставится в соответствие положительное число, объективно характеризующее степень возможности появления этого события.

Большинство определений математической вероятности может быть разделено на следующие: статистическое, классическое и геометрическое.

Многочисленными наблюдениями над самыми разнообразными случайными событиями установлен следующий достоверный факт: если над одним и тем же случайным событием в одних и тех же условиях проводить много серий из большого числа испытаний каждая, то наблюдаемая в каждой такой серии частота события будет колебаться от серии к серии в сравнительно узких пределах, будет, как говорят в теории вероятностей, “устойчивой”. При этом пределы, в которых будет колебаться устойчивая частота случайного события, будут тем теснее, чем больше число испытаний в каждой серии. Это свидетельствует о наличии статистической закономерности в изучаемом явлении.

Пусть в одних и тех же условиях проведена серия из $n*$ испытаний, в каждом из которых могло появиться или не появиться интересующее нас событие *А*. Пусть событие *А* появилось при этом в m испытаниях.

Относительной частотой Р(А)* события *А* в данной серии испытаний называется отношение $m*$ (числа испытаний, в которых появилось событие *А)* к $n*$ (общему числу проведенных испытаний), то есть

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

Из данного определения следует, что относительная частота случайного события всегда заключена между нулем и единицей:

$$0 \le P*(A) \le 1$$

Статистической вероятностью Р(А) события *А* называется предел, к которому стремится относительная частота *Р*(А)* при неограниченном увеличении числа испытаний, то есть

$$P(A) = \lim_{n\to\infty} P*(A) = \lim_{n\to\infty} \frac{m*}{n*}$$

При больших n статистическое определение позволяет в приблизительных расчетах относительную частоту *Р*(А)* принимать за вероятность события *А*. Недостатком этого определения вероятности является необходимость проведения большого числа опытов в одинаковых условиях. [3]

Вопросы и задания:

1. Как можно сравнить между собой случайные величины?
2. На какие части можно разделить большое разнообразие определений понятия «*вероятность*»?
3. В чем состоит недостаток статистического определения вероятности?

Текст 3. Классическое определение вероятности

Классическое определение вероятности сводит понятие вероятности к понятию равновероятности (равновозможности) событий, которое считается основным и не подлежит формальному определению.

Это определение применимо в случаях, когда удается выделить полную группу несовместных и равновероятных событий - элементарных исходов. Для примера рассмотрим урну с шарами.

Пусть в урне содержится 7 одинаковых, тщательно перемешаных шаров, причем 2 из них - красных, 1 - синий и 4 - белых. Испытание будет заключаться в том, что из урны наудачу берется один шар. Каждое событие, которое может наступить в проводимом испытании, является элементарным исходом. В данном примере семь элементарных исходов, которые мы обозначим $E_1, E_2, ..., E_7$. Исходы E_1, E_2 – появление красного шара, E_3 - появление синего шара, E_4, E_5, E_6, E_7 - появление белого шара. В нашем примере события E_1, E_2, ..., E_7 – попарно несовместны. Кроме того, они еще и равновозможны в данном испытании. Пусть событие A заключается в том, что на удачу взятый из урны шар оказался цветным (красным или синим). Те элементарные исходы, при которых интересующее нас событие A наступает, называют *исходами, благоприятствующими* событию A. В нашем примере исходами, благоприятствующими событию A, являются исходы E_1, E_2 и E_3. Разумно в качестве меры возможности появления события A, то есть вероятности $P(A)$ принять число, равное отношению исходов, благоприятствующих наступлению события A, числу всех возможных исходов. В нашем примере

$$P(A) = \frac{3}{7}.$$

Рассмотренный пример привел нас к определению вероятности, которое принято называть *классическим*.

Вероятностью события A называют отношение числа m благоприятствующих этому событию исходов к общему числу n всех элементарных исходов:

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

Классическое определение вероятности служит хорошей математической моделью тех случайных экспериментов, число исходов которых конечно, а сами исходы – равновозможны [3].

Вопросы и задания:

1. Какие события называются *равновозможными*; *несовместными*?

2. Что такое *благоприятный исход* события?

3. Прочитайте классическую формулу для определения вероятности события.

Текст 4. Основные формулы комбинаторики

При подсчете числа элементарных исходов, используется комбинаторика. *Комбинаторика* изучает количество комбинаций, подчиненных определенным условиям, которое можно составить из элементов, заданного конечного множества. Приведем наиболее употребительные из них.

Перестановками называются комбинации, состоящие из одних и тех же n различных элементов и отличающиеся только порядком их расположения.

$$P_n = n!,$$

где $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ (n! – читается «n – *факториал*»).

Размещениями называется комбинации, составленные из n различных элементов по m элементов, которые отличаются либо составом, либо их порядком элементов.

$$A_n^m = \frac{n!}{(n-m)!} = n(n-1)(n-2)\ldots(n-m+1).$$

Сочетаниями называется комбинации, составленные из n различных элементов по m элементов, которые отличаются хотя бы одним элементом. Порядок размещения элементов значения не имеет.

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

Подчеркнем, что числа размещений, перестановок и сочетаний связаны равенством

$$A_n^m = P_m C_n^m.$$

При решении задач комбинаторики используются следующие правила:

Правило суммы. Если некоторый объект *A* может быть выбран из совокупности объектов *m* способами, а другой объект *B* может быть выбран *n* способами, то выбрать либо *A,* либо *B* можно $m + n$ способами.

Правило произведения. Если объект *A* можно выбрать из совокупности объектов *m* способами и после каждого такого выбора объект *B* можно выбрать *n* способами, то пара объектов *(A, B)* в указанном порядке может быть выбрана $m \cdot n$ способами. [1]

Вопросы и задания:

1. Что изучает наука комбинаторика?
2. Какая комбинация называется *перестановкой*?
3. Как рассчитать число *размещений?*
4. Прочитайте формулу для расчета с*очетаниями*

Текст 5. Теоремы сложения и умножения вероятностей

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

Пример 1. Произведены два выстрела. Событие А – попадание при первом выстреле, *событие B* – попадание при втором выстреле. Тогда событие *A + B* – попадание при первом выстреле, или при втором, или в обоих выстрелах.

Теорема. Вероятность появления хотя бы одного из двух совместных событий, безразлично какого, равна сумме вероятностей этих событий без учета вероятности их совместного появления:

$$P(A+B)=P(A)+P(B)-P(AB). \quad (7.1)$$

Если события несовместны, то вероятность их одновременного наступления исключена. Тогда слагаемое $P(AB)$ из формулы (7.1) будет равно нулю, следовательно, формула (7.1) примет вид:

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

Теорема сложения вероятности может быть обобщена на любое конечное число совместных событий. Например, для трех совместных событий

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

Сумма вероятностей событий A_1, A_2, …, A_n, образующих полную группу, равна единице:

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

Произведением двух событий А и В называют событие *АВ*, состоящее в совместном появлении (совмещении) этих событий. Например, если *А* – деталь годная, *В* – деталь окрашенная, то *АВ* – деталь годна и окрашена.

Произведением нескольких событий называют событие, состоящее в совместном появлении всех этих событий. Например, если *А, В, С* – появление «герба» соответственно в первом, втором и третьем бросаниях монеты, то *АВС* – выпадение «герба» во всех трех испытаниях.

Условной вероятностью $P_A(B)$ называют вероятность событий *В, вычисленную в предположении, что событие А уже наступило.*

Теорема. Вероятность совместного появления двух событий равна произведению вероятности одного из них на условную вероятность другого, вычисленную в предположении, что первое событие уже наступило:

$$P(AB)=P(A)\cdot P_A(B). \tag{6.2}$$

Заметим, что порядок, в котором расположены события, может быть выбран любым, т.е. безразлично какое событие считать первым, вторым и т.д.

Для независимых событий теорема умножения (6.2) имеет вид

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

т.е. вероятность совместного появления двух независимых событий равна произведению вероятностей этих событий.

На практике о независимости событий заключают по смыслу задачи. Например, вероятности поражения цели каждым из двух орудий не зависят от того, поразило ли цель другое орудие, поэтому события «первое орудие поразило цель» и «второе орудие поразило цель» независимы. [1]

Вопросы и задания:

1. Поясните, что значит сумма событий, произведение событий. Приведите пример.
2. Сформулируйте *теорему сложения вероятностей* для совместных событий.
3. Чему равна сумма вероятностей *полной группы событий*?
4. Сформулируйте *теорему умножения вероятностей* для зависимых событий.
5. Как определить *зависимость* событий?

