

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

**ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ
СТАТИСТИКА**

Методические указания для подготовки к практическим занятиям

Направление подготовки 03.03.02 Физика
Направленность (профиль)
«Физика Земли и космоса»

**Ставрополь
2026**

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	
РАЗДЕЛ 1. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ.....	4
Тема 1. Элементы комбинаторики.....	4
РАЗДЕЛ 2. ТЕОРИЯ ВЕРОЯТНОСТЕЙ	6
Тема 2. Основные понятия теории вероятностей	6
Тема 3. Теоремы сложения и умножения вероятностей.	8
Тема 4. Формула полной вероятности.....	9
Тема 5. Теорема гипотез и формула Бернулли	10
Тема 6. Предельные теоремы	11
РАЗДЕЛ 3. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ	12
Тема 7. Дискретная случайная величина	12
Тема 8. Непрерывная случайная величина	17
Тема 9. Основные законы распределения случайных величин	19
РАЗДЕЛ 4. СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ.....	25
Тема 10. Основные понятия математической статистики.....	25
Тема 11. Оценка параметров генеральной совокупности по ее выборке.....	28
Тема 12. Статистическая проверка гипотез.....	31
СПИСОК ЛИТЕРАТУРЫ	

ВВЕДЕНИЕ

Актуальность учебной дисциплины «Теория вероятностей и математическая статистика» обусловлена необходимостью формирования у будущих специалистов общепрофессиональных компетенций, практических умений на которых базируется профессиональная деятельность в физике. В современных условиях важнейшим элементом подготовки бакалавров в области физики является формирование у них самостоятельности и культуры научного мышления, способности вести научную и профессиональную деятельность, навыков постановки целей, задач при решении проблемных ситуаций.

Целью практических занятий является формирование профессиональных компетенций бакалавра по направлению 03.03.02 Физика, связанных с освоением необходимого математического аппарата, с помощью которого разрабатываются и исследуются теоретические и экспериментальные модели

Основные задачи дисциплины:

Развитие навыков математического мышления студентов.

- Овладение методов исследования и решения исследовательских задач.
- Выработка у студентов умения самостоятельно расширять свои математические знания.
- Развитие навыков использования математических методов и основ математического моделирования в практической деятельности.

В результате изучения дисциплины у обучающихся формируются следующие компетенции (в соответствии с образовательным стандартом):

ИД-2_{ОПК-1}

использует математический аппарат и методы математического и численного моделирования для описания, анализа, теоретического и экспериментального исследования физических систем, явлений и процессов в своей профессиональной деятельности.

Цель данного практикума – описание последовательности, основного содержания, методики выполнения практических занятий по учебному курсу и (или) самостоятельной подготовки к ним. Занятия проводятся в форме практических.

РАЗДЕЛ 1. ЭЛЕМЕНТЫ КОМБИНАТОРИКИ.

Тема 1. Элементы комбинаторики.

Комбинаторика — область математики, изучающая количество комбинаций, подчиненных определенным условиям, которые можно составить из элементов заданного конечного множества.

Предметы, из которых состоят соединения, называются **элементами**.

Множеством называется совокупность элементов, объединенных по какому-либо признаку. Например, множество студентов, множество точек из окружности или множество действительных чисел и др.

Группы, составленные из каких-либо элементов и отличающиеся одна от другой либо самими элементами, либо порядком их следования называются **соединениями**.

В комбинаторике рассматриваются три типа соединений: **число размещений**, **число перестановок** и **число сочетаний**.

I. Число размещений

Размещением из n различных элементов по m в каждом называются такие соединения, которые отличаются друг от друга либо самими элементами (хотя бы одним), либо порядком их следования. Число размещений обозначается: A_n^m

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

Факториал — это произведение натуральных чисел от 1 до n . Факториал обозначается символом «!».

Размещение с повторением из n элементов по m элементов можно составить не только из различных элементов, но и из m каких угодно и, как угодно, повторяющихся элементов. Число размещений с повторяющимися элементами, находятся по формуле:

$$\bar{A}_n^m = n^m$$

II. Число перестановок

Перестановками из n различных элементов называют такие соединения, которые состоят из одних и тех же n различных элементов и отличаются только порядком их расположения. Число всех возможных перестановок обозначаются:

$$P_n = n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$$

В комбинаторике, принято считать, что: $0! = 1$.

Перестановка с повторяющимися элементами. Пусть, среди n возможных элементов имеется n_1 элементов первого типа, n_2 элементов второго типа и n_k элементов k -го типа. Тогда, число перестановок с повторением будет равно:

$$P_{n_1, n_2, \dots, n_k} = \frac{(n_1 + n_2 + \dots + n_k)!}{n_1! n_2! \cdot \dots \cdot n_k!}$$

III. Число сочетаний

Сочетанием из n элементов по m в каждом называются такие соединения, которые отличаются друг от друга по крайней мере одним элементом. Число сочетаний определяется по формуле:

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

Сочетание с повторением из n элементов по m , может содержать любой элемент сколь угодно раз от 1 до m включительно, или не содержать его совсем. Число сочетаний с повторением можно найти по формуле:

$$\bar{C}_n^m = C_{n+m-1}^m = \frac{(n+m-1)!}{m! \cdot (n-1)!}$$

Одной из наиболее распространённых комбинаторных формул является формула числа сочетаний. Для упрощения подсчётов и для доказательства некоторых утверждений удобно использовать следующие свойства сочетаний:

- 1) $C_n^m = C_n^{n-m}$ (свойство симметрии);
- 2) $C_{n+1}^{m+1} = C_n^m + C_n^{m+1}$ (свойство Паскаля);
- 3) $C_n^0 + C_n^1 + C_n^2 + C_n^3 + \dots + C_n^n = 2^n$;
- 4) $C_n^0 = C_0^0 = 1$;
- 5) $C_n^1 = n$.

Для решения комбинаторных задач, кроме основных формул комбинаторики, пользуются двумя важными правилами: правило суммы и правила произведения.

Правило суммы. Если некоторый объект A можно выбрать n способами, а объект B можно выбрать m способами, причем первые и вторые способы не пересекаются, то любой из указанных объектов (A или B) можно выбрать $(n + m)$ способами.

Правило произведения. Если из некоторого конечного множества первый объект A можно выбрать n способами и после каждого такого выбора второй объект B можно выбрать m способами, то оба объекта (A и B) в указанном порядке можно выбрать $(n \cdot m)$ способами.

Типовые задания.

Пример 1. В шахматном турнире принимали участие 15 шахматистов, причем каждый из них сыграл только одну партию с каждым из остальных. Сколько всего партий было сыграно в этом турнире?

Решение: В одной игре участвуют 2 человека, следовательно, нужно вычислить, сколькими способами можно отобрать 2-х человек из 15, при чем порядок в таких парах не важен. Воспользуемся формулой для нахождения числа сочетаний (выборки, отличающихся только составом) из различных элементов по m элементов

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

По условию: $n=15$, $m=2$. Следовательно число сочетаний из 15 элементов по 2 без повторения равно:

$$C_{15}^2 = \frac{15!}{2!(15-2)!} = 105$$

Ответ. 105 партий.

Пример 2. В пассажирском поезде 9 вагонов. Сколькими способами можно рассадить в поезде 4 человека, при условии, что все они должны ехать в различных вагонах?

Решение: Т.к. все пассажиры должны ехать в разных вагонах, требуется отобрать 4 вагона из 9 с учетом порядка (вагоны отличаются №), эти выборки – размещения из n различных элементов по m элементов, где $n=9$, $m=4$. Число таких размещений находим по формуле:

$$A_n^m = \frac{n!}{(n-m)!} \quad A_9^4 = \frac{9!}{(9-4)!} = 3024$$

Ответ: 3024 способами можно рассадить в поезде 4 человека.

Пример 3. В организации имеется две вакансии на должность сотрудника, одна вакансия на должность помощника сотрудника и две вакансии на должность секретаря. Сколько существует вариантов распределения должностей, если имеется 5 кандидатов?

Решение: так как всего имеется 5 вакансий и кандидатов тоже 5, то состав не важен, а соединения будут отличаться только по порядку расположения элементов. Следовательно,

для решения задачи необходимо найти число перестановок из 5 элементов с повторениями:

$$P_n(n_1, n_2, \dots, n_m) = \frac{n!}{n_1! \cdot n_2! \cdot \dots \cdot n_m!}$$

По условию задачи: $n_1 = 2, n_2 = 1, n_3 = 2, n = 5$, тогда

$$P_5(2,1,2) = P_n(2;1;2) = \frac{5!}{2_1! \cdot 1! \cdot 2!} = 30$$

РАЗДЕЛ 2. ТЕОРИЯ ВЕРОЯТНОСТЕЙ.

Тема 2. Основные понятия теории вероятностей

Теория вероятности – это математическая наука, изучающая закономерности в случайных явлениях. Основными понятиями теории вероятности являются испытания (опыт), события и вероятность событий.

Испытание (опыт) – это осуществление определенного комплекса условий, при которых производится наблюдение.

Событие – качественный результат испытания, который может произойти или не произойти.

Обычно, события в теории вероятности обозначаются прописными буквами латинского алфавита: $A, B, C, \dots$

Приведем примеры событий:

A - попадание в цель; B - появление герба, при бросании монеты; C - появление ошибки измерения.

Чтобы количественно сравнить между собой события по степени их возможности, очевидно, нужно с каждым событием связать определенное число, которое тем больше, чем более возможно событие, такое число называется **вероятностью события**.

События, которые могут произойти в результате испытания, принято классифицировать на три вида: достоверные, невозможные, случайные.

Достоверное событие – это событие, которое в результате испытаний непременно должно произойти.

Противоположность достоверному событию является невозможное событие.

Невозможное событие – это событие, которое при испытании не может произойти.

Событие называется **случайным**, если при испытании оно может либо произойти, либо не произойти.

Пример. Выпадение не более 6 очков при бросании одной игральной кости. Если приписать достоверному событию вероятность равную единице, то все другие события возможные, но не достоверные будут характеризоваться вероятностью меньше единицы, составляющую какую – то долю. Появление 12 очков, при выбрасывании одной игральной кости - невозможное событие. Вероятность такого события равна нулю. Выпадение 5 очков - случайное событие.

События A и B называются **несовместными**, если при одном испытании появление одного из них исключает появление другого.

События A и B называются **совместными**, если при испытании появление одного из них не исключает появления другого.

События A и B называются **равновозможными**, если нет оснований считать, что в данном испытании одно из них более возможно, чем другое.

Говорят, что несколько событий $(A_1, A_2, A_3, \dots, A_n)$ объединяют полную **группу событий**, если в результате опыта, события являются попарно несовместными и, непременно, должно появиться хотя бы одно из них.

Два несовместных события, образующие полную группу событий, называются **противоположными**.

Если какой-либо опыт обладает симметрией возможных исходов, по случай (шанс), представляет собой исчерпывающую систему равновозможных и исключаящих друг друга исходов хода.

Случай называется **благоприятным** некоторого события, если появление случая влечет за собой появление данного события.

Вероятностью события A называется отношение числа случаев, благоприятствующих ему, к общему числу случаев, т.е.

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

Геометрические вероятности $P(A) = \frac{l}{L}$, $P(A) = \frac{S_g}{S_G}$, $P(A) = \frac{V_g}{V_G}$

Статистическая вероятность, относительная частота $W(A) = \frac{m}{n}$

Урновая схема $P(A) = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}$

Основные свойства вероятности

- а) Вероятность любого события $0 \leq P(A) \leq 1$.
- б) Вероятность достоверного события $P(A) = 1$.
- в) Вероятность невозможного события $P(A) = 0$.

Типовые задания.

Пример 1. В ящике находится 14 кубиков: 8 желтых и 6 красных. Наудачу вынимают один кубик. Какова вероятность того, что вынутый кубик будет желтого цвета?

Решение: В данном случае число благоприятных исходов $m = 8$, а общее число исходов $n = 14$. Тогда искомая вероятность будет равна:

$$P(A) = \frac{m}{n} = \frac{8}{14} = 0,57 \approx 0,6.$$

Пример 2. На отрезок АВ длины $L=40$ см помещен маленький отрезок CD длиной $l=10$ см. Найти вероятность того, что точка, случайно брошенная на отрезок АВ, попадет также и на отрезок CD. Предполагается, что вероятность попадания точки на отрезок пропорциональна длине отрезка и не зависит от его расположения на числовой оси.

Решение: Событие A – точка, случайно брошенная на отрезок АВ, попадет также и на отрезок CD.

$$P(A) = \frac{l}{L} = \frac{10}{40} = \frac{1}{4}$$

Пример 3. На плоскости начерчены две концентрические окружности, радиусы которых 5 и 10 см соответственно. Найти вероятность того, что точка, брошенная наудачу в большой круг, попадет в кольцо, образованное построенными окружностями. Предполагается, что вероятность попадания точки в плоскую фигуру пропорциональна площади этой фигуры и не зависит от ее расположения относительно большого круга.

Решение: A - точка, брошенная наудачу в большой круг, попадет в кольцо, образованное построенными окружностями.

Площадь кольца (фигуры g)

$$S_g = \pi(R^2 - r^2) = \pi(10^2 - 5^2) = 75\pi$$

Площадь большого круга (фигуры G)

$$S_G = \pi R^2 = \pi 10^2 = 100\pi$$

Искомая вероятность

$$P(A) = \frac{S_g}{S_G} = \frac{75\pi}{100\pi} = 0,75.$$

Тема 3. Теоремы сложения и умножения вероятностей.

Теоремы сложения вероятностей

Суммой событий A и B называется событие, состоящее в появлении хотя бы одного из событий A и B и обозначается $A + B$.

теорема 1. Вероятность суммы двух несовместимых событий равна сумме вероятности этих событий, т.е.

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

теорема 2. Вероятность суммы двух совместных событий равна сумме вероятностей этих событий без вероятности их совместного появления этих событий, т.е.

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

Следствие 1. Если события $A_1, A_2, A_3, \dots, A_n$ образуют полную группу несовместных событий, то сумма их вероятностей равна единице, т.е.

$$\sum_{i=1}^n P(A_i) = 1.$$

Следствие 2. Сумма вероятностей противоположных событий равна единице, т.е.

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

Теоремы умножения вероятностей

Произведением событий A и B называется событие, состоящее в появлении обоих событий A и B . Обозначается $A \cdot B$.

теорема 1. Вероятность совместного появления двух независимых событий равна произведению их вероятностей:

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

теорема 2. Вероятность совместного появления двух зависимых событий равна произведению вероятности одного из них на условную вероятность другого, найденного в предположении, что первое событие произошло:

$$P(A \cdot B) = P(A) \cdot P\left(\frac{B}{A}\right),$$

где $P(B/A)$ – условная вероятность события B , при условии того, что событие A произошло.

Типовые задания.

Пример 1. В урне находятся шары: 7 синих, 5 зеленых и 6 белых. Из урны наудачу вынимается один шар. Найти вероятность того, что этот шар будет зеленый или белый.

Решение. Пусть событие A – изъятие зеленого шара, вероятность этого события равна: $P(A) = \frac{5}{18}$. Событие B – изъятие белого шара, вероятность события B равна: $P(B) = \frac{6}{18}$. События A и B являются несовместными, поэтому по теореме сложения для несовместных событий определяем искомую вероятность:

$$P(A + B) = P(A) + P(B) = \frac{5}{18} + \frac{6}{18} = \frac{11}{18} \approx 0,6.$$

Пример 2. Медицинская сестра обслуживает в палате четырех больных. Вероятность того, что в течение часа первый больной потребует внимания сестры равна 0,2; вероятность

второго больного равна 0,3; вероятность третьего больного 0,25; вероятность четвертого больного равна 0,1. Найдите вероятность того, что в течение часа все больные потребуют к себе внимания сестры.