Тема 2. Случайные величины

Текст 1. Описание случайных величин

Случайными называют величины, принимающие в результате опыта со случайным исходом то или иное значение, которое принципиально не может быть определено заранее. Случайная величина обладает некоторым набором возможных значений, но в итоге каждого отдельного исследования принимает лишь одно из них. Различают непрерывные и дискретные случайные величины. *Дискретная* случайная величина принимает на числовой оси конечный набор определенных значений x_i, с соответствующими вероятностями $P(x_i) = p_i$, при этом сумма вероятностей равна единице, так как ансамбль возможных исходов представляет собой полную группу событий. Если случайная величина может принимать в некотором интервале всевозможные значения, которые невозможно отделить друг от друга промежутками, то она является *непрерывной*. Погрешности измерений также являются случайными величинами, так как каждый отдельный результат измерения, а, следовательно, и значение погрешности заранее предсказать невозможно.

Полной характеристикой случайной величины является ее закон распределения. Интегральный закон распределения (функция распределения) представляет собой вероятность того, что случайная величина принимает значения, меньшие заданного: $F(x) = P(x' < x)$. Функция распределения для любой случайной величины при возрастании аргумента является неубывающей функцией и изменяется от $F(-\infty) = 0$ как вероятность невозможного события до $F(\infty) = 1$ как вероятность достоверного события.

Дифференциальный закон распределения (плотность распределения) характеризует вероятность появления случайной величины в узком интервале и для непрерывной дифференцируемой функции $F(x)$ определяется производной $p(x) = \dfrac{dF(x)}{dx} = F'(x)$. Плотность распределения – функция неотрицательная, т.е. $p(x) \ge 0$, и удовлетворяет условию нормировки:

Вероятность	p_1	p_2	$\ldots$	p_i	$\ldots$	p_n

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

Плотность распределения дискретной случайной величины представляется рядом распределения в виде таблицы, в верхней строке которой перечислены значения случайной величины в порядке их возрастания, а в нижней – соответствующие им вероятности: [2].

Значение случайной величины	x_1	x_2	…	x_i	…	x_n

Вопросы и задания:

1. Приведите пример случайной величины.
2. Чем отличается дискретная случайная величина от непрерывной случайной величины?
3. Что называется законом распределения?
4. В каком виде можно представить плотность распределения случайной величины?

Текст 2. Числовые характеристики дискретных случайных величин

Как уже известно, закон распределения полностью характеризует случайную величину. Однако часто закон распределения неизвестен и приходится ограничиваться меньшими сведениями. Иногда даже выгоднее пользоваться числами, которые описывают случайную величину суммарно; такие числа называют *числовыми характеристиками случайной величины.* К числу важных числовых характеристик относится математическое ожидание.

Математическое ожидание, как будет показано далее, приблизительно равно среднему значению случайной величины. Для решения многих задач достаточно знать математическое ожидание. Например, если известно, что математическое ожидание числа выбиваемых очков у первого стрелка больше, чем у второго, то первый стрелок в среднем выбивает больше очков, чем второй, и, следовательно, стреляет лучше второго. Хотя математическое ожидание дает о случайной величине значительно меньше сведений, чем закон ее распределения, но для решения задач, подобных приведенной и многих других, знание математического ожидания оказывается достаточным.

Математическим ожиданием дискретной случайной величины называют сумму произведений всех ее возможных значений на их вероятности.

Пусть случайная величина X может принимать только значения x_1, x_2, ..., x_n, вероятности которых соответственно равны p_1, p_2, ..., p_n. Тогда математическое ожидание M(X) случайной величины X определяется равенством

$$\mathrm{M}(\mathrm{X}) = x_1 p_1 + x_2 p_2 + \ldots + x_\mathrm{n} p_\mathrm{n}.$$

Если дискретная случайная величина X принимает счетное множество возможных значений, то

$$M(X) = \sum_{i=1}^{\infty} x_i p_i \,,$$

причем математическое ожидание существует, если ряд в правой части равенства сходится абсолютно.

Замечание. Из определения следует, что математическое ожидание дискретной случайной величины есть неслучайна (постоянная) величина. Рекомендуем запомнить это утверждение, те как далее оно используется многократно.

Итак, математическое ожидание числа появлений события в одном испытании равно вероятности этого события. [4]

Вопросы и задания:

1. Что полностью характеризует случайную величину?
2. Что показывает *математическое ожидание* случайной величины?
3. Прочитайте формулы для вычисления математического ожидания дискретной случайной величины.
4. Является ли *математическое ожидание* величиной случайной?

Текст 3. Целесообразность введения числовой характеристики рассеяния случайной величины

Легко указать такие случайные величины, которые имеют одинаковые математические ожидания, но различные возможные значения. Рассмотрим, например, дискретные случайные величины X и Y, заданные следующими законами распределения:

X	-0,01	0,01
p	0,5	0,5

Y	-100	100
p	0,5	0,5

Найдем математические ожидания этих величин:

$$M(X) = -0{,}01 \cdot 0{,}5 + 0{,}01 \cdot 0{,}5 = 0,$$
$$M(Y) = -100 \cdot 0{,}5 + 100 \cdot 0{,}5 = 0.$$

Здесь математические ожидания обеих величин одинаковы, а возможные значения различны, причем X имеет возможные значения, близкие к математическому ожиданию, а Y – далекие от своего математического ожидания. Таким образом, зная лишь математическое ожидание случайной величины, еще нельзя судить ни о том, какие возможные значения она может принимать, ни о том, как они рассеяны вокруг математического ожидания. Другими словами, математическое ожидание полностью случайную величину не характеризует.

По этой причине наряду с математическим ожиданием вводят и другие числовые характеристики. Так, например, для того чтобы оценить, как рассеяны возможные значения случайной величины вокруг ее математического ожидания, пользуются, в частности, числовой характеристикой, которую называют *дисперсией*.

Прежде чем перейти к определению и свойствам дисперсии, введем понятие отклонения случайной величины от ее математического ожидания.

Пусть X – случайная величина и $M(X)$ – ее математическое ожидание. Рассмотрим в качестве нон случайной величины разность $X - M(X)$.

Отклонением называют разность между случайной величиной и ее математическим ожиданием.

Пусть закон распределения X известен:

X	x_1	x_2	...	x_n
p	p_1	p_2	...	p_n

Напишем закон распределения отклонения. Для того чтобы отклонение приняло значение $x_1 - M(X)$, достаточно, чтобы случайная величина приняла значение x_1. Вероятность же этого события равна p_1; следовательно, и вероятность того, что отклонение примет значение

$x_1 - M(X)$, также равна p_1. Аналогично обстоит дело и для остальных возможных значений отклонения.

Таким образом, отклонение имеет следующий закон распределения:

$X - M(X)$	$x_1 - M(X)$	$x_2 - M(X)$	...	$x_n - M(X)$
p	p_1	p_2	...	p_n

Приведем важное свойство отклонения.

Теорема. Математическое ожидание отклонения равно нулю:

$$M[X - M(X)] = 0.$$

Доказательство. Пользуясь свойствами математического ожидания (математическое ожидание разности равно разности математических ожиданий, математическое ожидание постоянной равно самой постоянной) и приняв во внимание, что *M(X)* – постоянная величина, имеем

$$M[X - M(X)] = M(X) - M[M(X)] = M(X) - M(X) = 0.$$ [4]

Вопросы и задания:

1. В каких случаях недостаточно знать математическое ожидание для характеристики случайной величины?
2. Что такое «*отклонение случайной величины*»?
3. Перечислите, используемые при доказательстве теоремы свойства математического ожидания.
4. Прочитайте формулу, используемую при доказательстве теоремы.

Текст 4. Дисперсия дискретной случайной величины

На практике часто требуется оценить рассеяние возможных значений случайной величины вокруг ее среднего значения. Например, в артиллерии важно знать, насколько кучно лягут снаряды вблизи цели, которая должна быть поражена.

На первый взгляд может показаться, что для оценки рассеяния проще всего вычислить все возможные значения отклонения случайной величины и затем найти их среднее значение. Однако такой путь ничего не даст, так как среднее значение отклонения, т. е. $M[X - M(X)]$, для любой случайной величины равно нулю. Это свойство объясняется тем, что одни возможные отклонения положительны, а другие – отрицательны; в результате их взаимного погашения среднее значение отклонения равно нулю. Эти соображения говорят о целесообразности заменить возможные отклонения их абсолютными значениями или их квадратами. Так и поступают на деле.

Правда, в случае, когда возможные отклонения заменяют их абсолютными значениями, приходится оперировать с абсолютными величинами, что приводит иногда к серьезным затруднениям. Поэтому чаще всего идут по другому пути, т. е. вычисляют среднее значение квадрата отклонения, которое и называют *дисперсией*.

Дисперсией (рассеянием) дискретной случайной величины называют математическое ожидание квадрата отклонения случайной величины от ее математического ожидания:

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

Пусть случайная величина задана законом распределения

X	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

Тогда квадрат отклонения имеет следующий закон распределения:

$[X - M(X)]^2$	$[x_1 - M(X)]^2$	$[x_2 - M(X)]^2$	⋯	$[x_n - M(X)]^2$
p	p_1	p_2	…	p_n

По определению дисперсии,

$$D(X) = M[X - M(X)]^2 =$$
$$= [x_1 - M(X)]^2 \cdot p_1 + [x_2 - M(X)]^2 \cdot p_2 + ... + [x_n - M(X)]^2 \cdot p_n$$

Таким образом, для того чтобы найти дисперсию, достаточно вычислить сумму произведений возможных значений квадрата отклонения на их вероятности.

Замечание. Из определения следует, что дисперсия дискретной случайной величины есть неслучайная (постоянная) величина. В дальнейшем читатель узнает, что дисперсия непрерывной случайной величины также есть постоянная величина. [4]

Вопросы и задание:

1. Какими числовыми характеристиками для описания случайной величины часто пользуются?
2. Какую числовую характеристику называют *дисперсией?*
3. Прочитайте формулу для вычисления *дисперсии.*
4. Является ли *дисперсия* величиной случайной?

Текст 5. Другие числовые характеристики случайных величины

Характеристики случайной величины, построенные на основании выборочных данных, называются *выборочными или точечными оценками.* Свойства случайной величины могут характеризоваться различными начальными и центральными моментами, вычисляемыми в случае дискретной случайной величины по следующим формулам.

Начальный момент порядка k:

$$a_k = \sum_i \bar{\bar{x}}_i^{\,p} p_i = \sum_{j=1}^{m} x_j^k p_j$$

Центральный момент порядка k:

$$\mu_k = \sum_i (\bar{\bar{x}}_i - \alpha_1)^k p_i = \sum_{j=1}^{m} (x_j - \alpha_1)^k p_j$$

Важнейшие из них – *математическое ожидание $M(X)=m_X$* и дисперсия $D(X)=\sigma^2(X)$, где через σ обозначено *среднеквадратическое отклонение*, являются частными случаями моментов:

$$\bar{x} = \alpha_1, \ \ \bar{D} = \mu_2, \ \ \bar{\sigma} = \sqrt{\bar{D}}$$

Выделяют также несмещенную выборочную дисперсию:

$$S^2 = \frac{n}{n-1}\bar{D}$$

Если *выборочное* математическое ожидание случайной величины дает нам «ее среднее значение» или точку на координатной прямой, «вокруг которой разбросаны» значения рассматриваемой случайной величины, то *выборочная* дисперсия характеризует «степень разброса» значений.

Используются также оценки коэффициента асимметрии:

$$A = \frac{\mu_3}{\sigma^3}$$

и коэффициента эксцесса (островершинности):

$$E = \frac{\mu_4}{\sigma^4} - 3$$

как степени отклонения полигона частот от плотности нормального распределения непрерывной случайной величины, для которой они равны нулю. [2]

Вопросы и задания:

1. Прочитайте формулу для расчета *начального момента.*
2. Прочитайте формулу для расчета *начального момента.*
3. Чем являются *математическое ожидание* и *дисперсия*?
4. Приведите примеры других *числовых характеристик* случайной величины.

Тема 3. Элементы математической статистики

Текст 1. Статистическое распределение выборки

Пусть из генеральной совокупности извлечена выборка, причем x_1 наблюдалось n_1 раз, $x_i - n_i$ раз, $x_k - n_k$ раз и $\sum n_i = n$ – объем выборки.

Совокупность полученных результатов исследования выборки называется *статистической совокупностью* или *простым статистическим рядом*. Этот ряд может состоять, например, из значений результатов измерений диаметра валов привода, температуры прокатки подшипниковой стали или числа отказов электрорадиоаппаратуры за 1000 часов работы. Названные признаки являются случайными величинами. Их обозначают заглавными латинскими буквами X, Y, ... Ряд измерений объема n состоит из n значений, которые обозначаются малыми латинскими буквами, снабженными индексом, указывающим порядковый номер измерения: $x_1, x_2, \ldots, x_n$.

Наблюдаемые значения x_i называют *вариантами,* а последовательность вариант, записанных в возрастающем порядке, – *вариационным рядом.* Если закон распределения измеряемой случайной величины не известен заранее, то первым шагом к осмысливанию ряда наблюдений является построение *вариационного ряда.*

Различают дискретные и интервальные вариационные ряды. Для представления дискретных случайных величин используют дискретные вариационные ряды, для представления непрерывных случайных величин - в основном интервальные.

Число повторений значения x_i называют *частотам* n_i, а их отношения к объему выборки $\frac{n_i}{n} = W_i$ – *относительными частотами*. [1]

Вопросы и задания:

1. Что такое *варианта*?
2. Что называется *вариационным рядом*?
3. Чему равен *объем выборки*?
4. Как называется число повторений значения x_i?
5. Прочитайте формулу для расчета *относительной частоты*.

Текст 2. Полигон и гистограмма

Для наглядности строят различные графики статистического распределения и, в частности, полигон и гистограмму.

Полигоном частот называют ломаную, отрезки соединяют точки $(x_1;n_1)$, $(x_2;\ n_2)$, ..., $(x_k;n_k)$. Для построения полигона частот на оси абсцисс откладываются варианты x_i, а на оси ординат – соответствующий им частоты n_i. Точки $(x_i;\ n_i)$ соединяют отрезками прямых и получают полигон частот.

Полигоном относительных частот называют ломанную, отрезки которой соединяют точки $(x_1;\ W_1)$, $(x_2;\ W_2)$, ..., $(x_k;\ W_k)$. Для построения полигона относительных частот на оси абсцисс откладывают варианты x_i, а на оси ординат – соответствующие им относительные частоты W_i. Точки $(x_i;\ W_i)$ соединяют отрезками прямых, получают полигон относительных частот.

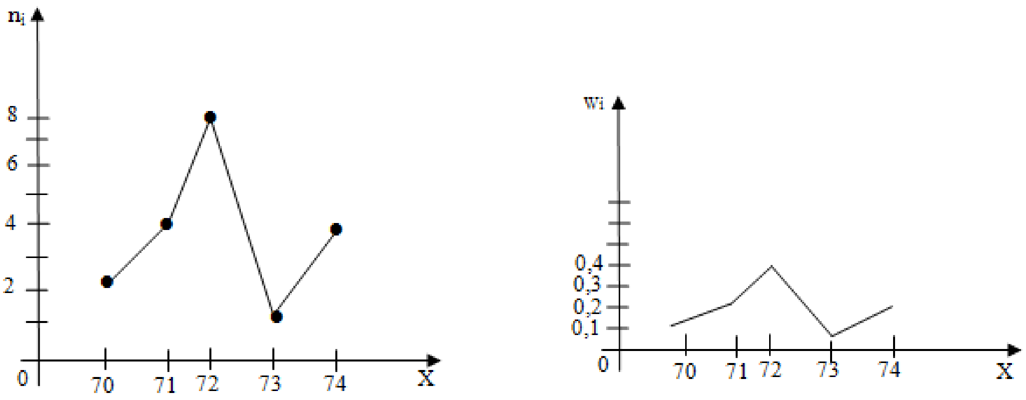


Рисунок 7.1 – Полигон частот и полигон относительных частот

В случае *непрерывного признака* целесообразно строить гистограмму, для чего интервал, в котором заключены все наблюдаемые значения признака, разбивают на несколько частичных интервалов длиной h и находят для каждого частичного интервала n_i – сумму частот вариант, попавших в i-й интервал.

Гистограммой частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению n_i/h (плотность частоты).

Для построения гистограммы частот на оси абсцисс откладывают частичные интервалы, а над ними проводят параллельные оси абсцисс на расстоянии n_i/h.