Решение: Считая требования больных независимыми друг от друга, по теореме умножения находим, что искомая вероятность

$$P(A \cdot B \cdot C \cdot D) = P(A) \cdot P(B) \cdot P(C) \cdot P(D)$$

или

$$P(A \cdot B \cdot C \cdot D) = 0,2 \cdot 0,3 \cdot 0,25 \cdot 0,1 = 0,0015.$$

Тема 4. Формула полной вероятности.

Формула полной вероятности является следствием теорем умножения и теорем сложения вероятностей.

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

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

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i)$$

где $P(H_i)$ - вероятность i -й гипотезы; $P(A/H_i)$ - вероятность события A при условии реализации этой гипотезы. Данное выражение носит название **формулы полной вероятности**.

Типовые задания.

Пример 1. Имеются 3 одинаковые с виду урны. В первой 4 белых и 2 черных шаров, во второй 3 белых и 1 черный шар, в третьей 1 белый и 1 черный шар. Некто подходит наугад к одной из урн и вынимает один шар. Найти вероятность того, что этот шар белый.

Решение. Пусть событие A состоит в том, что появится белый шар. Рассмотрим три гипотезы: H_1 – выбор шара из первой урны, H_2 – выбор шара из второй урны, H_3 – выбор шара из третьей урны.

Учитывая, что урна выбиралась наугад, находим соответствующие вероятности:

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

Находим условные вероятности изъятия белого шара для каждой урны: $P(A/H_1) = \frac{4}{6}$; $P(A/H_2) = \frac{3}{4}$; $P(A/H_3) = \frac{1}{2}$.

Определяем вероятность события A , пользуясь формулой полной вероятности:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3) = \frac{1}{3} \cdot \frac{4}{6} + \frac{1}{3} \cdot \frac{3}{4} + \frac{1}{3} \cdot \frac{1}{2} = \frac{2}{9} + \frac{1}{6} + \frac{1}{6} = 0,5.$$

Пример 2. На трех автоматических станках изготавливаются одинаковые детали. Известно, что 30 % продукции производится первым станком, 25% – вторым 45% – третьим. Вероятность изготовления детали, отвечающей стандарту, на первом станке равна 0.99, на втором - 0.988 и на третьем - 0.98. Изготовленные в течение дня на трех станках не рассортированные детали находятся на складе. Определить вероятность того, что наудачу взятая деталь не соответствует стандарту.

Решение. Пусть событие A – «наудачу выбранная деталь». Рассмотрим три гипотезы: H_1 – выбранная деталь изготовлена на первом станке, H_2 – выбранная деталь изготовлена на втором станке, H_3 – выбранная деталь изготовлена на третьем станке:

$$P(H_1) = \frac{30}{100} = 0,3; P(H_2) = \frac{25}{100} = 0,25; P(H_3) = \frac{45}{100} = 0,45$$

Находим вероятности нестандартной детали для каждого станка:

$$P(A/H_1) = 1 - 0,99 = 0,01; P(A/H_2) = 1 - 0,988 = 0,012;$$

$$P(A/H_3) = 1 - 0,98 = 0,02.$$

По формуле полной вероятности находим вероятность того, что наудачу взятая деталь не соответствует стандарту:

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3) = 0,3 \cdot 0,01 + 0,25 \cdot 0,012 + 0,45 \cdot 0,02 = 0,015.$$

Тема 5. Теорема гипотез и формула Бернулли

Как следствие формулы полной вероятности и теоремы умножения является теорема гипотез.

Пусть событие A может наступить лишь при условии появления одного из несовместных событий (гипотез) $H_1, H_2, \dots, H_n$, образующих полную группу событий. Если событие A уже произошло, то вероятность гипотез может быть переоценена по **формуле Байеса**:

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{\sum_{i=1}^n P(H_i) \cdot P(A/H_i)},$$

где $i = 1, 2, 3, \dots, n$.

Если производится несколько испытаний, причем вероятность события в каждом испытании не зависит от исхода других испытаний, то такие испытания или опыты называются **независимыми** относительно события A .

На практике часто приходится сталкиваться с вопросами повторения испытаний, например, проверка методики лечения, испытания на надежность и др. В таких задачах необходимо найти вероятность того, что событие A наступит n раз в m испытаниях.

Пусть произведено несколько независимых испытаний, причем вероятность появления события A в каждом из этих испытаний одинакова и равна p . Тогда вероятность того, что событие A не произойдет, будет равна q :

$$q = 1 - p.$$

Если производится n независимых опытов, в каждом из которых событие A появляется с вероятностью p , то вероятность того, что событие A появится m раз, оценивается **формулой Бернулли**:

$$P(A) = C_n^m \cdot p^m \cdot q^{n-m},$$

где C_n^m выражает общее число сложных событий, в которых событие A наступило m раз.

Типовые задания.

Пример 1. Пусть имеется два набора деталей. В первом наборе 12 деталей, из них 9 стандартных, во втором 10 деталей, из них 8 стандартных. Некто выбирает набор наугад и извлекает одну стандартную деталь. Найти вероятность того, что она взята из первого набора.

Решение. Пусть событие A состоит в том, что вынута стандартная деталь. Событие H_1 состоит в том, что некто выбрал первый набор, а событие, H_2 – второй набор. Эти события равновероятны, поэтому

$P(H_1) = P(H_2) = \frac{1}{2}$. Условные вероятности изъятия стандартной детали из первого и второго наборов соответственно равны: $P(A/H_1) = \frac{9}{12} = \frac{3}{4}$; $P(A/H_2) = \frac{8}{10} = \frac{4}{5}$. Определим искомую вероятность по формуле Байеса:

$$P(H_1/A) = \frac{P(H_1) \cdot P(A/H_1)}{\sum_{i=1}^n P(H_i) \cdot P(A/H_i)} = \frac{P(H_1) \cdot P(A/H_1)}{P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2)} = \frac{\frac{1}{2} \cdot \frac{3}{4}}{\frac{1}{2} \cdot \frac{3}{4} + \frac{1}{2} \cdot \frac{4}{5}} = \frac{48}{100} = 0,48.$$

Пример 2. Вероятность рождения мальчика равна 0,815. Определить вероятность того, что из 20 наугад выбранных новорожденных будет 13 мальчиков.

Решение: Вероятность рождения мальчиков $p = 0,815$, тогда $q = 1 - 0,815 = 0,185$. Учитывая, что $n = 20$, а $m = 13$, используя, формулу Бернулли, получим:

$$P(A) = C_n^m \cdot p^m \cdot q^{n-m} = C_{20}^{13} \cdot 0,815^{13} \cdot 0,185^{20-13} = \frac{20!}{13! \cdot 7!} \cdot 0,815^{13} \cdot 0,185^7 = 0,04.$$

Тема 6. Предельные теоремы

Локальная теорема Муавра:

Если вероятность наступления некоторого события в n независимых испытаниях постоянна и равна p ($0 < p < 1$), то вероятность $P_n(m)$ того, что в этих испытаниях событие A наступит ровно m раз, удовлетворяет при $n \rightarrow \infty$ соотношению $\sqrt{npq}P_n(m): \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \rightarrow 1$ равномерно для всех m для которых $x = x_{mn} = \frac{m-np}{\sqrt{npq}}$ находится в каком-либо конечном интервале.

Локальная теорема:

Если вероятности $p_1, p_2, \dots, p_k$ появления соответственно событий $A_1^{(s)}, A_2^{(s)}, \dots, A_k^{(s)}$ в s -ом испытании не зависят от номера и отличны от 0 и от 1 ($0 < p_i < 1$) $i = 1, 2, \dots, k$, то вероятность $P_n(m_1, m_2, \dots, m_k)$ того, что при n независимых испытаниях события $A_i^{(s)}$ $i = 1, 2, \dots, k$ появятся m_i раз ($m_1 + m_2 + \dots + m_k = n$) удовлетворяющих соотношению

$$\sqrt{n^{k-1}}P_n(m_1, m_2, \dots, m_k): \frac{e^{\frac{1}{2}\sum_{i=1}^k q_i x_i^2}}{(2\pi)^{\frac{k-1}{2}} \sqrt{p_1 p_2 \dots p_k}} \rightarrow 1$$

$n \rightarrow \infty$

Равномерно для всех m_i ($i = 1, 2, \dots, k$) для которых $x_i = \frac{m_i - np_i}{\sqrt{np_i q_i}}$ находятся в произвольных конечных интервалах $a_i < x_i < b_i$

Интегральная теорема Муавра-Лапласа:

Если μ есть число наступлений события в независимых испытаниях в каждом из которых вероятность этого события равна p , причем $0 < p < 1$, то равномерно относительно a и b ($-\infty < a < b < +\infty$) при $n \rightarrow \infty$ имеет место соотношение

$$P\left\{a < \frac{\mu - np}{\sqrt{npq}} < b\right\} \rightarrow \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{z^2}{2}} dz$$

Теорема Пуассона:

Если $p_n \rightarrow 0$ при $n \rightarrow \infty$, то $P(\mu_n = m) - \frac{A_n^m}{m!} e^{-A_n} \rightarrow 0$ где $A_n = np_n$

Типовые задания.

Пример 1. Производится n независимых испытаний, при каждом из которых вероятность наступления события A равна p . Чему равна вероятность того, что частота наступления события A отклонится от вероятности p не более чем на a ?

Решение:

$$P\left\{\frac{\mu}{n} < a\right\} = P\left\{-a\sqrt{\frac{n}{pq}} < \frac{\mu - np}{\sqrt{npq}} < a\sqrt{\frac{n}{pq}}\right\} \sim \frac{1}{\sqrt{2\pi}} \int_{-a\sqrt{\frac{n}{pq}}}^{a\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}} dx = \frac{2}{\sqrt{2\pi}} \int_0^{a\sqrt{\frac{n}{pq}}} e^{-\frac{x^2}{2}} dx$$

Пример 2. Вероятность попадания в цель при каждом выстреле равна 0,001. Найти вероятность, попадания в цель двумя и более пулями, если число выстрелов равно 5000.

Решение: Считая каждый выстрел за испытание и попадание в цель – за событие, мы можем для вычисления вероятности $P\{\mu_n \gg 2\} = \sum_{m=2}^{\infty} P_n(m) = 1 - P_n(0) - P_n(1)$

По теореме Пуассона $P_n(0) \approx e^{-5}, P_n(1) \approx 5e^{-5}$. Поэтому $P\{\mu_n \gg 2\} = 1 - 6e^{-5} \approx 0.9596$

Максимальное значение вероятности принимает при $m = 4$ и $m = 5$. Эти вероятности равны с точностью до четвертого знака после запятой

$$P(4) = P(5) = 0.1751$$

Вычисления по точной формуле дают точность до четвертого знака, следовательно $P\{\mu_n \gg 2\} = 0,9575$

РАЗДЕЛ 3. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Тема 7. Дискретная случайная величина

Наряду с понятием случайного события в теории вероятности используется наиболее удобное понятие *случайной величины*.

Случайная величина (СВ) – величина, которая в результате испытания может принимать, то или иное числовое значение, наперед не известное за ранее.

Случайные величины будем обозначать заглавными буквами латинского алфавита $X, Y, Z \dots$, а их возможные значения – соответствующими малыми буквами $x_1, x_2, x_i; y_1, y_2, y_i; z_1, z_2, z_i$.

Случайные величины бывают дискретные и непрерывные.

Дискретная случайная величина (ДСВ) – это величина, которая принимает отдельные, изолированные возможные значения с определенными вероятностями. Число возможных значений ДСВ может быть конечно и бесконечно, но счетно.

Пример. Число родившихся девочек (или мальчиков) из 100 новорожденных, температура воздуха на следующий день и т.д.

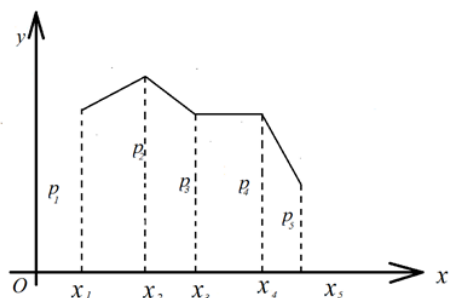
Законом распределения СВ называется соотношение, устанавливающее связь между возможными значениями случайной величины и их вероятностями.

Закон распределения случайной величины можно задать в виде таблицы, формулы или графически. При табличном задании закона распределения в первой строке таблицы перечислены все значения случайной величины в порядке возрастания, а в нижней – соответствующие им вероятности.

X_i	x_1	x_2	x_3	...	x_n
P_i	p_1	p_2	p_3	...	p_n

Такую таблицу будем называть **рядом распределения** случайной величины X . Поскольку все возможные значения можно рассматривать как полную группу несовместных событий, то сумма соответствующих вероятностей обязательно должна быть равна единице (условие нормировки):

$$p_1 + p_2 + p_3 + \dots + p_n = \sum_{i=1}^n p_i = 1$$



Для наглядности ряд распределения случайной величины можно изобразить графически. Для этого в прямоугольной системе координат по оси Ox откладывают значения случайной величины, а по оси ординат Oy – соответствующие им вероятности. Полученные точки соединяются

Рис. 1. Многоугольник или полигон распределения вероятностей

отрезками прямых. Построенная таким образом фигура называется **многоугольником** или **полигоном распределения** вероятностей.

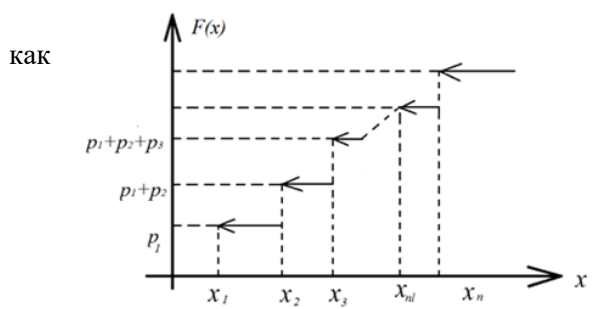


Рис.2 График функции распределения дискретной случайной величины

Многоугольник распределения, также и ряд распределения, полностью характеризует случайную величину. Он является одним из форм закона распределения.

Задать закон распределения дискретной случайной величины можно в виде функции распределения вероятностей (интегральной функции распределения) $F(x)$, где $F(x) = P(X < x) = \sum_{i, x_1 < x} P(X = x_1)$.

График функции распределения дискретной случайной величины имеет ступенчатый вид (рис.2):

Числовые характеристики непрерывных и дискретных случайных величин

Закон распределения (функция распределения и ряд распределения или плотность распределения) полностью описывают поведение случайной величины. Но в ряде задач достаточно знать некоторые числовые характеристики исследуемой величины (например, ее среднее значение и возможное отклонение от него). Рассмотрим основные числовые характеристики случайных величин.

Характеристики положения

Среди числовых характеристик случайных величин нужно прежде всего отметить те, которые характеризуют положение случайной величины на числовой оси, т.е. указывают некоторое среднее, около которого группируются все возможные значения случайной величины. К таким характеристикам относятся математическое ожидание, мода и медиана. Из характеристик положения в теории вероятностей важнейшую роль играет **математическое ожидание** случайной величины, которое иногда называют просто **средним значением** случайной величины.

Рассмотрим дискретную случайную величину X , имеющую возможные значения $x_1, x_2, \dots, x_n$ с вероятностями $p_1, p_2, \dots, p_n$.

Математическим ожиданием $M(x)$ дискретной случайной величины называется сумма произведений всех возможных значений случайной величины на их вероятности:

$$M(x) = x_1 \cdot p_1 + x_2 \cdot p_2 + \dots + x_n \cdot p_n = \sum_{i=1}^n x_i \cdot p_i$$

Свойства математического ожидания

1. Математическое ожидание постоянной величины равно самой постоянной: $M(C) = C$.
2. Константу (постоянную величину) можно вынести за знак математического ожидания: $M(k \cdot x) = k \cdot M(x)$.
3. Математическое ожидание алгебраической суммы конечного числа случайных величин равно сумме их математических ожиданий:

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

4. Математическое ожидание произведения конечного числа случайных величин равно произведению их математических ожиданий:

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

5. Математическое ожидание отклонения случайной величины от ее математического ожидания всегда равно нулю, т.е. $M[x - M(x)] = 0$.

6. Математическое ожидание числа появлений события X в n независимых испытаниях равно произведению числа испытаний на вероятность появления события в каждом испытании, т.е. $M(X) = n \cdot P$.

Кроме важнейшей из характеристик положения - математического ожидания, на практике применяются и другие характеристики положения случайной величины, в частности мода и медиана.

Модой M называют значение случайной величины, имеющее у дискретной величины наибольшую вероятность, а у непрерывной – наибольшую плотность вероятности.

Если кривая распределения имеет один максимум, то мода равна значению случайной величины, соответствующей этому максимуму. Такая кривая называется **унимодальной (одномодальной)**. Если кривая распределения имеет два или несколько одинаковых максимумов, то она соответственно называется **двухмодальной, или многомодальной**.

Медианой случайной величины называют такое ее значение Me , для которого функция распределения равна 0,5. Это означает, что вероятность случайной величины принять значение меньше медианы в точности равна вероятности этой величины принять значение, большее медианы.

Характеристики разброса

Характеристиками рассеяния возможных значений случайной величины вокруг математического ожидания служат дисперсия и среднее квадратичное отклонение.

Дисперсия, есть мера рассеивания или мера разброса случайной величины.

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

$$D(x) = M(x - M(x))^2 \quad \text{или} \quad D(x) = M(x^2) - [M(x)]^2$$

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

Рассмотрим **основные свойства дисперсии**.

1. Дисперсия постоянной величины равна нулю:

$$D(C) = 0.$$

2. Постоянный множитель можно выносить за знак дисперсии, предварительно возведя его в квадрат:

$$D(k \cdot x) = k^2 \cdot D(x)$$

3. Дисперсия алгебраической суммы независимых случайных величин равна сумме дисперсий слагаемых:

$$D(x_1 \pm x_2 \pm \dots \pm x_n) = D(x_1) + D(x_2) + \dots + D(x_n)$$

4. Дисперсия случайной величины равна разности между математическим ожиданием квадрата случайной величины и квадрата ее математического ожидания:

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

Наряду с дисперсией в качестве меры рассеивания широко используется среднее квадратичное отклонение случайной величины x .

Средним квадратическим отклонением случайной величины x называют квадратный корень из дисперсии:

$$\sigma = \sqrt{D(x)}.$$

Дисперсия имеет размерность, равную квадрату размерности случайной величины. Поэтому в тех случаях, когда желательно, чтобы оценка рассеяния имела размерность случайной величины, вычисляют среднее квадратическое отклонение, а не дисперсию. Например, если X выражается в метрах, то $\sigma(X)$ будет выражаться тоже в метрах, а $D(x)$ – в квадратных метрах.

Моменты. Характеристики формы

В теории вероятностей для описания основных свойств, применяют моменты двух видов: начальные и центральные.

Моменты случайных величин служат для описания свойств плотности распределения случайной величины. Моменты содержат меньше информации о случайной величине по сравнению с плотностью распределения, но часто более удобны при решении прикладных задач.

Начальным моментом порядка k случайной величины X называют математическое ожидание величины X^k , т.е.

$$\nu_k = M(X^k).$$

В частности, $\nu_1 = M(X)$, $\nu_2 = M(X^2)$. Следовательно, $D(X) = \nu_2 - \nu_1^2$.

Начальные моменты высших порядков, используются для вычисления центральных моментов.

Центрированной случайной величиной называется отклонение случайной величины X от ее математического ожидания:

$$X = x - M(x)$$

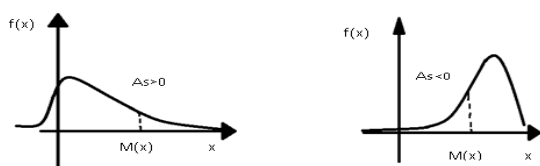
Центральным моментом порядка k случайной величины X называют математическое ожидание величины $(X - M(X))^k$:

$$M_k = M[X - M(X)]^k \text{ или } M_k = \sum_{i=1}^n (x_i - M(x))^k \cdot p_i$$

В частности, $M_1 = 0$, $M_2 = D(X)$. Следовательно, $M_2 = \nu_2 - \nu_1^2$.

Асимметрией называют отношение центрального момента третьего порядка к кубу среднего квадратического отклонения:

$$A_S(X) = \frac{M_3(X)}{[\sigma(X)]^3}.$$



Асимметрия положительна, если “длинная часть” кривой распределения расположена справа от математического ожидания; асимметрия отрицательна, если “длинная часть” кривой расположена слева от математического ожидания.

Практически определяют знак асимметрии по расположению кривой распределения относительно моды (рис.3).

Четвертый центральный момент характеризует «крутость» или плосковершинность (остро) и характеризуется коэффициентом эксцесса.

Четвертый центральный момент характеризует «крутость» или плосковершинность (остро) и характеризуется коэффициентом эксцесса.

Эксцессом называют характеристику, которая определяется равенством:

$$E_k(X) = \frac{M_4(X)}{[\sigma(X)]^4} - 3.$$

Для нормального распределения эксцесс равен нулю.

Следовательно, если эксцесс некоторого распределения отличен от нуля, то кривая этого распределения отличается от нормальной кривой: если эксцесс положительный, то кривая имеет более высокую и “острую” вершину, чем нормальная кривая; если эксцесс отрицательный, то сравниваемая кривая имеет более низкую и “плоскую” вершину, чем нормальная кривая (рис.4).

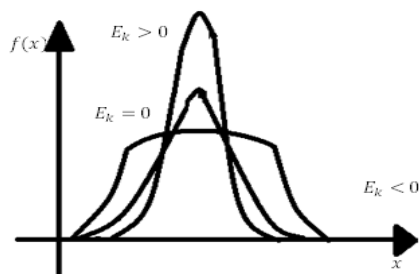


Рис. 4. Эксцесс для нормального распределения случай-

Типовые задания.

Пример 1. Дискретная случайная величина задана законом распределения:

X_i	1	2	3	7	9	10	12
P_i	0,04	0,26	0,31	0,09	0,18	0,11	0,01

Найти математическое ожидание, дисперсию, среднее квадратическое отклонение, моду и медиану случайной величины.

Решение. Найдем сначала математическое ожидание, воспользовавшись формулой:

$$M(x) = \sum_{i=1}^n x_i \cdot p_i = 1 \cdot 0,04 + 2 \cdot 0,26 + 3 \cdot 0,31 + 7 \cdot 0,09 + 9 \cdot 0,18 + 10 \cdot 0,11 + 12 \cdot 0,01 = 4,96$$

Затем, находим дисперсию по формуле:

$$D(x) = M(x^2) - [M(x)]^2 = (1^2 - 4,96^2) + (2^2 - 4,96^2) + (3^2 - 4,96^2) + (7^2 - 4,96^2) + (9^2 - 4,96^2) + (10^2 - 4,96^2) + (12^2 - 4,96^2) = 215,789$$

Находим среднее квадратическое отклонение:

$$\sigma = \sqrt{D(x)} = \sqrt{215,789} = 14,69$$

В этом ряду распределения мода равна $Mo = 3$, т.к. для дискретной случайной величины, мода равна наибольшему значению вероятности величины.

Медиана равна $Me = 7$.

Пример 2. Определить характеристики положения, характеристики разброса, асимметрию и эксцесс дискретной случайной величины, представленной законом распределения:

x_i	0	1	2	3
p_i	0,1	0,3	0,4	0,2

Решение:

Найдем характеристики положения.

Математическое ожидание дискретной случайной величины:

$$M(x) = \sum_{i=1}^n x_i \cdot p_i = 0 \cdot 0,1 + 1 \cdot 0,3 + 2 \cdot 0,4 + 3 \cdot 0,2 = 1,7$$

Мода случайной величины $Mo = X_i(p_{max}) = 2$.

Медиана дискретной случайной величины: $Me = \frac{1+2}{2} = 1,5$.

Найдем характеристики разброса.

Дисперсия дискретной случайной величины:

$$D(x) = \sum_{i=1}^n (x_i - M(x))^2 \cdot p_i = (0 - 1,7)^2 \cdot 0,1 + (1 - 1,7)^2 \cdot 0,3 + (2 - 1,7)^2 \cdot 0,4 + (3 - 1,7)^2 \cdot 0,2 = 0,81.$$

Среднее квадратическое отклонение:

$$\sigma(x) = \sqrt{D(x)} = \sqrt{0,81} = 0,9$$

Определим коэффициент асимметрии:

$$A_s(X) = \frac{M_3(X)}{[\sigma(X)]^3}$$

Найдем сначала третий центральный момент:

$$M_3(X) = \sum_{i=1}^n (x_i - M(x))^3 \cdot p_i = (0 - 1,7)^3 \cdot 0,1 + (1 - 1,7)^3 \cdot 0,3 + (2 - 1,7)^3 \cdot 0,4 + (3 - 1,7)^3 \cdot 0,2 = -0,4913 - 0,1029 + 0,0108 + 0,4394 = -0,144$$

Тогда коэффициент асимметрии будет равен:

$$A_5(X) = \frac{M_3(X)}{[\sigma(X)]^3} = \frac{-0,144}{0,9^3} = -0,198$$

Находим эксцесс:

$$E_k(X) = \frac{M_4(X)}{[\sigma(X)]^4} - 3.$$

Найдем сначала четвертый центральный момент:

$$\begin{aligned} M_4(X) &= \sum_{i=1}^n (x_i - M(x))^4 \cdot p_i \\ &= (0 - 1,7)^4 \cdot 0,1 + (1 - 1,7)^4 \cdot 0,3 + (2 - 1,7)^4 \cdot 0,4 + (3 - 1,7)^4 \cdot 0,2 \\ &= 0,8352 + 0,0720 + 0,0324 + 0,5712 = 1,4817 \end{aligned}$$

Окончательно величина эксцесса будет равна:

$$E_k(X) = \frac{M_4(X)}{[\sigma(X)]^4} - 3 = \frac{1,4817}{0,9^4} - 3 = 2,2583 - 3 = -0,7417.$$

Тема 8. Непрерывная случайная величина

Непрерывная случайная величина (НСВ) – это величина, которая может принимать все значения из некоторого конечного или бесконечного промежутка.

Пример. Расстояние, которое может пролететь стрела.

Для задания случайной величины необходимо знать не только, какие значения может принимать эта величина, но и с какой вероятностью она принимает эти значения. Чтобы задавать вероятности значений случайных величин, разнообразных по своей природе, а также задавать их одним и тем же способом, вводят понятие функции распределения вероятностей случайной величины или просто функции распределения случайной величины.

Функцией распределения $F(x)$ случайной величины X называют функцию $F(x)$, определяющую вероятность того, что случайная величина примет значение, меньшее x :

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

Рассмотрим свойства функции распределения:

1. Функция распределения принимает значения: $0 \leq F(x) \leq 1$, т.к. функция может принимать только те значения, которые принимает вероятность.
2. Функция распределения является неубывающей, т.е. $F(x_2) \geq F(x_1)$, при $x_2 > x_1$.
3. $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow +\infty} F(x) = 1$. В частности, если все возможные значения X лежат в интервале $[a, b]$, то $F(x) = 0$ при $x \leq a$ и $F(x) = 1$ при $x \geq b$. Действительно, $X < a$ – событие невозможное, а $X < b$ – достоверное.
4. Вероятность попадания случайной величины в интервал $[x_1, x_2]$ равна приращению функции распределения на этом интервале, т.е:

$$P(x_1 \leq X < x_2) = F(x_2) - F(x_1)$$

Определение и свойства функции распределения сохраняются и для непрерывной случайной величины, для которой функцию распределения можно считать одним из видов задания закона распределения.

Еще одним способом задания закона распределения непрерывной случайной величины является так называемая плотность распределения (плотность вероятности, дифференциальная функция).

Плотность распределения случайной величины, есть производная от функции распределения $F(x)$:

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

Рассмотрим **свойства плотности распределения**:

1. Плотность вероятности – неотрицательная функция, т. е: $f(x) \geq 0$.
2. Вероятность попадания непрерывной случайной величины в интервал $[a, b]$ равна определенному интегралу от ее плотности вероятности, определяется по формуле:

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

3. Несобственный интеграл в бесконечных пределах от плотности распределения случайной величины равен 1, т.е.

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

Свойство №3 называется **условием нормировки непрерывной случайной величины**.

Для того, чтобы вычислить математическое ожидание непрерывной случайной величины используют формулу:

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

Характеристики положения.

Для **непрерывной случайной величины медиана** определяется из соотношения

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

Геометрически медиана представляет собой абсциссу точки, которая делит площадь, ограниченную кривой распределения пополам.

Если кривая распределения имеет один максимум, то мода равна значению случайной величины, соответствующей этому максимуму. Такая кривая называется унимодальной (одномодальной). Если кривая распределения имеет два или несколько случайной величины одинаковых максимумов, то она соответственно называется двухмодальной, или многомодальной.

Характеристики разброса.

Дисперсию непрерывной случайной величины можно определить, воспользовавшись выражением:

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

Средним квадратическим отклонением случайной величины x называют квадратный корень из дисперсии:

$$\sigma = \sqrt{D(x)}.$$

Характеристики формы.

Для непрерывных случайных величин центральный момент k - го порядка определяется выражением:

$$M_k = \int_{-\infty}^{\infty} (x - M(x))^k \cdot f(x) \cdot dx.$$

Типовые задания.

Пример 1. Плотность распределения случайной величины имеет следующий вид:

$$f(x) = \begin{cases} a \cdot \cos x, & 0 \leq x \leq \frac{\pi}{2} \\ 0, & x < 0, x > \frac{\pi}{2} \end{cases}.$$

Найти постоянный параметр a .

Решение: Плотность распределения должна удовлетворять условию нормировки:

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

Поэтому потребуем, чтобы выполнялось равенство

$$\int_0^{\frac{\pi}{2}} a \cdot \cos x \cdot dx = 1.$$

Отсюда $a = \frac{1}{\int_0^{\frac{\pi}{2}} \cos x \cdot dx}$. Найдём определённый интеграл:

$$\int_0^{\frac{\pi}{2}} \cos x \cdot dx = \sin x \Big|_0^{\frac{\pi}{2}} = \sin \frac{\pi}{2} - \sin 0 = 1 - 0 = 1.$$

Таким образом, искомый параметр: $a = 1$.