Площадь i-го частичного прямоугольника равна $h \cdot n_i / h = n_i$ – сумме частот вариант i-го интервала; следовательно, *площадь гистограммы частот равна сумме всех частот т. е. объему выборки*.

На рисунке 7.2 изображена гистограмма частот распределения.

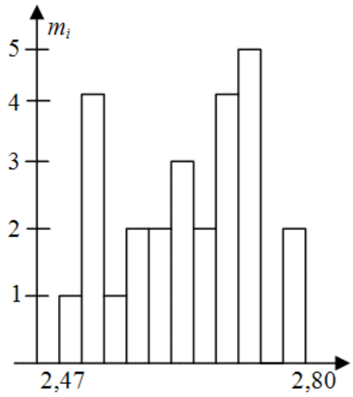


Рисунок 7.2 – Гистограмма частот

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению W_i / h (плотность относительной частоты).

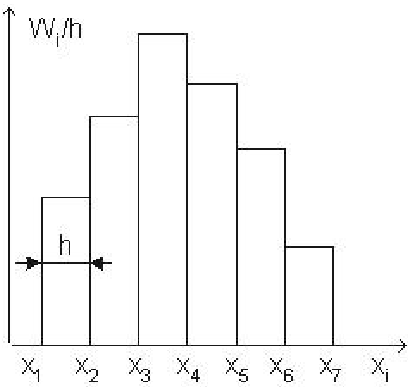


Рисунок 7.3 – Гистограмма относительных частот

Для построения гистограммы относительных частот на оси абсцисс откладывают частичные интервалы, а над ними проводят отрезки, параллельные оси абсцисс на расстоянии W_i / h. Площадь i-го частичного прямоугольника равна $h \cdot W_i / h = W_i$ – относительной частоте вариант, попавших в i-й интервал. Следовательно, *площадь гистограммы относительных частот равна сумме всех относительных частот, т. е. единице.* [1]

Вопросы:

1. Для чего строят *полигоны* и *гистограммы*?
2. Что такое *полигон*?
3. Чем *полигон частот* отличается от *полигона относительных частот*?
4. Чем *гистограмма частот* отличается от *гистограммы относительных частот*?
5. Чему равна площадь *гистограммы частот*?
6. Чему равна площадь *гистограммы относительных частот*?

Текст 3. Предмет и задачи математической статистики

Как известно, теория вероятностей изучает математические модели случайных явлений, при этом сама математическая модель считается заданной. Так, например, если изучается некоторое случайное событие A, то известно $P(A)$. Если же речь идёт о случайной величине X ,то известен закон распределения вероятностей в какой-либо форме и, как следствие, числовые характеристики исследуемой случайной величины. В практических задачах эти характеристики, как правило, неизвестны, но имеются некоторые экспериментальные данные о событии или случайной величине. Требуется на основании этих данных построить подходящую вероятностную модель изучаемого явления, то есть приближенно оценить неизвестные закон распределения и числовые характеристики исследуемой случайной величины на основе экспериментальных данных. Это и является задачей математической статистики.

Математической статистикой называется наука, занимающаяся методами обработки экспериментальных данных, полученных в результате наблюдений над случайными явлениями. При этом решаются следующие в порядке сложности и важности задачи:

- <u>описание явлений</u>, то есть упорядочение поступившего статистического материала, представление его в наиболее удобном для обозрения и анализа виде (таблицы, графики);
- <u>анализ и прогноз</u>, то есть приближенная оценка на основании статистических данных интересующих нас характеристик, например, математического ожидания и дисперсии наблюдаемой случайной величины и определение погрешности этих оценок;
- <u>выработка оптимальных решений</u>, то есть, например, определение числа опытов n, достаточного для того, чтобы ошибка от замены теоретических числовых характеристик их экспериментальными

оценками была не более заданной. В связи с этим возникает задача проверки правдоподобия гипотез о параметрах распределения и о законах распределения случайной величины, решением которой является возможность сделать один из выводов:

- отбросить гипотезу, как противоречащую опытным данным;
- принять гипотезу, считать ее приемлемой.

Математическая статистика, опираясь на размышляющий, оценивающий, сопоставляющий человеческий разум, помогает экспериментатору лучше разобраться в опытных данных, полученных в результате наблюдений над случайными явлениями; оценить, значимы или не значимы наблюдаемые факты; принять или отбросить те или иные гипотезы о природе случайных явлений. [3]

Вопросы:

1. Что такое *математическая статистика*?
2. Какие задачи решает *математическая статистика*?

Текст 4. Построение гистограммы и эмпирической функции распределения по выборке

Полный набор всех возможных N значений дискретной случайной величины X называют *генеральной совокупностью.* Однако в реальных условиях нельзя рассчитывать на такую подробную информацию. Часть генеральной совокупности из n элементов, отобранных случайным образом, называется *выборкой,* при этом число n называют *объемом выборки.* Различают выборки малого объема ($n < 30$) и большого объема ($n > 30$).

Вначале на основе результатов эксперимента строят *простой статистический ряд* – таблицу, состоящую из двух строк, в первой - порядковый номер измерения, во второй – его результат:

i	1	2	...	n
x_i	x_1	x_2	...	x_n

Для визуальной оценки распределения случайной величины производят группировку данных. Вначале x_k располагают в порядке возрастания, затем интервал наблюдаемых значений случайной величины разбивают на k последовательных непересекающихся частичных интервалов $\tilde{x}_0 \div \tilde{x}_1$, $\tilde{x}_1 \div \tilde{x}_2$, … $\tilde{x}_j \div \tilde{x}_{j+1}$, … $\tilde{x}_{k-1} \div \tilde{x}_k$, *далее* подсчитывают числа n_j – количество $x_к$, попавших в j-тый интервал. Полученный таким образом *группированный статистический ряд* отражают таблицами вида:

$\tilde{x}_j \div \tilde{x}_{j+1}$	$\tilde{x}_0 \div \tilde{x}_1$	$\tilde{x}_1 \div \tilde{x}_2$	...	$\tilde{x}_{k-1} \div \tilde{x}_k$
n_j	n_1	n_2	...	n_k

или, подсчитывая частоты $p_j = \dfrac{n_j}{n}$

$\tilde{x}_j \div \tilde{x}_{j+1}$	$\tilde{x}_0 \div \tilde{x}_1$	$\tilde{x}_1 \div \tilde{x}_2$	...	$\tilde{x}_{k-1} \div \tilde{x}_k$
p_j	p_1	p_2	...	p_k

или, определяя середину j - ого интервала $\bar{\bar{x}}_j = \tilde{x}_j + 0.5\Delta_j$, где $\Delta_j = \tilde{x}_{j+1} - \tilde{x}_j$ – длина j- ого интервала, получим ряд распределении в виде

$\bar{\bar{x}}_j$	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_2$	...	$\bar{\bar{x}}_k$
p_j	p_1	p_2	...	p_k

При этом частоты p_i удовлетворяют условию $\sum_{i=1}^{k} p_i = 1$. Деля частоту p_j на длину соответствующего интервала Δ_j, получим таблицу плотностей частоты f_j. Откладывая по оси абсцисс интервалы $\tilde{x}_j \div \tilde{x}_{j+1}$ и надстраивая на каждом интервале, как на основании, прямоугольник высотой f_j, то есть площадью p_j, получим ступенчатую фигуру - *гистограмму* частот - статистический аналог кривой плотности распределения. Еще более точной оценкой кривой плотности распределения является *полигон* частот - ломаная, отрезки которой соединяют точки $(\bar{\bar{x}}_j, f_j)$. В итоге ряд распределения можно представить таблицей вида

$\bar{\bar{x}}_j$	$\bar{\bar{x}}_1$	$\bar{\bar{x}}_2$	...	$\bar{\bar{x}}_k$
f_j	f_1	f_2	...	f_k

Иногда обходятся без группировки, но, располагая x_i в порядке возрастания, затем подсчитывают n_i – число повторов каждого значения случайной величины x_i и определяют в итоге частоту $p_i = \dfrac{n_i}{n}$, что приводит к ряду распределения вида

x_i	x_1	x_2	…	x_m
p_i	p_1	p_2	…	p_m

Здесь m – число разных значений наблюдаемой случайной величины.

Другим способом графического представления эмпирического закона распределения является *эмпирическая функция распределения* - оценка функции распределения дискретной случайной величины X, начисляемая по формуле

$$F(x) = \sum_{x_i < x} p_i$$

и являющаяся разрывной ступенчатой, равной нулю левее наименьшего наблюдаемого значения, испытывающей скачок величиной p_j при переходе через левую границу i - ого интервала и в итоге достигающей единицы правее наибольшего наблюдаемого значения. [3]

Вопросы и задания:

1. Дайте определение понятию *генеральная совокупность*.
2. Как построить *простой статистический ряд* распределения?
3. Какому условию должны удовлетворять частоты p_i?
4. Чем отличается *полигон* от *гистограммы?*

Тема 4. Основы теории оценивания

Текст 1. Статистические оценки параметров распределения

Пусть требуется изучить количественный признак генеральной совокупности. Допустим, что из теоретических соображений удалось установить, какое именно распределение имеет признак. Естественно возникает задача оценки параметров, которыми определяется это распределение. Например, если наперед известно, что изучаемый признак распределен в генеральной совокупности нормально, то необходимо оценить (приближенно найти) математическое ожидание и среднее квадратическое отклонение, так как эти два параметра полностью определяют нормальное распределение; если же есть основания считать, что признак имеет, например, распределение Пуассона, то необходимо оценить параметр λ, которым это распределение определяется.

Обычно в распоряжении исследователя имеются лишь данные выборки, например значения количественного признака $x_1, x_2, \ldots, x_n$, полученные в результате n наблюдении (здесь и далее наблюдения предполагаются независимыми). Через эти данные и выражают оцениваемый параметр. Рассматривая $x_1, x_2, \ldots, x_n$ как независимые случайные величины $X_1, X_2, \ldots, X_n$, можно сказать, что найти статистическую оценку неизвестного параметра теоретического распределения – это значит найти функцию от наблюдаемых случайных величин, которая и дает приниженное значение оцениваемого параметра. Например, для оценки математического ожидания нормального распределения служит функция (среднее арифметическое наблюдаемых значений признака)

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

Итак, статистической оценкой неизвестного параметра теоретического распределения называют функцию наблюдаемых случайных величин.[4]

Вопросы и задания:

1. Какие задачи решает теория оценивания?

2. На основании каких данных выражают оцениваемый параметр?

Текст 2. Точность оценки, доверительная вероятность (надежность). Доверительный интервал

Точечной называют оценку, которая определяется одним числом. Все оценки, рассмотренные выше, – точечные. При выборке малого объема точечная оценка может значительно отличаться от оцениваемого параметра, т.е. приводить к грубым ошибкам. По этой причине при небольшом объеме выборки следует пользоваться интервальными оценками.

Интервальной называют оценку, которая определяется двумя числами – концами интервала. Интервальные оценки позволяют установить точность и надежность оценок.

Пусть найденная по данным выборки статистическая характеристика Θ^* служит оценкой неизвестного параметра Θ. Будем считать Θ постоянным числом (Θ может быть и случайной величиной). Ясно, что Θ^* тем точнее определяет параметр Θ, чем меньше абсолютная величина разности $|\Theta - \Theta^*|$. Другими словами, если $\delta > 0$ и $|\Theta - \Theta^*| < \delta$, то чем меньше δ, тем оценка точнее. Таким образом, положительное число δ характеризует *точность оценки.*

Однако статистические методы не позволяют категорически утверждать, что оценка Θ^* удовлетворяет неравенству $|\Theta - \Theta^*| < \delta$; можно лишь говорить о вероятности γ, с которой это неравенство осуществляется. *Надежностью (доверительной вероятностью)* оценки Θ^* называют вероятность γ, с которой осуществится неравенство $|\Theta - \Theta^*| < \delta$. Обычно надежность оценки задается наперед, причем в качестве γ берут число, близкое к единице. Наиболее часто задают надежность, равную 0,95; 0,99 и 0,999.

Пусть вероятность того, что $|\Theta - \Theta^*| < \delta$, равна γ:

$$P\left[|\Theta - \Theta^*| < \delta\right] = \gamma .$$

Заменив неравенство $|\Theta - \Theta^*| < \delta$ равносильным ему двойным неравенством $-\delta < \Theta - \Theta^* < \delta$, или $\Theta^* - \delta < \Theta < \Theta^* + \delta$, имеем

$$P\left[\Theta^* - \delta < \Theta < \Theta^* + \delta\right] = \gamma .$$

Это соотношение следует понимать так: вероятность того, интервал $(\Theta^*-\delta, \Theta^*+\delta)$ заключает в себе (покрывает) неизвестный параметр Θ, равна γ.

Доверительным называют интервал $(\Theta^*-\delta, \Theta^*+\delta)$, который покрывает неизвестный параметр с заданной надежностью γ.

Замечание. Интервал $(\Theta^*-\delta, \Theta^*+\delta)$ имеет случайные концы (их называют доверительными границами). Действительно, в разных выборках получаются различные значения Θ*. Следовательно, от выборки к выборке будут изменяться и концы доверительного интервала, т. е. доверительные границы сами являются случайными величинами – функциями от $x_1, x_2, \ldots, x_n$.

Так как случайной величиной является не оцениваемый пара метр Θ, а доверительный интервал, то более правильно говорить, не о вероятности попадания Θ в доверительный интервал, а о вероятности того, что доверительный интервал покроет Θ.

Метод доверительных интервалов разработал американский статистик Ю. Нейман, исходя из идей английского статистика Р. Фишера.[4]

Вопросы и задания:

1. Что называется *точечной оценкой*?
2. Что называется *интервальной оценкой*?
3. Дайте определение понятию «*доверительная вероятность*».
4. *Доверительный интервал*: что это?
5. Кто является основателем метода доверительных интервалов?

Текст 3. Оценка математического ожидания

Координата центра распределения определяет положение плотности распределения на числовой оси. Возможны различные определения понятия центра распределения. Как правило, за центр распределения принимают центр тяжести однородной плоской фигуры, ограниченной плотностью распределения, т.е. первый начальный момент, или математическое ожидание. Однако для островершинных распределений, близких к распределению Коши и не имеющих моментов, понятие математического ожидания не имеет смысла. В таких случаях в качестве характеристики положения можно принять центр симметрии как точку на оси, слева и справа от которой появление случайной величины равновероятно и составляет 0,5. Согласно определению, – это медиана. Кроме того, такой характеристикой могут быть центр размаха, середина интерквантильного промежутка и др., существующие для любых распределений. Тем не менее для имеющих моменты распределений, в большинстве случаев встречающихся на практике, в качестве оценки центра распределения чаще других используют математическое ожидание (генеральное среднее). Особенно велика роль среднего при обработке результатов измерения, так как оно совпадает с искомым значением.

Как было показано ранее, оценкой максимального правдоподобия математического ожидания нормального закона является среднее арифметическое, совпадающее с оценкой по методу моментов независимо от закона распределения. Среднее арифметическое - несмещенная оценка для любого распределения, так как математическое ожидание равно истинному значению:

$$M[\bar{x}] = M\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} M[x_i] = m$$

Дисперсия среднего арифметического:

$$D(\bar{x}) = \sigma_{\bar{x}}^2 = D\left\{\frac{1}{n}\sum_{i=1}^{n} x_i\right\} = \frac{1}{n^2}\sum_{i=1}^{n} D(x_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

Среднее квадратическое отклонение результата измерения:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Таким образом, случайная величина $\bar{x}$ имеет среднее значение m и стандартное отклонение $\sigma/\sqrt{n}$. Если среднее арифметическое получено по выборке из нормальной совокупности, то благодаря устойчивости нормального закона (сумма нормально распределенных случайных величин является нормальной) оно распределено также по нормальному закону. [2]

Вопросы и задания:

1. Что такое *математическое ожидание*?
2. Прочитайте формулу для расчета *математического ожидания.*
3. Прочитайте формулу для расчета *дисперсии*.

Тема 5. Проверка статистических гипотез

Текст 1. Статистические гипотезы

В прикладных исследованиях часто возникает необходимость формулировки, а затем и экспериментальной проверки некоторых предположительных утверждений. Например, применение новых медицинских препаратов и методик существенно повышает эффект лечения или применение новых технологий приводит к повышению производительности труда. Кроме того, непосредственно при выполнении статистических расчетов исходные данные подвергают детальному исследованию: удовлетворяют ли они тем предположениям, на которых основаны вычислительные статистические процедуры (например, алгоритмы оценивания), такие как независимость случайных наблюдений, равенство дисперсий погрешностей, одинаковость законов распределения нескольких случайных величин, модель закона распределения, из которого получена выборка, симметричность распределений и др. Все эти предположения можно рассматривать как гипотезы, которые необходимо проверить. Статистические гипотезы охватывают поведение случайных величин.