Следовательно, функция плотности вероятности примет вид:

$$f(x) = \begin{cases} \cos x, & 0 \leq x \leq \frac{\pi}{2} \\ 0, & x < 0, \quad x > \frac{\pi}{2} \end{cases}$$

Пример 2. Найти математическое ожидание, дисперсию и среднее квадратическое отклонение непрерывной случайной величины, заданной функцией распределения:

$$F(x) = \begin{cases} 0, & x \leq 1, \\ \frac{x^2 - x}{2}; & 1 \leq x \leq 2, \\ 1, & x \geq 2. \end{cases}$$

Решение. Найдём дифференциальную функцию распределения:

$$f(x) = F'(x) = \begin{cases} 0, & x \leq 1, \\ \frac{2x - 1}{2}; & 1 \leq x \leq 2, \\ 0, & x \geq 2. \end{cases}$$

Определяем математическое ожидание

$$M(x) = \int_1^2 x \cdot \left(\frac{2x-1}{2}\right) \cdot dx = \int_1^2 \frac{2x^2 - x}{2} \cdot dx = \frac{x^3}{3} \Big|_1^2 - \frac{x^2}{4} \Big|_1^2 = \left(\frac{8}{3} - \frac{1}{3}\right) - \left(\frac{4}{4} - \frac{1}{4}\right) = \frac{7}{3} - \frac{3}{4} = \frac{19}{12}.$$

Дисперсия непрерывной величины определяется выражением:

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

Подставляя численные значения, получаем:

$$D(x) = \int_1^2 x^2 \cdot \left(\frac{2x-1}{2}\right) \cdot dx - \left[\frac{19}{12}\right]^2 = \int_1^2 \frac{x^3}{2} \cdot dx - \frac{1}{2} \int_1^2 x^2 \cdot dx - \left[\frac{19}{12}\right]^2 = \frac{x^4}{4} \Big|_1^2 - \frac{x^3}{6} \Big|_1^2 - \left[\frac{19}{12}\right]^2 = \left(\frac{16}{4} - \frac{1}{4}\right) - \left(\frac{8}{6} - \frac{1}{6}\right) - \frac{361}{144} = \frac{15}{4} - \frac{7}{6} - \frac{361}{144} = \frac{372-361}{144} = \frac{11}{144}.$$

Среднее квадратическое отклонение равно: $\sigma = \sqrt{D(x)} = \sqrt{\frac{11}{144}} \approx 0,276$.

Тема 9. Основные законы распределения случайных величин

Наиболее полно случайная величина может быть охарактеризована законом распределения случайной величины в виде функции распределения или плотности распределения.

В зависимости от характера самих объектов имеют место следующие наиболее распространённые законы распределения случайных величин:

- 1) биномиальный закон распределения;
- 2) распределение Пуассона;
- 3) геометрическое распределение;

- 4) нормальный закон распределения (закон Гаусса);
- 5) показательное распределение;
- 6) распределение «хи-квадрат»;
- 7) распределение Стьюдента;
- 8) распределение Фишера-Снедекора.

Биномиальное распределение

В некоторых случаях вероятности значений дискретных случайных величин подчиняются определенным законам и могут быть заданы *аналитически* (в виде формул).

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

Рассмотрим последовательность n идентичных повторных испытаний, удовлетворяющих следующим условиям, которые носят название схемы Бернулли:

- 1). Все n испытания – независимы. Это значит, что вероятность наступления события в любом из n повторных испытаний не зависит от результатов других испытаний.
- 2). В каждом испытании может наступить или не наступить некоторое событие A , вероятность наступления которого p остается неизменной.
- 3). Противоположное событие $\bar{A}$ имеет вероятность $q = 1 - p$, которая тоже остается неизменной от испытания к испытанию.
- 4). Эти события A и $\bar{A}$ - взаимно несовместные и противоположные, называемые успех и неуспех.

Тогда для вычисления вероятности того, что в n независимых повторных испытаниях, удовлетворяющих условиям 1-4, событие A наступит ровно m раз, при $X = m = 0, 1, 2, \dots, n$ (в любой последовательности), вычисляется по формуле Бернулли:

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

где p – вероятность успеха в каждом испытании; $q = 1 - p$ - вероятность неуспеха в каждом испытании; C_n^m - число сочетаний из n по m .

Так как правая часть формулы Бернулли представляет общий член биномиального разложения $(q + p)^n$, то этот закон распределения называют *биномиальным*.

Используя формулу Бернулли, можно построить ряд распределения дискретной случайной величины, подчиняющейся биномиальному закону распределению, т.е., удовлетворяющей условиям 1-4. Вероятность того, что случайная величина X принимает значение, равное k , определяется по формуле

$$P(X = k) = C_n^k p^k q^{n-k}.$$

Биномиальный закон задается рядом распределения, который имеет вид, аналогичный приведенному в таблице:

Число успехов $X=m$	0	1	2	...	m	...	n
Вероятность $P_{n,m}$	$C_n^0 p^0 q^n = P_{n,0}$	$C_n^1 p^1 q^{n-1} = P_{n,1}$	$C_n^2 p^2 q^{n-2} = P_{n,2}$	...	$C_n^m p^m q^{n-m} = P_{n,m}$	...	$C_n^n p^n q^0 = P_{n,n}$

Распределение Пуассона

Дискретную случайную величину называют *Пуассоновской*, если ее закон распределения имеет следующий вид:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \text{ где } k = \overline{0, \infty} \text{ и } \lambda = \text{const (постоянное значение)}.$$

Придавая k целые неотрицательные значения $k = 0, 1, 2, \dots, n$, можно записать ряд распределения вероятностей, который называется *законом распределения Пуассона*.

m	0	1	2	...	m	...	n
$P_{n,m}$	$e^{-\lambda}$	$\lambda e^{-\lambda}$	$\frac{\lambda^2 e^{-\lambda}}{2!}$	...	$\frac{\lambda^m e^{-\lambda}}{m!}$	...	$\frac{\lambda^n e^{-\lambda}}{n!}$

Примеры Пуассоновских случайных величин:

Число вызовов на автоматическую станцию за промежуток времени T .

Число частиц распада некоторого радиоактивного вещества за промежуток времени T .

Число телевизоров, которые поступают в мастерскую за промежуток времени T в большом городе.

Число автомобилей, которые поступят к стоп-линии перекрестка в большом городе.

Замечание: В сериях независимых испытаний (когда n велико, p мало) для вычисления вероятности наступления события ровно k раз используют формулу Пуассона:

$P_n(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$, где $\lambda = np$, то есть среднее число появлений событий остается постоянным.

Геометрическое распределение

Пусть производятся независимые испытания, в каждом из которых вероятность появления события A равна p ($0 < p < 1$) и, следовательно, вероятность его «непоявления» $q = 1 - p$. Испытания заканчиваются, как только появится событие A . Таким образом, если событие A появилось в k -м испытании, то в предшествующих $k - 1$ испытаниях оно не появлялось.

Обозначим через X дискретную случайную величину – число испытаний, которые нужно провести до первого появления события A . Очевидно, возможными значениями X являются натуральные числа $x_1 = 1, x_2 = 2, \dots$. Пусть в первых $k - 1$ испытаниях событие A не наступило, а в k -м испытании появилось. Вероятность этого «сложного события», по теореме умножения вероятностей независимых событий, $P(X = k) = q^{k-1} p$.

Дискретная случайная величина имеет *геометрическое распределение*, если ее закон распределения имеет следующий вид:

$$P(X = k) = q^{k-1} p, \text{ где } k = \overline{1, \infty}.$$

Замечание 1: Полагая $k = 1, 2, \dots$, получим геометрическую прогрессию с первым членом p и знаменателем q ($0 < q < 1$). По этой причине распределение называют геометрическим.

Замечание 2: Ряд $\sum_{k=1}^{\infty} q^{k-1} p$ сходится и сумма его равна единице. Действительно

$$\text{сумма ряда равна } \frac{p}{(1-q)} = \frac{p}{p} = 1.$$

Нормальный закон распределения

Нормальный закон распределения является самым распространенным из законов распределения, наиболее часто встречающимся в случайных явлениях природы. Объясняется это тем, что, как правило, случайная величина принимает те или иные значения под воздействием большого числа взаимно независимых малых причин, другими словами, *является суммой очень большого числа взаимно независимых случайных величин*, малых по сравнению со всей суммой.

Нормальный закон распределения задается дифференциальной функцией:

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

Параметры $a = M(X)$ и $\sigma^2 = D(X)$ являются соответственно математическим ожиданием и средним квадратическим отклонением случайной величины X , распределенной по нормальному закону.

В этом случае говорят, что X имеет нормальное распределение с параметрами a и σ^2 .

Таким образом, если известны математическое ожидание и дисперсия случайной величины и известно, что она подчиняется нормальному закону, то известен полностью закон распределения этой величины – дифференциальная функция распределения. Нормальный закон распределения иногда называют законом Гаусса. График плотности нормального распределения называют *нормальной кривой* или *кривой Гаусса*.

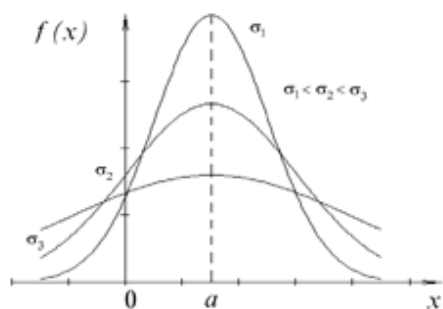


Рис. 1 График нормального закона распределения

Максимум функции плотности вероятностей

нормального распределения равен $\frac{1}{\sigma\sqrt{2\pi}}$. Отсю-

да следует, что с возрастанием σ максимальная ордината нормальной кривой убывает, а сама кривая становится более пологой, то есть сжимается к оси Ox ; при убывании σ нормальная кривая становится более «островершинной» и растягивается в положительном направлении оси Oy :

Практически достоверно, что величина, распределенная по нормальному закону, отклоняется от математического ожидания не более чем на утроенное среднее квадратическое отклонение, т.е.

$$P\{|x - a| < 3\sigma\} \approx 1.$$

Таким образом, имеет место, так называемое, *правило «трех сигм»*: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения.

Серединным, или вероятным, отклонением случайной величины называют такое число a_0 , что вероятности неравенств

$$|x - M(x)| < a_0 \text{ и } |x - M(x)| \geq a_0$$

одинаковы (т. е. равны по $\frac{1}{2}$ каждая).

Вероятность попадания в заданный интервал нормальной случайной величины

Пусть случайная величина X распределена по нормальному закону. Тогда вероятность того, что X примет значение, принадлежащее интервалу (α, β) , равна:

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-(x-a)^2/(2\sigma^2)} dx.$$

Введем новую переменную $z = (x-a)/\sigma$. Отсюда, $x = \sigma z + a$, $dx = \sigma dz$. Найдем новые пределы интегрирования. Если $x = \alpha$, то $z = (\alpha-a)/\sigma$; если $x = \beta$, то $z = (\beta-a)/\sigma$.

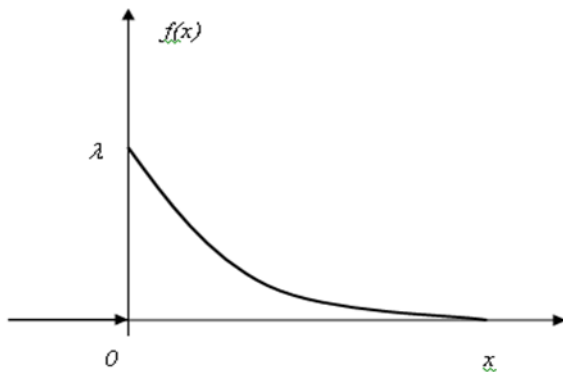
Таким образом, имеем

$$\begin{aligned} P(\alpha < X < \beta) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-z^2/2} (\sigma dz) = \frac{1}{\sqrt{2\pi}} \int_{(\alpha-a)/\sigma}^0 e^{-z^2/2} dz + \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-z^2/2} dz - \frac{1}{\sqrt{2\pi}} \int_0^{(\alpha-a)/\sigma} e^{-z^2/2} dz. \end{aligned}$$

Пользуясь функцией Лапласа

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

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



Показательное распределение

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

Рис. 2 График функции показательного распределения

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

где λ – постоянная положительная величина.

Например, время безотказной работы компьютерной системы есть случайная величина, имеющая показательное распределение с параметром λ , физический смысл которого – среднее число отказов в единицу времени. Интервал между последовательными поступлениями вызовов на автоматическую телефонную станцию, интервал между посещениями врача-педиатра – это примеры показательных случайных величин.

Интегральная функция распределения показательного закона имеет вид:

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

Распределение «Хи-квадрат»

Пусть $x_1, x_2, x_3, \dots, x_n$ – нормально распределенные независимые случайные величины, причем математическое ожидание каждой из них равно нулю, а дисперсия равна единице, т.е. закон распределения, имеет вид:

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

Возьмем сумму квадратов этих случайных величин:

$$\chi^2 = \sum_{i=1}^n x_i^2.$$

Принято считать, что сумма квадратов этих величин распределена по закону «хи-квадрат» с $k = n$ степенями свободы.

Плотность распределения χ^2 с $k = n$ степенями свободы имеет вид:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \cdot e^{-\frac{x}{2}} \cdot x^{\frac{k}{2}-1}, & x > 0, \end{cases}$$

где $\Gamma(x) = \int t^{x-1} e^{-t} dt$ – гамма-функция.

Распределение χ^2 определяется одним параметром – числом степеней свободы k . С ростом числа степеней свободы, распределение χ^2 достаточно медленно приближается к нормальному распределению.

Распределение Стьюдента

Распределение Гаусса хорошо описывает отклонение измеряемой величины в эксперименте при большом количестве повторных измерений. На практике нередко обходятся двумя-тремя измерениями. Результаты образуют очень малую выборку. В этом случае распределение Стьюдента гораздо лучше описывает получаемые результаты.

Распределение Стьюдента – это непрерывное одномерное распределение с одним параметром – количеством степеней свободы. Форма распределения Стьюдента похожа на форму нормального распределения (чем больше число степеней свободы, тем ближе распределение к нормальному). Отличием является то, что хвосты распределения Стьюдента медленнее стремятся к нулю, чем хвосты нормального распределения.

Пусть Z – нормально распределенная случайная величина с математическим ожиданием равным нулю и дисперсией равной единице. Пусть V – независимая случайная величина, которая

распределена по закону χ^2 с k степенями свободы. Тогда величина $t = \frac{Z}{\sqrt{\frac{V}{k}}}$, имеет распреде-

ление, называемое t – распределением или распределением Стьюдента с k степенями свободы. Плотность распределения Стьюдента имеет вид:

$$f(x) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \cdot \Gamma\left(\frac{k}{2}\right)} \cdot \left(1 + \frac{x^2}{k}\right)^{-\frac{k+1}{2}}.$$

С ростом числа степеней свободы распределение Стьюдента достаточно быстро приближается к нормальному распределению.

Таблицы распределения Стьюдента позволяют при данном числе степеней свободы k по вероятности q определить значение t_q , для которого выполняется соотношение:

$$P(|t| > t_q) = q.$$

Распределение Фишера-Снедекора

Пусть U и V – независимые случайные величины, распределенные по закону χ^2 со степенями свободы k_1 и k_2 . Тогда величина $F = \frac{U/k_1}{V/k_2}$ имеет распределение, которое называется распределением Фишера-Снедекора со степенями свободы k_1 и k_2 .

Плотность распределения Фишера-Снедекора со степенями свободы k_1 и k_2 имеет вид:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{\Gamma\left(\frac{k_1+k_2}{2}\right) \cdot k_1^{\frac{k_1}{2}} \cdot k_2^{\frac{k_2}{2}}}{\Gamma\left(\frac{k_1}{2}\right) \cdot \Gamma\left(\frac{k_2}{2}\right)} \cdot \frac{x^{\frac{k_1-2}{2}}}{(k_2+k_1x)^{\frac{k_1+k_2}{2}}}, & x > 0 \end{cases}$$

Типовые задания.

Пример 1. Случайная величина X распределена по нормальному закону с $M(X)=30$ и $\sigma(X)=10$. Найти вероятность того, что случайная величина X примет значение, принадлежащее интервалу $(10,50)$.

Решение: $P(10 < X < 50) = \Phi\left(\frac{50-30}{10}\right) - \Phi\left(\frac{10-30}{10}\right) = 2\Phi(2)$.