Решение о том, справедлива ли данная гипотеза, основывается на знании выборки заданного объема, а также, возможно, и на знании априорной информации о законе распределения, из которого получена выборка, если таковая имеется.

Статистическую гипотезу, являющуюся утверждением о параметрах распределения некоторой случайной величины, например о среднем значении, дисперсии и др., называют *параметрической*.

Статистическую гипотезу называют *непараметрической*, если она является: а) утверждением о некоторых свойствах распределения исследуемых случайных величин, например о совпадении функции распределения двух и более случайных величин, принадлежности выборки данному классу распределений и т.п.; б) независимым от вида распределения утверждением о параметрах случайных величин, например о равенстве двух и более средних арифметических или дисперсий.

Методы проверки статистических гипотез называют статистическими *критериями*.

Гипотезу, которая проверяется, называют *нулевой* (H_0). Гипотезу, которая противопоставляется нулевой, называют *альтернативной* (H_1). Нулевая гипотеза рассматривается как утверждение, которое является существенно важным, если она отвергается, т.е. теория должна быть отвергнута, если есть хотя бы один противоречащий факт, но необязательно должна быть принята, если такой пример не найден. По-другому, в основу проверки статистических гипотез положен принцип, согласно которому маловероятные события считаются практически невозможными, и наоборот, те события, вероятность которых близка к единице, принимаются за достоверные. Дозволенная степень риска, связанная с пренебрежением событиями с малой вероятностью, зависит от разного рода обстоятельств и прежде всего от практической важности следствий, вытекающих из наступления таких событий.

Если конкурирующие гипотезы полностью определяют распределение случайной величины, например, значение математического ожидания m равно m_0 (гипотеза H_0) или m_1 (гипотеза H_1), то они являются *простыми*. В противном случае – *сложными*. В общем случае, если распределение имеет l параметров и гипотеза утверждает, что k из них имеют заданные значения, то она является простой при $k = l$ и сложной при $k < l$. Если по некоторой гипотезе случайная величина распределена по нормальному закону с известным математическим ожиданием и неизвестной дисперсией, то гипотеза сложная.

При проверке гипотез необходимо применить критерий, с помощью которого по полученному значению статистики можно сделать разумный выбор между нулевой и альтернативной гипотезами. Построение критерия начинается с выбора такого множества или области, попадание в которую найденной статистики означает принятие нулевой гипотезы. Такую область (множество) называют *областью принятия гипотезы*, дополнительную – *критической*, при попадании в которую нулевая гипотеза отвергается. При проверке гипотезы H_0 против H_1 возможны два вида ошибок. Ошибка *первого рода* – это отклонение верной гипотезы H_0, *второго рода* – принятие неверной гипотезы H_1. Вероятность ошибки первого рода называют *уровнем значимости* критерия и обозначается он $\alpha = P$ (отвергается верная гипотеза H_0). Ошибка первого рода должна совершаться как можно реже, поэтому уровень значимости α выбирается малым: 0,10; 0,05; 0,01 и т.д. При этом желательно, чтобы выбранный критерий минимизировал вероятность ошибки второго рода $\beta = P$ (принимается гипотеза H_1 при верной гипотезе H_0). Вероятность дополнительного события, т.е. правильного отклонения гипотезы H_1, называют *мощностью критерия*. Численно мощность критерия $\pi = 1 - \beta$, т.е. равна вероятности того, что наблюдаемое значение статистики попадет в критическую область при верной альтернативной гипотезе H_1 [2].

Вопросы и задания:

1. Что такое *гипотеза*?
2. Какими бывают *статистические гипотезы*?
3. Приведите пример сложной гипотезы.
4. При принятии гипотезы существует риск совершить ошибки. Какие *ошибки* могут иметь место в статистике?
5. Что такое *уровень значимости*?

Текст 2. Важные понятия при проверке статистических гипотез.

При проверке гипотеза может подтвердиться, а может и не подтвердиться. Проверка статистических гипотез сопряжена с возможностью допустить ошибку.

Ошибка первого рода состоит в том, что будет отвергнута верная гипотеза.

Ошибка второго рода состоит в том, что будет принята ложная гипотеза.

Вероятность совершения ошибки первого рода обозначается *б* и называется *уровнем значимости*. Уровень значимости α обычно задается близким к нулю (например, 0,05; 0,01; 0,02 и т.д.). Чем меньше уровень значимости α, тем меньше вероятность отвергнуть проверяемую гипотезу H_0, когда она верна, т.е. совершить ошибку первого рода.

При этом желательно, чтобы выбранный критерий минимизировал вероятность ошибки второго рода β (принимается гипотеза H_1 при верной гипотезе H_0). Вероятность не отклонить ложную гипотезу обозначается β. Вероятность дополнительного события, т.е. правильного отклонения гипотезы H_1, называют *мощностью критерия.* Численно мощность критерия $\pi = 1 - \beta$, равна вероятности того, что наблюдаемое значение статистики попадет в критическую область при верной альтернативной гипотезе. H_1.

Проверка любой статистической гипотезы осуществляется с помощью статистического критерия.

Статистический критерий – это случайная величина (статистика), которая используется с целью проверки нулевой гипотезы.

Множество всех возможных значений выбранного статистического критерия разделяется на два непересекающихся подмножества: критичная область и область принятия гипотезы (рисунок 7.4).

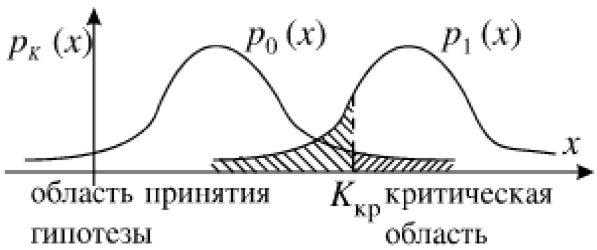


Рисунок 7.4 – Проверка статистической гипотезы

Критическая область – это множество возможных значений статистического критерия, при которых нулевая гипотеза отвергается.

Область принятия гипотезы – это множество возможных значений статистического критерия, при которых нулевая гипотеза принимается.

В том случае, если наблюдаемое значение статистического критерия (рассчитанное по выборочной совокупности) принадлежит критической области, нулевую гипотезу отвергают. Если же наблюдаемое значение статистического критерия принадлежит области принятия гипотезы, то нулевая гипотеза принимается.

Критические точки (квантили) – это точки, которые разграничивают критическую область и область принятия гипотезы.

Порядок проверки статистической гипотезы таков:

1) задается уровень значимости α, выбирается статистический критерий K и вычисляется (обычно по таблицам для закона распределения K) значение $K_{кр}$; определяется вид критической области;

2) по выборке вычисляется наблюдаемое значение критерия $K_{набл}$;

3) если $K_{набл}$ попадает в критическую область, нулевая гипотеза отвергается; при попадании $K_{набл}$ в область принятия гипотезы нулевая гипотеза принимается. [1]

Вопросы и задания:

1. Сформулируйте ошибки, которые возникаю при проверке статистических гипотез.
2. Что такое *статистический критерий*?
3. Объясните понятия «*критическая область*» и «*Область принятия гипотезы*».
4. Приведёте последовательность действий при проверке статистических гипотез.

Текст 3. Гипотезы о значениях числовых характеристик

Гипотезы о равенстве среднего значения m и дисперсии σ^2 определенным числам m_0 и σ_0^2. Они возникают, например, при проверке качества функционирования измерительных устройств. Если m_0 - номинальное значение измеряемого параметра и $m \neq m_0$, то это означает, что прибор дает систематическую ошибку. Точность прибора определяется значением σ_0 и, если $\sigma > \sigma_0$, то это означает, что качество прибора не отвечает стандартным требованиям.

Сравнение средних — это проверка гипотезы о равенстве средних значений совокупностей, из которых получены выборки.

На практике часто встречаются ситуации, когда среднее значение данных одного эксперимента отличается от среднего значения данных другого, хотя условия эксперимента являются схожими. Тогда возникает вопрос, можно ли считать это расхождение незначимым, т.е. чисто случайным, или оно вызвано существенным различием двух генеральных совокупностей. Например, такие вопросы возникают при исследовании надежности технических систем, где результаты сравниваются с предыдущими измерениями; при контроле качества изделий, изготовленных на разных предприятиях; в финансах – при сравнении уровня доходности различных активов.

Пример 1. Было произведено $n_1 = 12$ измерений диаметра вала (в мм). При этом оказалось, что среднее $\overline{x_1} = 10,2$, а стандартное среднее квадратичное отклонение $\sigma_1 = 0,05$. Затем вал поместили в условия с высокой температурой и провели $n_2 = 8$ измерений диаметра его оси. Среднее на этот раз оказалось равным $\overline{x_2} = 10,25$, а стандартное отклонение $\sigma_2 = 0,06$. Можно ли сделать вывод, что диаметр вала существенно увеличивается при увеличении температуры? Доверительная вероятность $P = 0,95$.

Гипотезы о дисперсиях в науке, технике, экономике, медицине, социологических исследованиях играют очень большую роль, поскольку измеряемое дисперсией рассеяние характеризует изменяемость интересующего признака в определенных условиях. Среднее значение сглаживает, затушевывает эти изменения, а дисперсия их выявляет.

Дисперсия характеризует точность приборов, технологических процессов, риск, связанный с отклонением доходности от заданного уровня, и т.д.

Сравнение дисперсий в измерительной практике принципиально необходимо при объединении двух и более рядов измерений.

Пример 2. Точность работы станка-автомата проверяется по дисперсии σ^2 контролируемого размера изделий, которая не должна превышать 0,15. По данным из 25 отобранных изделий вычислена несмещенная дисперсия $\overline{S^2} = 0,25$. Выяснить, обеспечивает ли станок требуемую точность. Доверительная вероятность P = 99 %. [1]

Вопросы и задания:

1. В чем заключается гипотеза о значениях числовых характеристик?
2. Когда возникает вопрос о равенстве средних значений случайных величин?
3. Когда возникает вопрос о равенстве дисперсий двух случайных величин?
4. Какую информацию несет дисперсия в реальных условиях?
5. Приведите собственный пример гипотезы о равенстве средних значений.
6. Приведите собственный пример гипотезы о равенстве дисперсий.

8. Список терминов и определений

Термин	Определение	Перевод на китайский язык
Глава 1. Основные понятия теории вероятности		
Теория вероятностей	Математическая наука, позволяющая по вероятности одних случайных событий находить вероятности других случайных событий, связанных каким–либо образом между собой	概率论- 概率论是研究随机性或不确定性等现象的数学. 概率论是用来模拟在同一环境下会产生不同结果的实验。如果一个随机试验所包含的单位事件是有限的，那么每个单位事件发生的可能性均相等。
Исход опыта	результат любого опыта	实验结果- 任何实验的结果
Событие	Исход или группа исходов, удовлетворяющих определённым требованиям	事件- 试验次数增多，我们可以观察到某种规律性的结果符合特定条件
Благоприятный исход.	Исходы опытов, соответствующие установленному событию	有利结果- 实验的结果符合特定的情況
Достоверное событие	Событие, которое всегда происходит в рассматриваемом эксперименте	确定事件- 在实验过程中总是发生的事件
Невозможное событие	Событие, которое никогда не может наступить в рассматриваемом эксперименте.	不可能事件- 一在实验中不可能发生的事件
Случайное событие	Событие, которое при проведении эксперимента (опыта) может наступить, а может и не наступить	随机事件 (概率论)- 在实验中有可能发生也有可能不发生的事件

Термин	Определение	Перевод на китайский язык
Полная группа событий	Несколько возможных событий, если в результате опыта обязательно должно произойти хотя бы одно из них	必然事件 一在实验中必然会发生的事件
Несовместные события	Несколько событий в опыте, которые не могут произойти одновременно	互斥事件一在实验中有些事件不能同时发生
Равновозможные события	События, вероятности наступление которых равны между собой	同概率事件- 一些事件发生的概率相同
Вероятность события	1. Численная мера степени объективной возможности этого события. 2. Отношение благоприятного числа исходов опыта к общему числу всех равновозможных несовместных элементарных исходов, образующих полную группу	事件发生率- 1.一件事情发生的几率 2. 有利结果，同概率事件，互斥事件和必然事件之间的比例
Относительной частотой события	Отношение числа испытаний, в которых событие появилось, к общему числу фактически произведенных испытаний	实际发生事件在整个实验中所占的比例

Термин	Определение	Перевод на китайский язык
Глава 2. Случайные величины		
Случайная величина (СВ)	Величина, которая в результате опыта может принимать то или иное значение, причём до завершения опыта неизвестно какое在可能有这样或者那样意义的实验结果的可能性	随机变量 在不同的条件下由于偶然因素影响，其可能取各种不同的值,实验完成以前不确定性的.
Дискретная случайная величина (ДСВ).	Случайная величина, которая может принимать отдельные, изолированные значения с определенными вероятностями	离散性的随机变量-只取有限种可能值的随机变量
Закон распределения случайной величины	Любое правило, устанавливающее соответствие между значениями случайной величины и вероятностями ее наступления	概率分布-随机变量X小于任何已知实数x的事件可以表示成的函数
Ряд распределения	Совокупность всех возможных значений x_i и соответствующих им вероятностей $p_i = P(X = x_i)$,	分布序列 所有的x_i数值及相应的概率 $p_i = P(X = x_i)$
Многоугольник распределения.	Графическое изображение ряда распределения. Для его построения возможные значения случайной величины (x_i) откладываются по оси абсцисс, а вероятности p_i – по оси ординат; точки A_i с координатами (x_i,p_i) соединяются ломаными линиями	频数多边形 - 概率分布级数的记录.为了绘图频数多边形应该绘图随机变量(x_i)的可能值绘图在横坐标轴, 概率值(p_i)的可能值绘图在纵坐标轴. A_i点(x_i,p_i)位置把折线接上

Термин	Определение	Перевод на китайский язык
Функция распределения	Функция F(x), значение которой в точке x равно вероятности того, что случайная величина X будет меньше этого значения x, то есть $F(x) = P(X < x)$	累积分布函数 又叫累计分布函数，是概率密度函數的积分，能完整描述一個实际随机变量X的概率分佈。一般以大寫“對於所有實數x，累积分布函数定義如下 $F(x) = P(X < x)$
Непрерывная случайная величина (НСВ).	Случайная величина, возможные значения которой непрерывно заполняют некоторые конечные или бесконечные промежутки	连续随机变量- 若随机变量的分布函数可表示成一个非负可积函数的积分，则称为连续型随机变量
Плотность ю распределения (или плотность ю вероятности)	Производная функции распределения случайной величины	概率密度函数- 机率密度箱编号
Математическое ожидание	Числовая характеристика, обозначающая среднее значение случайной величины	期望值,是随机试验在同样的机会下重复多次的结果计算出的等同“期望”的平均值
Центрированная случайная величина	Отклонение случайной величины от ее математического ожидания	中心的随机变量- 不适合数学期望的随机数
Дисперсия	Числовая характеристика: математическое ожидание квадрата соответствующей центрированной случайной величины, характеризует разброс данной случайной величины относительно ее среднего значения	变异量(数) - 一个随机变量的方差描述的是它的离散程度，也就是该变量离其期望值的距离

Термин	Определение	Перевод на китайский язык
Глава 3. Элементы математической статистики		
Математическая статистика	Раздел математики, посвященный математическим методам сбора, систематизации, обработки и интерпретации статистических данных	数理统计学,是研究有效地运用数据收集与数据处理、多种模型与技术分析与统计分析等.
Статистические данные	Сведения о числе объектов в какой-либо более или менее обширной совокупности, обладающих теми или иными признаками.	统计材料—所有统计数据的总和
Выборочная совокупность (выборка)	Совокупность случайно отобранных объектов	标本 (统计)-随机选择的对象总和
Генеральная совокупность	Совокупность объектов, из которых производится выборка	总体-选择对象的总和
Объем совокупности (выборочной или генеральной)	Число объектов совокупности (n или N)	样本量-样本中所包含的个体数目n或N为样本容量
Варианты (ед. Варианта)	Наблюдаемые значения случайной величины	函数-描述每个输入值对应唯一输出值的这种对应关系
Вариационный ряд	Последовательность вариант, записанных в возрастающем порядке	有序样本-按递增次序写得函数的连贯性