По таблице значений функции Лапласа находим $\Phi(2) = 0,4772$. Отсюда искомая вероятность:

$$P(10 < X < 50) = 2 \cdot 0,4772 = 0,9544.$$

Пример 2. Найти симметричный интервал, в который случайная величина, распределенная по закону Стьюдента с 12-ю степенями свободы, попадает вероятностью 0,9.

Решение: Очевидны соотношения:

$$P(-x < t < x) = P(|t| < x) = 1 - P(|t| \geq x) = 0,9$$

Из последнего равенства следует:

$$P(|t| \geq x) = 0,1 \quad (n = 12).$$

Определяем из таблицы IV в приложении: $x = 1,782$.

РАЗДЕЛ 4. СТАТИСТИЧЕСКИЙ АНАЛИЗ РЕЗУЛЬТАТОВ ИССЛЕДОВАНИЙ

Тема 10. Основные понятия математической статистики

Математическая статистика – это наука, изучающая закономерности, которым подчиняются случайные массовые явления на основе результатов наблюдений.

Результаты наблюдений (измерений) называют **статистическими данными**.

Основные задачи математической статистики:

- 1) Задача определения распределения закона случайной величины или системы случайных величин по статистическим данным.
- 2) Задача проверки правдоподобия гипотез.

3) Задача нахождения неизвестных параметров распределения.

Итак, основная задача математической статистики состоит в создании методов сбора и обработки статистических данных для получения научных и практических выводов.

Рассмотрим основные понятия математической статистики.

Генеральной совокупностью называется совокупность всех исследуемых объектов.

На практике, обследование всех объектов генеральной совокупности не производится, так как этот процесс излишне трудоемкий. Обычно из всей совокупности отбирают ограниченное число объектов, которые подвергают изучению.

Выборочной совокупностью или **выборкой** называется совокупность объектов, отобранных случайным образом из генеральной совокупности.

Объем генеральной совокупности N и объем выборки n – число объектов в рассматриваемой совокупности.

Например. Если из 2000 людей, болеющих каким-то заболеванием, отобрано для обследования 150 человек, то объем генеральной совокупности $N = 2000$, а объем выборки $n = 150$.

Для того, чтобы по данным выборки можно было уверенно судить об интересующих признаках генеральной совокупности, выборка должна правильно представлять пропорции генеральной совокупности. Данное требование формируется так: выборка должна быть представительной или репрезентативной.

Способы составления выборки:

- повторный;
- бесповторный.

Повторной называется выборка, при которой, отобранный объект возвращается в генеральную совокупность перед отбором следующего.

Бесповторной называется выборка, при которой объект не возвращается в генеральную совокупность.

Дополнительные способы отбора выборки

1) **простой случайный отбор** – отбор, при котором объекты извлекаются по одному из всей генеральной совокупности.

2) **типичность, типический отбор** – это отбор, при котором объекты извлекаются не из всей генеральной совокупности, а из каждой ее «типической части».

3) **механический отбор** – генеральная совокупность механически делятся на столько групп, сколько объектов должно войти в выборку, а из каждой группы выбирается один объект.

4) **серийный отбор** – отбор, при котором объекты из генеральной совокупности выбирают не по одному, а сериями, которые подвергаются статистическому отбору.

Пусть имеется общая генеральная совокупность N , извлечена выборка $x_1, x_2, \dots, x_k$ объемом n .

x_1	x_2	...	x_i
n_1	n_2	...	n_i

Наблюдаемые значения x_i называют **вариантами**, а последовательность вариантов записанных в возрастающем порядке называют **вариационным рядом**.

Число наблюдений варианты n_i называют **частотой встречаемости**. Отношение частоты встречаемости к объему выборки называют **относительной частотой**: $\omega = \frac{n_i}{n}$.

Статистическим распределением выборки называется перечень вариантов и соответствующих им частот или относительных частот.

Пример 1. Выборка задана в виде распределения частот:

x_i	4	7	8	12
n_i	5	2	3	10

Найти распределения относительных частот.

Решение. Найдем объем выборки: $n = 5 + 2 + 3 + 10 = 20$. Найдем относительные частоты, для этого разделим частоты на объем выборки:

$$\omega_1 = \frac{n_1}{n} = \frac{5}{20} = 0,25; \omega_2 = \frac{n_2}{n} = \frac{2}{20} = 0,1; \omega_3 = \frac{n_3}{n} = \frac{3}{20} = 0,15;$$

$$\omega_4 = \frac{n_4}{n} = \frac{10}{20} = 0,5.$$

Запишем распределение относительных частот:

x_i	4	7	8	12
ω_i	0,25	0,1	0,15	0,5

Контроль: $0,25 + 0,1 + 0,15 + 0,5 = 1$.

Для наглядного представления статистического распределения пользуются графическими изображениями. Они помогают лучше представить характер распределения элементов выборки, а иногда сделать предварительные предположения о законе распределения генеральной совокупности. Такими графиками являются полигон частот и гистограмма.

Полигон частот – ломаная линия, отрезки которой соединяют точки с координатами $(x_1, n_1), (x_2, n_2), \dots, (x_i, n_i)$.

Для построения полигона частот на оси абсцисс откладываются варианты x_i , а по оси ординат – соответствующие им частоты n_i . Построенные точки с координатами (x_i, n_i) соединяют отрезками.

Для непрерывного признака графической иллюстрацией служит гистограмма.

Гистограммой относительных частот называют ступенчатую фигуру, состоящую из смежных прямоугольников, основания которых одинаковы и равны ширине класса, а высота равна частоте попадания в интервал n_i или относительной частоте ω_i .

Ширину интервала можно определить по формуле Стерджеса:

$$k = \frac{x_{\max} - x_{\min}}{1 + 3,32 \cdot \lg n},$$

где $x_{\max}$ – максимальное, а $x_{\min}$ – минимальное значение вариантов, n – объем статистической совокупности.

Выборные характеристики. Характеристики положения.

Модой (Mo) называется наиболее вероятное значение случайной величины или то значение этой величины, частота которого наибольшая.

Например, распределение лейкоцитов по числу поглощенных ими бактерий имеет вид:

число погл. бакт.	0	1	2	3	4	5	6	7	8	9	10
кол-во лейкоцитов	2	10	12	35	40	91	108	135	240	180	99

Наибольшая частота 240 отвечает признаку, равному 8, следовательно, $Mo = 8$.

Для определения моды интервальных рядов служит формула:

$$Mo = x_{\text{ниж}} + i \cdot \left(\frac{n_2 - n_1}{2n_2 - n_1 - n_3} \right),$$

где $x_{\text{ниж}}$ – нижняя граница модального класса, т.е. класса с наибольшей частотой встречаемости n_2 ; n_2 – частота модального класса; n_1 – частота класса, предшествующего модальному; n_3 – частота класса, следующего за модальным; i – ширина классового интервала.

Медиана Me – это такое среднее значение, которое делит совокупность величин x_i на две равные по количеству членов части, причем в одной из них все значения x_i меньше медианы, а в другой – больше.

Выборочная средняя – это среднее арифметическое значение вариантов статистического ряда: $\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i$.

Характеристики рассеяния вариантов вокруг своего среднего

Выборочная дисперсия – среднее арифметическое квадратов отклонения полученных значений признака от их среднего значения.

$$S_B^2 = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot n_i$$

Среднее квадратическое отклонение (стандартное отклонение) – это квадратный корень из выборочной дисперсии:

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

Коэффициент вариации CV – это отношение среднего квадратического отклонения к средней величине признака, выраженное в процентах:

$$CV = \frac{S_B}{\bar{x}_B} \cdot 100\%.$$

Коэффициент вариации – это мера относительной изменчивости случайной величины, например, частоту сердечных сокращений (ЧСС, уд/мин), артериальное давление (АД, мм. рт. ст.) и температуру (t °C) в единых единицах – процентах.

Вариационный размах $\Delta = x_{\max} - x_{\min}$ – это разность между наибольшим и наименьшим значениями признака.

Типовые задания

Пример 1. В результате измерения физической величины одним прибором (без систематических ошибок) получили следующие результаты: 9, 11, 13, 15, 17. Найти выборочную дисперсию ошибок прибора.

Решение. Найдем выборочную среднюю:

$$\bar{x}_B = \frac{9 \cdot 1 + 11 \cdot 1 + 13 \cdot 1 + 15 \cdot 1 + 17 \cdot 1}{5} = \frac{65}{5} = 13$$

Найдем выборочную дисперсию:

$$S_B^2 = \frac{(9-13)^2 + (11-13)^2 + (13-13)^2 + (15-13)^2 + (17-13)^2}{5} = \frac{40}{5} = 8.$$

Пример 2. Сравните 2 варьирующихся признака. Один характеризуется средней $\bar{x}_1 = 46,8$ и средним квадратическим отклонением $S_1 = 4,85$, другой – величинами $\bar{x}_2 = 84,6$ и $S_2 = 5,7$. Какой признак варьируется сильнее?

Решение. Найдем коэффициенты вариации для двух признаков по формуле: $CV = \frac{S_B}{\bar{x}_B} \cdot 100\%$.

$$CV_1 = \frac{4,85}{46,8} \cdot 100\% = 10,36\%, \quad CV_2 = \frac{5,7}{84,6} \cdot 100\% = 6,74\%$$

Так как $CV_1 > CV_2$, то первый признак варьируется сильнее.

Тема 11. Оценка параметров генеральной совокупности по ее выборке

Смысл статистических методов заключается в том, чтобы по выборке ограниченного объема n , т.е. по некоторой части генеральной совокупности, высказать обоснованное суждение о ее свойствах в целом.

Числовые значения, характеризующие генеральную совокупность, называются **параметрами**. Одна из задач математической статистики – определение параметров большого массива по исследованию его части.

Статистическое оценивание может выполняться двумя способами:

- 1) Точечная оценка – оценка, которая дается для некоторой определенной точки.
- 2) Интервальная оценка – по данным выборки оценивается интервал, в котором лежит истинное значение с заданной вероятностью.

Точечная оценка параметров генеральной совокупности

Статистической оценкой θ^* неизвестного параметра θ теоретического распределения называют функцию $f(X_1, X_2, \dots, X_n)$ от наблюдаемых случайных величин $X_1, X_2, \dots, X_n$.

Точечной называют статистическую оценку, которая определяется одним числом $\Theta^* = f(x_1, x_2, \dots, x_n)$, где $x_1, x_2, \dots, x_n$ – результаты n наблюдений над количественным признаком X (выборка).

Качество оценки устанавливается по трем свойствам: быть *состоятельной*, *эффективной* и *несмещенной*.

Точечная оценка называется *состоятельной*, если при увеличении объема выборки выборочная характеристика стремится к соответствующей характеристике генеральной совокупности.

Точечная оценка называется *эффективной*, если она имеет наименьшую дисперсию выборочного распределения по сравнению с другими аналогичными оценками.

Точечную оценку называют *несмещенной*, если ее математическое ожидание равно оцениваемому параметру при любом объеме выборки.

Несмещенной оценкой генеральной средней (математического ожидания) служит выборочная средняя $\bar{x}_B$:

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i$$

x_i – варианты выборки; n_i – частота встречаемости варианты x_i ; n – объем выборки.

Выборочная средняя является несмещенной оценкой генеральной средней, т.к. $M(\bar{x}_B) = \bar{x}_{\text{ген}}$.

Выборочная дисперсия S_B^2 не обладает свойством несмещенности. Это смещенная оценка генеральной дисперсии $\sigma_{\text{ген}}^2$.

$M(S_B^2) = \frac{n-1}{n} \cdot \sigma_{\text{ген}}^2 \neq \sigma_{\text{ген}}^2$ – это и означает, что выборочная дисперсия S_B^2 является смещенной оценкой $\sigma_{\text{ген}}^2$.

На практике используют исправленную выборочную дисперсию S^2 , которая является несмещенной оценкой дисперсии генеральной совокупности:

$$S^2 = \frac{n-1}{n} \cdot S_B^2(x) = \frac{n}{n-1} \cdot \frac{\sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot n_i}{n},$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^k (x_i - \bar{x}_B)^2 \cdot n_i.$$

Кроме того, в расчетах используют S – исправленное среднее квадратическое отклонение ($S = \sqrt{S^2}$), называемое стандартным отклонением, и ошибку выборочной средней (стандартную ошибку среднего) $m_{\bar{x}}$:

$$m_{\bar{x}} = \frac{S}{\sqrt{n}}.$$

Интервальная оценка параметров генеральной совокупности

В ряде случаев требуется не только найти для исследуемого параметра подходящее численное значение, но и оценить его надежность и точность. Требуется знать – к каким ошибкам может привести замена параметра его точечной оценкой, и с какой уверенностью можно ожидать, что ошибки не выйдут за известные пределы. Чтобы дать представление о точности и надежности оценки параметров, в математической статистике прибегают к интервальным оценкам.

Интервальная оценка – это числовой интервал, который определяется двумя числами – границами интервала, содержащий неизвестный параметр генеральной совокупности.

Доверительный интервал – это интервал значений признака, который с известной вероятностью содержит оцениваемый параметр генеральной совокупности.

Доверительная вероятность p – это вероятность того, доверительный интервал содержит значение оцениваемого параметра. Доверительную вероятность принято устанавливать на уровнях 90%, 95% и 99%. Эти вероятности признаны достаточными для

уверенного суждения о генеральных параметрах на основании известных выборочных показателей.

Кроме доверительной вероятности часто используют «противоположное» значение – *уровень значимости*: $\alpha = 1 - p$, который выражает вероятность непопадания генеральной средней в доверительный интервал.

Для нормального закона распределения, зная величину выборочной средней и ее ошибку, можно определить границы, в которых с той или иной вероятностью находится параметр генеральной совокупности – математическое ожидание. Эти границы называются доверительными и определяются по формуле:

$$\bar{x}_B - \frac{S}{\sqrt{n}} \cdot t_{\alpha,f} < \mu < \bar{x}_B + \frac{S}{\sqrt{n}} \cdot t_{\alpha,f},$$

где $t_{\alpha,f}$ – величина нормированного отклонения, определяемая по таблицам распределения Стьюдента. Величина $t_{\alpha,f}$ определяется вероятностью попадания генерального параметра в указанный интервал и числом степеней свободы $f = n - 1$, где n – объем выборки.

Так, при $n = 30$, $f = 30 - 1 = 29$ для $p = 0,95$, $t_{0,05;29} = 2,05$; для $p = 0,99$, $t_{0,01;29} = 2,76$, для $p = 0,999$, $t_{0,001;29} = 3,66$.

Из приведенных данных видно, что с увеличением доверительной вероятности границы интервала раздвигаются, т.е. увеличивается надёжность попадания параметра в указанный интервал, но уменьшается точность его определения. Поэтому для сохранения высокой точности доверительной вероятности и повышения точности доверительной оценки необходимо увеличить объем выборки.

Типовые задания.

Пример 1. Найти несмещенную оценку генеральной средней, дисперсии генеральной совокупности и стандартное отклонение по выборке объема 10, описывающую вес девочек: 27; 24; 23; 27; 25; 28; 26; 23; 24; 25.

Решение: Несмещенной оценкой генеральной средней служит выборочная средняя:

$$\bar{x}_B = \frac{1}{n} \sum_{i=1}^k x_i \cdot n_i = \frac{1}{10} \cdot (27 \cdot 2 + 24 \cdot 2 + 23 \cdot 2 + 25 \cdot 2 + 28 \cdot 1 + 26 \cdot 1) = 25,2.$$

Несмещенная оценка дисперсии генеральной совокупности равна исправленной дисперсии:

$$S^2 = \frac{1}{9} \cdot [(27 - 25,2)^2 \cdot 2 + (24 - 25,2)^2 \cdot 2 + (23 - 25,2)^2 \cdot 2 + (25 - 25,2)^2 \cdot 2 + (28 - 25,2)^2 + (26 - 25,2)^2] = 3,08.$$

Определим стандартное отклонение по формуле $S = \sqrt{S^2}$, т.е. $S = \sqrt{3,08} = 1,75$.