Термин	Определение	Перевод на китайский язык
Частота	Число повторений значения варианты x_i	频率-函数 x_i 重复发生的 次数
Относительная частота	Отношение частоты к объему выборки $W_i = \frac{n_i}{n}$	相对频率-频率比样本量的关系式
Полигон частот	Ломаная линия, отрезки которой соединяют точки $(x_1;n_1)$, $(x_2;n_2)$, ..., $(x_k;n_k)$.	频数多边图-折线的一段接上坐标$(x_1;n_1)$, $(x_2;n_2)$, ..., $(x_k;n_k)$.
Гистограмма частот	Ступенчатая фигура, состоящую из прямоугольников, основаниями которых служат частичные интервалы длиною h, а высоты равны отношению n_i/h (плотность частоты).	频率直方图-是一种二维统计图表，它的两个坐标分别是统计样本和该样本对应的某个属性的度量

Глава 4. Основы теории оценивания

Выборочный метод	Из генеральной совокупности случайным образом извлекается ограниченное число объектов, образующих выборку. По детальному исследованию выборки определяются свойства (характеристики) генеральной совокупности.	统计抽查法,是应用概率论与数理统计理论的原理，遵循随机原则，从被查总体中抽出样本进行检查，并根据对样本的检查结果推断总体的一种抽样检查方法。
Оценка истинного значения параметра Θ	Полученное по результатам эксперимента приближенное значение $\tilde{\Theta}$ искомого параметра	参数点估计-从经验接受的近似值 $\tilde{\Theta}$ 的请求参数

Термин	Определение	Перевод на китайский язык
Точечная оценка	Единственное числовое значение, которое и принимается за значение искомого параметра	点推定-点估算 惟一的所需参数的量数
Генеральное среднее $\bar{x}_г$	среднее арифметическое значений признака генеральной совокупности	平均值(平均数)$\bar{x}_г$- 为集中趋势的最常用测度值，目的是确定一组数据的均衡点。
Выборочное среднее $\bar{x}_в$	среднее арифметическое значение признака выборочной совокупности	样本平均值,是一个向量，是随机变数观测值的几何平均。
Генеральная дисперсия $D_г$	среднее арифметическое квадратов отклонений значений признака генеральной совокупности от их среднего значения $\bar{x}_г$	普通变量$D_г$- 普通标准差离均与 $\bar{x}_г$的平均值的算术平均
Генеральное среднее квадратическое отклонение	квадратный корень из генеральной дисперсии	普通标准差离均- 差平方和的平均
Выборочная дисперсия $D_в$	Среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения $\bar{x}_в$	总体方差,是一组资料中各数值与其算术平均数离差平方和的平均数
Выборочное среднее квадратическое отклонение	Среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения $\bar{x}_в$	抽选的標準差- 離均差平方和的平均。方差的算术平方根，反映组内个体间的离散程度。测量到分布程度的结果
Исправленная дисперсия	Несмещенная оценка генеральной дисперсии	修正变量-的无偏估计

Термин	Определение	Перевод на китайский язык
Функция правдоподобия	совместная плотность распределения выборки $L = \prod_{i=1}^{n} f(x_i, \tilde{\Theta}).$	似然函数- 表示模型参数中的似然性 $L = \prod_{i=1}^{n} f(x_i, \tilde{\Theta}).$ 选录 的分布密度
Теоретический момент	Момент, вычисленные аналитическим путем через плотность f(x) или функцию распределения F(x)	理论的矩 - 算出f(x)密度或者F(x) 累积分布函数用分析的方法的一种矩
Доверительный интервал	Интервал $[\Theta_{i1}; \Theta_{i2}]$, между границами которого с заданной вероятностью Р = 1 – α заключено истинное значение xi	置信区间,是对这个样本的某个总体参数的 $[\Theta_{i1}; \Theta_{i2}]$区间估计 P = 1 – α 在这些限度xi的真值.
Доверительная вероятность, Р	Степень уверенности в том, что доверительный интервал действительно будет содержать истинное (неизвестное) значение	置信区间的置信水平 P- 事件本身并没有什么概率，事件之所以指派有概率只是指派概率的人头脑中所具有的信念证据
Уровень значимости, α	α = 1 – Р	置信度-而1-α, α = 1 – Р
Глава 5. Проверка статистических гипотез		
Статистическая гипотеза	Любое предположение о виде неизвестного закона распределения или о параметрах известных распределений	统计 上 对 参数 的 假设, 就是对 一个 或 多个 参数 的 论述.
Параметрические гипотезы	Гипотезы о значениях параметров распределений или о сравнительной величине параметров двух распределений	参数假设 - 设A是关于总体分布的一项命题 如果h0可以通过有限个实参数来描述，则称为参数假设

Термин	Определение	Перевод на китайский язык
Непараметрические гипотезы	Гипотезы о виде распределения	非参数假设 -分布种类假设
Проверить статистическую гипотезу	Проверить, согласуются ли выборочные данные с выдвинутой гипотезой	假设检定-会希望根据结果对未知的真正参数值做出适当的推论
Нулевая (основная или проверяемая) гипотеза	Выдвинутая гипотеза H_0	零假设 -提出的H_0假设
Конкурирующая (альтернативная) гипотеза	Гипотеза H_1, которая противоречит нулевой гипотезе	对立假设-对方零假设的H_1假设
Ошибка первого рода	состоит в том, что будет отвергнута верная гипотеза	第一型错误-拒绝了实际上成立的，为“弃真”的错误
Ошибка второго рода	состоит в том, что будет принята ложная гипотеза.	第二型错误,为在进行假设检验时，原假设不正确而接受原假设的错误。
Мощность критерия	Вероятность дополнительного события, т.е. правильного отклонения гипотезы H_1	统计功效 -在假设检验中，拒绝原假设后，接受正确的替换假设的概率
Статистический критерий	Случайная величина (статистика), которая используется с целью проверки нулевой гипотезы	检验统计量-在零假设情况下，这项统计量服从一个给定的概率分布，而这在另一种假设下则不然

Термин	Определение	Перевод на китайский язык
Критическая область	Множество возможных значений статистического критерия, при которых нулевая гипотеза отвергается	关键领域 - 一些统计直, 拒绝零假设的数值
Область принятия гипотезы	Множество возможных значений статистического критерия, при которых нулевая гипотеза принимается	接受带- 一些统计直, 接受零假设的数值
Критические точки (квантили)	Точки, которые разграничивают критическую область и область принятия гипотезы	临界点（分位点）,是界定临界区和接受假设区的点
Сравнение средних	проверка гипотезы о равенстве средних значений совокупностей, из которых получены выборки.	平均数比较- 平均值等式的假设检验
Сравнение дисперсий	проверка гипотезы о равенстве дисперсий двух совокупностей, из которых получены выборки	随机变量比较- 两总和变量的等式的假设的检验
Критерий согласия	Статистический критерий проверки о предполагаемом законе неизвестного распределения	拟合优度, 是指回归直线对观测值的拟合程度。

9. Список используемых источников

1. Вавилова Г.В. Математическая обработка результатов измерения: учебное-методическое пособие/ Г.В. Вавилова; Национальный исследовательский Томский политехнический университет. - Томск: Изд-во Томского политехнического университета, 2013.- 160 с.

2. Садовский Г.А. Теоретические основы информационно-измерительной техники: учебное пособие / Г.А. Садовский. – М.: Высшая школа, 2008. – 478 с. : ил.

3. Лазарева Л.И. Теория вероятностей. Математическая статистика: учебное пособие/ Л. И. Лазарева, А. А. Михальчук; Томский политехнический университет. — Томск: Изд-во ТПУ, 2000. – 136 с.

4. Гмурман В.Е. Теория вероятностей и математическая статистика: учебное пособие / В. Е. Гмурман. – М.: Высшая школа, 1997. – 479 с. : ил.

Учебное издание

ВАВИЛОВА Галина Васильевна
ГНЕЗДИЛОВА Наталия Сергеевна

РКИ: ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебно-методическое пособие

Научный редактор кандидат наук, доцент кафедры РКИ,
Замятина Е.В.
Компьютерная верстка
Дизайн обложки

Подписано к печати Формат 60х84/16. Бумага «Снегурочка».
Печать XEROX. Усл.печ.л. 9,01. Уч.-изд.л. 8,16.
Заказ . Тираж 100 экз.

Национальный исследовательский Томский политехнический университет
Система менеджмента качества Томского политехнического университета сертифицирована NATIONAL QUALITY ASSURANCE по стандарту ISO 9001:2008

ИЗДАТЕЛЬСТВО ТПУ. 634050, **г.** Томск, **пр.** Ленина, 30
Тел./факс: 8(3822)56-35-35, www.ipu.ru