Пример 2. У 20 студентов, сдававших экзамен, сердце билось в среднем со скоростью 96 ударов в минуту. Стандартное отклонение выборки было равно 5 ударам в минуту. Найти доверительный интервал для генерального среднего при $p \geq 0,95$.

Решение: По условию задачи $\bar{x}_B = 96$; $n = 20$; $S = 5$. Доверительный интервал определяем по формуле: $\bar{x}_B - \frac{S}{\sqrt{n}} \cdot t_{\alpha,f} < \mu < \bar{x}_B + \frac{S}{\sqrt{n}} \cdot t_{\alpha,f}$. Найдем $t_{\alpha,f}$ из таблицы распределения Стьюдента (см. приложение, таблица IV) при уровне значимости $\alpha \leq 0,05$ и числе степеней свободы $f = n - 1$;

$f = 20 - 1 = 19$, $t(\alpha \leq 0,05, f = 19) = 2,093$. Тогда

$$96 - \frac{5}{\sqrt{20}} \cdot 2,093 < \mu < 96 + \frac{5}{\sqrt{20}} \cdot 2,093$$

Окончательно получаем:

$$93,66 < \mu < 98,34.$$

Тема 12. Статистическая проверка гипотез

Рассмотрим теперь второе направление математической статистики – проверку статистических гипотез. И одно исследование не обходится без сравнений.

Статистическая гипотеза – это любое предположение о виде неизвестного распределения или о параметрах известных распределений.

Гипотезы принято обозначать буквой H с индексами. Будем предполагать, что у нас имеется 2 непересекающиеся гипотезы H_0 и H_1 , где H_0 – **нулевая гипотеза** (или **основная**), H_1 – **альтернативная** или **конкурирующая гипотеза**.

Задача проверки статистических гипотез состоит в том, чтобы на основе выборок $x_1, x_2, x_3, \dots, x_n$ принять (т.е. считать справедливой) либо нулевую гипотезу H_0 , либо конкурирующую гипотезу H_1 .

При проверке статистических гипотез может быть принято неправильное решение, т.е. могут быть допущены ошибки двух родов. **Ошибка первого рода** состоит в том, что будет отвергнута правильная гипотеза. **Ошибка второго рода** состоит в том, что будет принята неправильная гипотеза.

Для проверки принятой гипотезы используют **статистический критерий** – это правило, позволяющее, основываясь только на выборке, принять либо отвергнуть нулевую гипотезу H_0 .

Все критерии различий условно подразделены на две группы параметрические и непараметрические критерии.

Критерии различия называют **параметрическим** если он основан на конкретном типе распределения генеральной совокупности (как правило нормальном) или использует параметры этой совокупности (среднее, дисперсии и т.д.).

Критерий различия называют **непараметрическим** если он не базируется на предположении о типе распределения генеральной совокупности и не использует параметры этой совокупности.

Общая постановка задачи проверки гипотез

Принятие статистического решения разбивается на этапы или шаги. Рассмотрим этот алгоритм:

- 1) По выборочным данным формулируют основную гипотезу H_0 и альтернативную гипотезу H_1 .
- 2) Определяют объем выборки N .
- 3) Задают нужный уровень значимости α . Уровень значимости α – это вероятность ошибки отвергнуть нулевую гипотезу H_0 , если на самом деле эта гипотеза верна. Так, при $\alpha \leq 0,05$ ошибка возможна в 5% случаев, при $\alpha \leq 0,01$ ошибка возможна в 1% случаев и т.д.
- 4) В зависимости от гипотезы H_0 определяют статистический критерий, имеющий известное распределение.

Критерий – это случайная величина K , которая служит для проверки гипотезы H_0 . Критерий зависит от двух параметров: от числа степеней свободы и от уровня значимости α . Фактическую величину критерия получают по данным наблюдения $K_{\text{набл}}$.

- 5) По таблицам определяют критическое значение, превышение которого при справедливости гипотезы маловероятно $K_{\text{крит}}(\alpha, f)$.
- 6) Сравнивают $K_{\text{набл}}$ и $K_{\text{крит}}(\alpha, f)$.

Для параметрических критериев: если $K_{\text{набл}} > K_{\text{крит}}(\alpha, f)$, то отвергают гипотезу H_0 и принимают гипотезу H_1 , если $K_{\text{набл}} < K_{\text{крит}}(\alpha, f)$, то принимают гипотезу H_0 .

Для непараметрических критериев: если $K_{\text{набл}} < K_{\text{крит}}(\alpha, f)$, то принимают гипотезу H_0 .

- 7) Формулируют общий вывод исходя из поставленной задачи: различие статистически значимо или незначимо.

Все критерии различий условно подразделены на две группы параметрические и непараметрические критерии.

Проверка гипотез относительно средних

Сравнение средних совокупностей имеет важное практическое значение. Часто встречаются случаи, когда средний результат одной серии экспериментов отличается от среднего результата другой серии экспериментов. При этом возникает вопрос, можно ли объяснить обнаруженное расхождение средних неизбежными случайными ошибками эксперимента или оно вызвано некоторыми закономерностями.

Пусть имеются две совокупности, характеризующиеся генеральными средними $\bar{x}_0$ и $\bar{y}_0$ и известными дисперсиями σ_x^2 и σ_y^2 . Необходимо проверить гипотезу H_0 о равенстве генеральных средних, т.е. $H_0: \bar{x}_0 = \bar{y}_0$. Для проверки гипотезы H_0 из этих совокупностей взяты две независимые выборки объемов n_1 и n_2 , по которым найдены средние арифметические $\bar{x}$ и $\bar{y}$ и выборочные дисперсии S_x^2 и S_y^2 .

В случае справедливости гипотезы H_0 разность $\bar{x} - \bar{y}$ имеет нормальный закон распределения с математическим ожиданием:

$$M(\bar{x} - \bar{y}) = M(\bar{x}) - M(\bar{y}) = \bar{x}_0 - \bar{y}_0 = 0$$

и дисперсией:

$$\sigma_{\bar{x}-\bar{y}}^2 = \sigma_{\bar{x}}^2 + \sigma_{\bar{y}}^2 = \frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}$$

(дисперсия разности независимых случайных величин равна сумме их дисперсий, а дисперсия средней n независимых слагаемых в n раз меньше дисперсии каждого).

Поэтому при выполнении гипотезы H_0 статистика

$$t = \frac{(\bar{x} - \bar{y}) - M(\bar{x} - \bar{y})}{\sigma_{\bar{x}-\bar{y}}} = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{\sigma_x^2}{n_1} + \frac{\sigma_y^2}{n_2}}}$$

имеет стандартное нормальное распределение.

В случае конкурирующей гипотезы $H_1: \bar{x}_0 > \bar{y}_0$ (или $\bar{x}_0 < \bar{y}_0$) выбирают одностороннюю или двустороннюю критическую область и находят критическое значение статистики: $\Phi(t_{кр}) = \Phi(t_{1-2\alpha}) = 1 - 2\alpha$ - или

$$\Phi(t_{кр}) = \Phi(t_{1-\alpha}) = 1 - \alpha.$$

Если фактически наблюдаемое значение статистики t больше критического, определенного на уровне значимости α (по абсолютной величине), т.е. $|t| > t_{кр}$, то гипотеза H_0 отвергается. Если $|t| \leq t_{кр}$, то делается вывод, что нулевая гипотеза не противоречит имеющимся наблюдениям.

Проверка гипотез для дисперсий

Пусть генеральные совокупности X_1 и X_2 распределены нормально и имеют дисперсии σ_1^2 и σ_2^2 . Необходимо проверить нулевую гипотезу о равенстве дисперсий, т.е. $H_0: \sigma_1^2 = \sigma_2^2$ относительно конкурирующей $H_1: \sigma_1^2 \neq \sigma_2^2$.

По независимым выборкам объемов n_1 и n_2 , извлеченным из этих совокупностей, найдены исправленные выборочные дисперсии $\hat{S}_1^2$ и $\hat{S}_2^2$. Задача проверки гипотезы сводится к сравнению дисперсий $\hat{S}_1^2$ и $\hat{S}_2^2$. При заданном уровне значимости α , надо проверить нулевую гипотезу о равенстве генеральных дисперсий нормальных совокупностей.

В качестве критерия проверки нулевой гипотезы о равенстве генеральных дисперсий используем случайную величину F , равную отношению большей исправленной выборочной дисперсии к меньшей, т.е.

$$F_{\text{набл}} = \frac{\hat{s}_0^2}{\hat{s}_M^2}.$$

Величина F , при условии справедливости нулевой гипотезы имеет распределение Фишера-Снедекора со степенями свободы $f_1 = n_1 - 1$ и $f_2 = n_2 - 1$, где n_1 – объем выборки, по которой вычислена большая выборочная дисперсия.

Из таблиц находим $F_{\text{крит}}(\alpha, f_1, f_2)$ далее сравнивается $F_{\text{крит}}$ и $F_{\text{набл}}$: если $F_{\text{набл}} < F_{\text{крит}}(\alpha, f_1, f_2) \Rightarrow H_0$ (принимается нулевая гипотеза), генеральные дисперсии различаются незначимо.

Проверка гипотез о законах распределения

Выше рассматривались гипотезы, относящиеся к отдельным параметрам распределения случайных величин, причем модели законов распределения этих величин предполагались известными. Однако во многих практических задачах модель закона распределения заранее не известна и возникает задача выбора модели, согласующейся с результатами наблюдений случайной величины.

Пусть высказано предположение, что неизвестная функция распределения $F_{\text{эмп}}(x)$ исследуемой случайной величины X имеет вполне определенную модель $F_{\text{теор}}(x)$, т.е. высказана гипотеза: $H_0: F_{\text{эмп}}(x) = F_{\text{теор}}(x)$. В качестве теоретической модели может быть рассмотрена нормальная, показательная, равномерная или какая-либо другая модель. Это определяется сущностью изучаемого явления, а также результатами предварительной обработки полученных экспериментальных данных (видом гистограммы, полигона частот, соотношением основных числовых характеристик).

Выдвигается альтернативная гипотеза, что данная генеральная совокупность не распределена по закону $F_{\text{теор}}(x)$, т.е. $H_1: F_{\text{эмп}}(x) \neq F_{\text{теор}}(x)$.

Задается уровень значимости, например, $(\alpha \leq 0,05)$.

Если хотим проверить, согласуются эмпирические данные с нашим гипотетическим предложением относительно теоретической функции распределения или нет, то используем критерий согласия.

Критерием согласия называют критерий, который позволяет установить, является ли расхождение эмпирического и теоретического распределений случайным или значимым, т. е. согласуются ли данные наблюдений с выдвинутой статистической гипотезой или не согласуются. Распределение генеральной совокупности, которое она имеет в силу выдвинутой гипотезы, называют *теоретическим*. Обычно эмпирические и теоретические частоты различаются в силу того, что:

1. расхождение случайно и связано с ограниченным количеством наблюдений;
2. расхождение неслучайно и объясняется тем, что статистическая гипотеза о теоретическом законе распределения – ошибочна.

Таким образом, критерии согласия позволяют отвергнуть или подтвердить правильность выдвинутой гипотезы о модели распределения экспериментальных данных.

Рассмотрим один из них, использующий χ^2 распределение и получивший название **критерий согласия Пирсона**.

В наиболее часто используемом на практике критерии χ^2 – Пирсона в качестве меры расхождения U берется величина χ^2 , равная сумме квадратов отклонений частот (статистических вероятностей) w_i от гипотетических p_i , рассчитанных по предлагаемому распределению, взятых с некоторыми весами c_i .

$$U = \chi^2 = \sum_{i=1}^m c_i \cdot (w_i - p_i)^2$$

Веса c_i вводятся таким образом, чтобы при одних и тех же отклонениях $(w_i - p_i)^2$ больший вес имели отклонения, при которых p_i мала, и меньший вес – при которых p_i ве-

лика. Очевидно, этого удастся достичь, если взять c_i обратно пропорциональными вероятностям p_i . Взяв в качестве весов $c_i = \frac{n}{p_i}$, можно доказать, что при $n \rightarrow \infty$ статистика

$$U = \chi^2 = \sum_{i=1}^m \frac{n}{p_i} \cdot (w_i - p_i)^2 \quad \text{или} \quad U = \chi^2 = \sum_{i=1}^m \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$$

имеет χ^2 -распределение с $k = m - r - 1$ степенями свободы, где m – число интервалов эмпирического распределения (вариационного ряда); r – число параметров теоретического распределения, вычисленных по экспериментальным данным.

Числа $n_i = n \cdot w_i$ и $n \cdot p_i$ называются соответственно **эмпирическими** и **теоретическими частотами**.

Схема применения критерия χ^2 сводится к следующему:

1. Определяется мера расхождения эмпирических и теоретических частот χ^2 .
2. Для выбранного уровня значимости α по таблице χ^2 -распределения находят критическое значение $\chi_{\alpha, k}^2$ при числе степеней свободы $k = m - r - 1$.
3. Если фактически наблюдаемое значение χ^2 больше критического, т.е. $\chi^2 > \chi_{\alpha, k}^2$, то гипотеза H_0 отвергается, если $\chi^2 \leq \chi_{\alpha, k}^2$, гипотеза H_0 не противоречит опытным данным.

Замечание: статистика

$$\chi^2 = \sum_{i=1}^m \frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$$

имеет χ^2 -распределение лишь при $n \rightarrow \infty$, поэтому необходимо, чтобы в каждом интервале было достаточное количество наблюдений, по крайней мере 5 наблюдений. Если в каком-нибудь интервале число наблюдений $n_i < 5$, имеет смысл объединить соседние интервалы, чтобы в объединенных интервалах n_i было не меньше 5.

Критерий Колмогорова. На практике кроме критерия χ^2 часто используется критерий Колмогорова, в котором в качестве меры расхождения между теоретическими и эмпирическими распределениями рассматривают максимальное значение абсолютной величины разности между эмпирической функцией распределения $F_n(x)$ и соответствующей теоретической функцией распределения: $D = \max |F_n(x) - F(x)|$, называемое **статистикой критерия Колмогорова**.

Доказано, что какова бы ни была функция распределения $F(x)$ непрерывной случайной величины X , при неограниченном увеличении числа наблюдений ($n \rightarrow \infty$) вероятность неравенства $P(D\sqrt{n} \geq \lambda)$ стремимся к пределу:

$$P(\lambda) = 1 - \sum_{k=-\infty}^{+\infty} (-1)^k \cdot e^{2k^2 \cdot \lambda^2}$$

Задавая уровень значимости α , из соотношения $P(\lambda_\alpha) = \alpha$ можно найти соответствующее критическое значение λ_α . В таблице приводятся критические значения λ_α кри-

Уровень значимости α	0,40	0,30	0,20	0,10	0,05	0,025	0,01	0,005	0,001	0,0005
Критическое значение λ_α	0,89	0,97	1,07	1,22	1,36	1,48	1,63	1,73	1,95	2,03

терия Колмогорова для некоторых α .

Схема применения критерия Колмогорова

1. Строятся эмпирическая функция распределения $F_n(x)$ и предполагаемая теоретическая функция распределения $F(x)$.

2. Определяется мера расхождения между теоретическим и эмпирическим распределением D и вычисляется величина $\lambda = D\sqrt{n}$.

3. Если вычисленное значение λ окажется больше критического λ_α , определенно-го на уровне значимости α , то нулевая гипотеза H_0 о том, что случайная величина X имеет заданный закон распределения, отвергается. Если $\lambda \leq \lambda_\alpha$, то считают, что гипотеза H_0 не противоречит опытным данным.

Непараметрические критерии

Критерий различия называют *непараметрическим*, если он не базируется на предположении о типе распределения генеральной совокупности и не использует параметры этой совокупности. Поэтому для непараметрических критериев предлагается также использовать такой термин как «критерий, свободный от распределения». Это функции, зависящие непосредственно от вариантов данной совокупности с их частотами. Служат для проверки рабочих гипотез независимо от формы распределения совокупностей, из которых взяты сравниваемые выборки. Непараметрические критерии рассматривают не только количественные, но и качественные признаки, многие из которых выражаются порядковыми номерами, рангами, индексами и пр.

Например, *критерий знаков, критерий Уилкоксона-Манна-Уитни*.

Рассмотрим подробнее *критерий знаков*.

Требования критерия:

- 1) законы распределения X и Y предполагаются неизвестными, но одинаковыми;
- 2) объем выборок составляет 100 элементов и больше и одинаков.

Критерий знаков применяют для проверки нулевой гипотезы H_0 : совокупности, из которых взяты сравниваемые выборки, имеют одну и ту же или одинаковые функции распределения.

Вводится величина $z_i = y_i - x_i$. Значения этой разности меняются от опыта к опыту по величине и знаку. Если влияние фактора незначимо, то вероятно появление как положительных n_+ , так и отрицательных n_- . Из этих величин выбирают наименьшее $n_N^{\text{набл}}$. Величина $n_N^{\text{набл}}(\min)$ – это наименьшее число из однозначных разностей.

Значение $n_N^{\text{крит}}$ находят по формуле или из таблиц. Если его определять по формуле, то это целая часть A ,

$$A = \frac{(N-1)}{2} - K \cdot \sqrt{N+1},$$

где N – объем выборки. $K = 0,98$ при $\alpha \leq 0,05$.

Сравниваем: если $n_N^{\text{набл}} > n_N^{\text{крит}} \Rightarrow H_0$, то принимается нулевая гипотеза. Различие недостоверно. Действие фактора незначимо.

Критерий U Уилкоксона-Манна-Уитни

Несвязанные или независимые выборки образуются, когда в целях эксперимента для сравнения привлекаются данные двух или более выборок, причем эти выборки могут быть взяты из одной или из различных генеральных совокупностей. Таким образом, для несвязанных выборок характерно, что в них обязательно входят разные испытуемые.

Для оценки достоверности различий между несвязными выборками используется ряд непараметрических критериев. Одним из наиболее распространенных является критерий U . Этот критерий применяют для оценки различий по уровню выраженности какого-либо признака для двух независимых (несвязанных) выборок. При этом выборки могут различаться по числу входящих в них испытуемых. Этот критерий особенно удобен в том случае, когда число испытуемых невелико и в обеих выборках не превышает величину 20, хотя таблицы критических значений для величин выборок не превышающих 60 человек испытуемых.

Алгоритм применения критерия Манна – Уитни:

1. Задается уровень значимости α .

2. Эмпирические данные ранжируются (записываются в порядке возрастания) по двум линиям (одинаковые значения записываются одно под другим), которые соответствуют исследуемым группам. Ранжирование производится одновременно для обеих групп.
3. Полученные ранги суммируются по каждой линии отдельно, и наблюдаемое значение критерия находится по формуле:

$$U_i = R_i - \frac{n_i \cdot (n_i + 1)}{2}; U_i(\min) = U_{\text{набл}}$$

4. По таблице для заданной надежности и соответствующих объемов выборок находится критическое значение критерия Манна-Уитни.
5. Если расчетное значение критерия $U_{\text{набл}}$ меньше или равно критическому значению $U_{\text{крит}}$, то различие между выборками есть статически значимым. В противном случае различие статистически недостоверно и выборочные значения можно считать одинаковыми.

Типовые задания.

Пример 1. В таблице представлены изменения выработки на одного основного рабочего. Проверить гипотезу, что выработка рабочих цеха распределена по нормальному закону при $\alpha = 0,05$.

Выработка, %	94-100	100-106	106-112	112-118	118-124	124-130	130-136	136-142
Число ра- бочих	3	7	11	20	28	19	10	2

Решение. Предположим нормальный закон распределения признака. Параметры нормального закона α и σ^2 , являющиеся соответственно математическим ожиданием и дисперсией случайной величины X , неизвестны, поэтому заменяем их «наилучшими» оценками по выборке – несмещенными и состоятельными оценками соответственно выборочной средней $\bar{x}$ и «исправленной» выборочной дисперсией $\hat{S}^2$. Так как число наблюдений $n = 100$ достаточно велико, то вместо «исправленной» $\hat{S}^2$ можно взять «обычную» выборочную дисперсию $\hat{S}^2$.

В данном случае выборочная средняя равна: $\bar{x} = 119,2(\%)$, выборочная дисперсия $\hat{S}^2 = 87,48$ и выборочное среднее квадратическое отклонение $S = 9,35(\%)$.

Для расчета вероятностей p_i попадания случайной величины X в интервал $[x_i, x_{i+1}]$ используем функцию Лапласа (см. приложение, таблица II) в соответствии со свойством нормального распределения:

$$p_i(x_i \leq X \leq x_{i+1}) = \frac{1}{2} \left[\Phi \left(\frac{x_{i+1} - \alpha}{\sigma} \right) - \Phi \left(\frac{x_i - \alpha}{\sigma} \right) \right] \approx \frac{1}{2} \left[\Phi \left(\frac{x_{i+1} - 119,2}{9,35} \right) - \Phi \left(\frac{x_i - 119,2}{9,35} \right) \right].$$

Например,

$$p_i(94 \leq X \leq 100) = \frac{1}{2} \left[\Phi \left(\frac{100 - 119,2}{9,35} \right) - \Phi \left(\frac{94 - 119,2}{9,35} \right) \right] \approx \frac{1}{2} [\Phi(-2,05) - \Phi(-2,69)] = \frac{1}{2} (-0,9596 + 0,9928) = 0,0166;$$

и соответствующая первому интервалу теоретическая частота $n \cdot p_1 = 100 \cdot 0,0166 \approx 1,7$.

$$p_i(100 \leq X \leq 106) = \frac{1}{2} \left[\Phi \left(\frac{106 - 119,2}{9,35} \right) - \Phi \left(\frac{100 - 119,2}{9,35} \right) \right] \approx \frac{1}{2} [\Phi(-1,42) - \Phi(-2,05)] = \frac{1}{2} (-0,8415 + 0,9596) = 0,059;$$

и соответствующая второму интервалу теоретическая частота $n \cdot p_2 = 100 \cdot 0,059 \approx 5,9$ и т.д.

Для определения статистики χ^2 удобно составить таблицу:

i	Интервал $[x_i, x_{i+1}]$	Эмпириче- ские частоты n_i	Вероят- ности p_i	Теоретиче- ские частоты $n \cdot p_i$	$(n_i - n \cdot p_i)^2$	$\frac{(n_i - n \cdot p_i)^2}{n \cdot p_i}$
1	94-100	3	0,017	1,7	5,76	0,758
2	100-106	7	0,059	5,9		
3	106-112	11	0,141	14,1	9,61	0,682
4	112-118	20	0,228	22,8	7,84	0,344
5	118-124	28	0,247	24,7	10,89	0,441
6	124-130	19	0,182	18,2	0,64	0,035
7	130-136	10	0,087	8,7	0,16	0,014
8	136-142	2	0,029	2,9		
	Σ	100	0,990	99,0	-	$\chi^2 = 2,27$

Учитывая, что рассматриваемом эмпирическом распределении частоты первого и последнего интервалов ($n_1 = 3, n_8 = 2$) меньше 5, при использовании критерия χ^2 – Пирсона целесообразно объединить указанные интервалы с соседними. Итак, фактически наблюдаемое значение статистики $\chi^2 = 2,27$.

Так как новое число интервалов (с учетом объединение крайних) $m = 6$, а нормальный закон распределения определяется $r = 2$ параметрами, то число степеней свободы $k = m - r - 1 = 6 - 2 - 1 = 3$. Соответствующее критическое значение статистики χ^2 по таблице V приложений $\chi_{0,05;3}^2 = 7,82$. Так как $\chi^2 < \chi_{0,05;3}^2$, то гипотеза о выбранном теоретическом нормальном законе с параметрами $N(119,2; 87,48)$ согласуется с опытными данными.

Пример 2. Два пловца X и Y в равных условиях 10 раз проплыли 25-метровую дистанцию. Проанализировав время прохождения дистанции первым $x_i(\text{м} \cdot \text{с})^{-1}$ и вторым $y_i(\text{м} \cdot \text{с})^{-1}$ пловцами, определить достоверность различия полученных данных. При $\alpha = 0,05$.

x_i	12,4	12,5	12,3	12,8	12,5	12,0	12,2	12,4	12,3	12,7
y_i	12,8	12,9	12,5	12,4	12,7	12,5	12,8	12,3	12,5	12,2

Решение. Ранжируются данные и присваиваются им ранги:

R_{x_i}	1	2,5	5	5	8	8	12	12	-	15,5	18	-	-
x_i	12,0	12,2	12,3	12,3	12,4	12,4	12,5	12,5	-	12,7	12,8	-	-
R_{y_i}	-	2,5	5	-	8	-	12	12	12	15,5	18	18	20
y_i	-	12,2	12,3	-	12,4	-	12,5	12,5	12,5	12,7	12,8	12,8	12,9

Суммируются ранги отдельно R_{x_i} и R_{y_i} .

$$R_{x_i} = 1 + 2,5 + 5 + 5 + 8 + 8 + 12 + 12 + 15,5 + 18 = 87,0,$$

$$R_{y_i} = 2,5 + 5 + 8 + 12 + 12 + 15,5 + 18 + 18 + 20 = 123,0.$$

Наблюдаемое значение критерия находится по формуле:

$$U_i = R_i - \frac{n_i \cdot (n_i + 1)}{2}.$$

$$U_1 = R_{x_i} - \frac{n_1 \cdot (n_1 + 1)}{2} = 87,0 - \frac{10 \cdot 11}{2} = 87,0 - 55 = 32,0$$

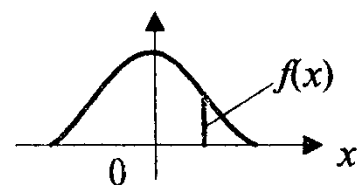
$$U_2 = R_{y_i} - \frac{n_2 \cdot (n_2 + 1)}{2} = 123,0 - \frac{10 \cdot 11}{2} = 123,0 - 55 = 68,0$$

$$U_{\text{набл}} = 32,0.$$

По таблице критических значений критерия Манна-Уитни (см. Приложение таб VIII) при заданном уровне значимости 0,05 и количестве эмпирических измерений 10 находим критическое значения критерия $U_{\text{крит}} = 27$.

Так как $U_{\text{набл}} > U_{\text{крит}}$, различие между выборками статистически недостоверно. Показатели обоих пловцов на 25 – метровой дистанции существенно не отличаются, пловцы – спортсмены одной квалификации

Значения функции $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

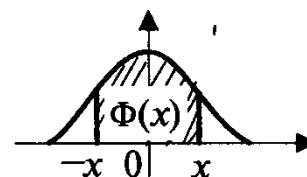


Целые и де- сятые доли x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,3989	0,3989	0,3989	0,3988	0,3986	0,3984	0,3982	0,3980	0,3977	0,3973
0,1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0,2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0,3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0,4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0,5	3521	3503	3485	3467	3448	3429	3410	3391	3372	3352
0,6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0,7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0,8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0,9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1,0	0,2420	0,2396	0,2371	0,2347	0,2323	0,2299	0,2275	0,2251	0,2227	0,2203
1,1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1,2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1,3	1714	1691	1669	1647	1626	1604	1582	1561	1539	1518
1,4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1,5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1,6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1,7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1,8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1,9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2,0	0,0540	0,0529	0,0519	0,0508	0,0498	0,0488	0,0478	0,0468	0,0459	0,0449
2,1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2,2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2,3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2,4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2,5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2,6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107

x	0	1	2	3	4	5	6	7	8	9
2,7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2,8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2,9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3,0	0,0044	0,0043	0,0042	0,0041	0,0039	0,0038	0,0037	0,0036	0,0035	0,0034
3,1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3,2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3,3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3,4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3,5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3,6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3,7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3,8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3,9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
4,0	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001	0,0001
4,1	0,0001338									
4,5	0,0000160									
5,0	0,0000015									

Таблица II

Значения функции Лапласа $\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$



Целые и деся- тые доли x	Сотые доли x									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0080	0,0160	0,0239	0,0319	0,0399	0,0478	0,0558	0,0638	0,0717
0,1	0,0797	0,0876	0,0955	0,1034	0,1113	0,1192	0,1271	0,1350	0,1428	0,1507
0,2	0,1585	0,1663	0,1741	0,1819	0,1897	0,1974	0,2051	0,2128	0,2205	0,2282
0,3	0,2358	0,2434	0,2510	0,2586	0,2661	0,2737	0,2812	0,2886	0,2960	0,3035
0,4	0,3108	0,3182	0,3255	0,3328	0,3401	0,3473	0,3545	0,3616	0,3688	0,3759
0,5	0,3829	0,3899	0,3969	0,4039	0,4108	0,4177	0,4245	0,4313	0,4381	0,4448
0,6	0,4515	0,4581	0,4647	0,4713	0,4778	0,4843	0,4907	0,4971	0,5035	0,5098
0,7	0,5161	0,5223	0,5285	0,5346	0,5407	0,5467	0,5527	0,5587	0,5646	0,5705
0,8	0,5763	0,5821	0,5878	0,5935	0,5991	0,6047	0,6102	0,6157	0,6211	0,6265
0,9	0,6319	0,6372	0,6424	0,6476	0,6528	0,6579	0,6629	0,6679	0,6729	0,6778
1,0	0,6827	0,6875	0,6923	0,6970	0,7017	0,7063	0,7109	0,7154	0,7199	0,7243
1,1	0,7287	0,7330	0,7373	0,7415	0,7457	0,7499	0,7540	0,7580	0,7620	0,7660
1,2	0,7699	0,7737	0,7775	0,7813	0,7850	0,7887	0,7923	0,7959	0,7984	0,8029
1,3	0,8064	0,8098	0,8132	0,8165	0,8198	0,8230	0,8262	0,8293	0,8324	0,8355
1,4	0,8385	0,8415	0,8444	0,8473	0,8501	0,8529	0,8557	0,8584	0,8611	0,8638
1,5	0,8664	0,8690	0,8715	0,8740	0,8764	0,8789	0,8812	0,8836	0,8859	0,8882
1,6	0,8904	0,8926	0,8948	0,8969	0,8990	0,9011	0,9031	0,9051	0,9070	0,9090
1,7	0,9109	0,9127	0,9146	0,9164	0,9181	0,9199	0,9216	0,9233	0,9249	0,9265
1,8	0,9281	0,9297	0,9312	0,9327	0,9342	0,9357	0,9371	0,9385	0,9392	0,9412
1,9	0,9426	0,9439	0,9451	0,9464	0,9476	0,9488	0,9500	0,9512	0,9523	0,9533
2,0	0,9545	0,9556	0,9566	0,9576	0,9586	0,9596	0,9606	0,9616	0,9625	0,9634
2,1	0,9643	0,9651	0,9660	0,9668	0,9676	0,9684	0,9692	0,9700	0,9707	0,9715
2,2	0,9722	0,9729	0,9736	0,9743	0,9749	0,9756	0,9762	0,9768	0,9774	0,9780
2,3	0,9786	0,9791	0,9797	0,9802	0,9807	0,9812	0,9817	0,9822	0,9827	0,9832
2,4	0,9836	0,9841	0,9845	0,9849	0,9853	0,9857	0,9861	0,9865	0,9869	0,9872
2,5	0,9876	0,9879	0,9883	0,9886	0,9889	0,9892	0,9895	0,9898	0,9901	0,9904
2,6	0,9907	0,9910	0,9912	0,9915	0,9917	0,9920	0,9922	0,9924	0,9926	0,9928
2,7	0,9931	0,9933	0,9935	0,9937	0,9939	0,9940	0,9942	0,9944	0,9946	0,9947
2,8	0,9949	0,9951	0,9952	0,9953	0,9955	0,9956	0,9958	0,9959	0,9960	0,9961
2,9	0,9963	0,9964	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972
3,0	0,9973	0,9974	0,9975	0,9976	0,9976	0,9977	0,9978	0,9979	0,9979	0,9980
3,1	0,9981	0,9981	0,9982	0,9983	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986
3,2	0,9986	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,3	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,4	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995	0,9995
3,5	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997	0,9997
3,6	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998	0,9998	0,9998
3,7	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
4,0	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999

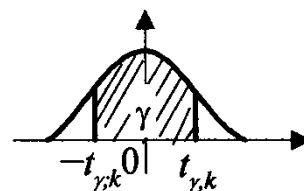
Таблица III

Значения функции Пуассона $P_m(\lambda) = \frac{\lambda^m}{m!} e^{-\lambda}$

$m \backslash \lambda$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
0	0,9048	0,8187	0,7408	0,6703	0,6065	0,5488	0,4966	0,4493	0,4066	0,3679
1	0,0905	0,1637	0,2223	0,2681	0,3033	0,3293	0,3476	0,3595	0,3659	0,3679
2	0,0045	0,0164	0,0333	0,0536	0,0758	0,0988	0,1216	0,1438	0,1647	0,1839
3	0,0002	0,0011	0,0033	0,0072	0,0126	0,0198	0,0284	0,0383	0,0494	0,0613
4	0,0000	0,0001	0,0003	0,0007	0,0016	0,0030	0,0050	0,0077	0,0111	0,0153
5	0,0000	0,0000	0,0000	0,0001	0,0002	0,0003	0,0007	0,0012	0,0020	0,0031
6	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0002	0,0003	0,0005
7	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001

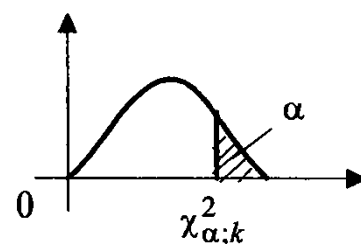
$m \backslash \lambda$	2,0	3,0	4,0	5,0	6,0	7,0	8,0	9,0	10,0
0	0,1353	0,0498	0,0183	0,0067	0,0025	0,0009	0,0003	0,0001	0,0001
1	0,2707	0,1494	0,0733	0,0337	0,0149	0,0064	0,0027	0,0011	0,0005
2	0,2707	0,2240	0,1465	0,0842	0,0446	0,0223	0,0107	0,0050	0,0023
3	0,1805	0,2240	0,1954	0,1404	0,0892	0,0521	0,0286	0,0150	0,0076
4	0,0902	0,1681	0,1954	0,1755	0,1339	0,0912	0,0572	0,0337	0,0189
5	0,0361	0,1008	0,1563	0,1755	0,1606	0,1277	0,0916	0,0607	0,0378
6	0,0120	0,0504	0,1042	0,1462	0,1606	0,1490	0,1221	0,0911	0,0631
7	0,0034	0,0216	0,0595	0,1045	0,1377	0,1490	0,1396	0,1171	0,0901
8	0,0009	0,0081	0,0298	0,0653	0,1033	0,1304	0,1396	0,1318	0,1126
9	0,0002	0,0027	0,0132	0,0363	0,0689	0,1014	0,1241	0,1318	0,1251
10	0,0000	0,0008	0,0053	0,0181	0,0413	0,0710	0,0993	0,1186	0,1251
11	0,0000	0,0002	0,0019	0,0082	0,0225	0,0452	0,0722	0,0970	0,1137
12	0,0000	0,0001	0,0006	0,0034	0,0113	0,0264	0,0481	0,0728	0,0948
13	0,0000	0,0000	0,0002	0,0013	0,0052	0,0142	0,0296	0,0504	0,0729
14	0,0000	0,0000	0,0001	0,0005	0,0022	0,0071	0,0169	0,0324	0,0521
15	0,0000	0,0000	0,0000	0,0002	0,0009	0,0033	0,0090	0,0194	0,0347
16	0,0000	0,0000	0,0000	0,0000	0,0003	0,0015	0,0045	0,0109	0,0217
17	0,0000	0,0000	0,0000	0,0000	0,0001	0,0006	0,0021	0,0058	0,0128
18	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0009	0,0029	0,0071
19	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004	0,0014	0,0037
20	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002	0,0006	0,0019
21	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0003	0,0009
22	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001	0,0004
23	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0002
24	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0001
25	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000

Таблица IV

Значения $t_{\gamma,k}$ -критерия Стьюдента

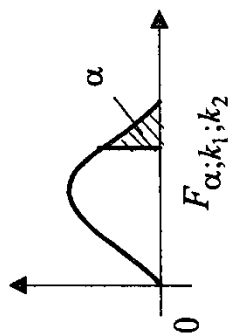
Число степеней свободы k	Вероятность γ											
	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95	0,98	0,99
1	0,16	0,32	0,51	0,73	1,00	1,38	1,96	3,08	6,31	12,71	31,82	63,66
2	14	29	44	62	0,82	06	34	1,89	2,92	4,30	6,96	9,92
3	14	28	42	58	76	0,98	25	64	35	3,18	4,54	5,84
4	13	27	41	57	74	94	19	53	13	2,78	3,75	4,60
5	13	27	41	56	73	92	16	48	01	57	36	03
6	0,13	0,26	0,40	0,55	1,72	1,91	1,13	1,44	1,94	2,45	3,14	3,71
7	13	26	40	55	71	90	12	41	89	36	00	50
8	13	26	40	55	70	89	11	40	86	31	2,90	35
9	13	26	40	54	70	88	10	38	83	26	82	25
10	13	26	40	54	70	88	09	37	81	23	76	17
11	0,13	0,26	0,40	0,54	0,70	0,88	1,09	1,36	1,80	2,20	2,72	3,11
12	13	26	39	54	69	87	08	36	78	18	68	05
13	13	26	39	54	69	87	08	35	77	16	65	01
14	13	26	39	54	69	87	08	34	76	14	62	2,98
15	13	26	39	54	69	87	07	34	75	13	60	95
16	0,13	0,26	0,39	0,53	0,69	0,86	1,07	1,34	1,75	2,12	2,58	2,92
17	13	26	39	53	69	86	07	33	74	11	57	90
18	13	26	39	53	69	86	07	33	73	10	55	88
19	13	26	39	53	69	86	07	33	73	09	54	86
20	13	26	39	53	69	86	06	32	72	09	53	84
21	0,13	0,26	0,39	0,53	0,69	0,86	1,06	1,32	1,72	2,08	2,52	2,83
22	13	26	39	53	69	86	06	32	72	07	51	82
23	13	26	39	53	68	86	06	32	71	07	50	81
24	13	26	39	53	68	86	06	32	71	06	49	80
25	13	26	39	53	68	86	06	32	71	06	48	79
26	0,13	0,26	0,39	0,53	0,68	0,86	1,06	1,31	1,71	2,06	2,48	2,78
27	13	26	39	53	68	85	06	31	70	05	47	77
28	13	26	39	53	68	85	06	31	70	05	47	76
29	13	26	39	53	68	85	05	31	70	04	46	76
30	13	26	39	53	68	85	05	31	70	04	46	75
40	0,13	0,25	0,39	0,53	0,68	0,85	1,05	1,30	1,68	2,02	2,42	2,70
60	13	25	39	53	68	85	05	30	67	00	39	66
120	0,13	0,25	0,39	0,53	0,68	0,84	1,04	1,29	1,66	1,98	2,36	2,62
∞	13	25	38	52	67	84	04	28	64	96	33	58

Таблица V

Значения $\chi^2_{\alpha;k}$ критерия Пирсона

Число степен- ней свобо- ды k	Вероятность α												
	0,99	0,98	0,95	0,90	0,80	0,70	0,50	0,30	0,20	0,10	0,05	0,02	0,01
1	0,00	0,00	0,00	0,02	0,06	0,15	0,45	1,07	1,64	2,71	3,84	5,41	6,64
2	0,02	0,04	0,10	0,21	0,45	0,71	1,39	2,41	3,22	4,60	5,99	7,82	9,21
3	0,11	0,18	0,35	0,58	1,00	1,42	2,37	3,66	4,64	6,25	7,82	9,84	11,3
4	0,30	0,43	0,71	1,06	1,65	2,20	3,36	4,88	5,99	7,78	9,49	11,7	13,3
5	0,55	0,75	1,14	1,61	2,34	3,00	4,35	6,06	7,29	9,24	11,1	13,4	15,1
6	0,87	1,13	1,63	2,20	3,07	3,83	5,35	7,23	8,56	10,6	12,6	15,0	16,8
7	1,24	1,56	2,17	2,83	3,82	4,67	6,35	8,38	9,80	12,0	14,1	16,6	18,5
8	1,65	2,03	2,73	3,49	4,59	5,53	7,34	9,52	11,0	13,4	15,5	18,2	20,1
9	2,09	2,53	3,32	4,17	5,38	6,39	8,34	10,7	12,2	14,7	16,9	19,7	21,7
10	2,56	3,06	3,94	4,86	6,18	7,27	9,34	11,8	13,4	16,0	18,3	21,2	23,2
11	3,05	3,61	4,58	5,58	6,99	8,15	10,3	12,9	14,6	17,3	19,7	22,6	24,7
12	3,57	4,18	5,23	6,30	7,81	9,03	11,3	14,0	15,8	18,5	21,0	24,1	26,2
13	4,11	4,76	5,89	7,04	8,63	9,93	12,3	15,1	17,0	19,8	22,4	25,5	27,7
14	4,66	5,37	6,57	7,79	9,47	10,8	13,3	16,2	18,1	21,1	23,7	26,9	29,1
15	5,23	5,98	7,26	8,55	10,3	11,7	14,3	17,3	19,3	22,3	25,0	28,3	30,6
16	5,81	6,61	7,96	9,31	11,1	12,6	15,3	18,4	20,5	23,5	26,3	29,6	32,0
17	6,41	7,26	8,67	10,1	12,0	13,5	16,3	19,5	21,6	24,8	27,6	31,0	33,4
18	7,02	7,91	9,39	10,9	12,9	14,4	17,3	20,6	22,8	26,0	28,9	32,3	34,8
19	7,63	8,57	10,1	11,6	13,7	15,3	18,3	21,7	23,9	27,2	30,1	33,7	36,2
20	8,26	9,24	10,8	12,4	14,6	16,3	19,3	22,8	25,0	28,4	31,4	35,0	37,6
21	8,90	9,92	11,6	13,2	15,4	17,2	20,3	23,9	26,2	29,6	32,7	36,3	38,9
22	9,54	10,6	12,3	14,0	16,3	18,1	21,3	24,9	27,3	30,8	33,9	37,7	40,3
23	10,2	11,3	13,1	14,8	17,2	19,0	22,3	26,0	28,4	32,0	35,2	39,0	41,6
24	10,9	12,0	13,8	15,7	18,1	19,9	23,3	27,1	29,6	33,2	36,4	40,3	43,0
25	11,5	12,7	14,6	16,5	18,9	20,9	24,3	28,2	30,7	34,4	37,7	41,7	44,3
26	12,2	13,4	15,4	17,3	19,8	21,8	25,3	29,2	31,8	35,6	38,9	42,9	45,6
27	12,9	14,1	16,1	18,1	20,7	22,7	26,3	30,3	32,9	36,7	40,1	44,1	47,0
28	13,6	14,8	16,9	18,9	21,6	23,6	27,3	31,4	34,0	37,9	41,3	45,4	48,3
29	14,3	15,6	17,7	19,8	22,5	24,6	28,3	32,5	35,1	39,1	42,6	46,7	49,6
30	14,9	16,3	18,5	20,6	23,4	25,5	29,3	33,5	36,2	40,3	43,8	48,0	50,9

Таблица VI

Значения $F_{\alpha; k_1; k_2}$ - критерия Фишера—Снедекора

		$\alpha=0,05$																	
$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	240	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84

Продолжение табл. VI

$k_1 \backslash k_2$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,15	2,07	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,13	2,06	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,12	2,04	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,10	2,03	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,83	1,75	0,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,83	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00

$\alpha=0,01$																			
1	4052	4999,5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40	99,42	99,43	99,45	99,46	99,47	99,47	99,48	99,49	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23	27,05	26,87	26,69	26,60	26,50	26,41	26,32	26,22	26,13
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55	14,37	14,20	14,02	13,93	13,84	13,75	13,65	13,56	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05	9,89	9,72	9,55	9,47	9,38	9,29	9,20	9,11	9,02
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87	7,72	7,56	7,40	7,31	7,23	7,14	7,06	6,97	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,99	6,84	6,72	6,62	6,47	6,31	6,16	6,07	5,99	5,91	5,82	5,74	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81	5,67	5,52	5,36	5,28	5,20	5,12	5,03	4,95	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26	5,11	4,96	4,81	4,73	4,65	4,57	4,48	4,40	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85	4,71	4,56	4,41	4,33	4,25	4,17	4,08	4,00	3,91
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54	4,40	4,25	4,10	4,02	3,94	3,86	3,78	3,69	3,60
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30	4,16	4,01	3,86	3,78	3,70	3,62	3,54	3,45	3,36
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10	3,96	3,82	3,66	3,59	3,51	3,43	3,34	3,25	3,17
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94	3,80	3,66	3,51	3,43	3,35	3,27	3,18	3,09	3,00

15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80	3,67	3,52	3,37	3,29	3,21	3,13	3,05	2,96	2,87
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69	3,55	3,41	3,26	3,18	3,10	3,02	2,93	2,84	2,75
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59	3,46	3,31	3,16	3,08	3,00	2,92	2,83	2,75	2,65
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51	3,37	3,23	3,08	3,00	2,92	2,84	2,75	2,66	2,57
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43	3,30	3,15	3,00	2,92	2,84	2,76	2,67	2,58	2,49
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37	3,23	3,09	2,94	2,86	2,78	2,69	2,61	2,52	2,42
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31	3,17	3,03	2,88	2,80	2,72	2,64	2,55	2,46	2,36
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26	3,12	2,98	2,83	2,75	2,67	2,58	2,50	2,40	2,31
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21	3,07	2,93	2,78	2,70	2,62	2,54	2,45	2,35	2,26
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17	3,03	2,89	2,74	2,66	2,58	2,49	2,40	2,31	2,21
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13	2,99	2,85	2,70	2,62	2,54	2,45	2,36	2,27	2,17
26	7,72	5,53	4,64	4,14	3,82	3,59	3,42	3,29	3,18	3,09	2,96	2,81	2,66	2,58	2,50	2,42	2,33	2,23	2,13
27	7,68	5,49	4,60	4,11	3,78	3,56	3,39	3,26	3,15	3,06	2,93	2,78	2,63	2,55	2,47	2,38	2,29	2,20	2,10
28	7,64	5,45	4,57	4,07	3,75	3,53	3,36	3,23	3,12	3,03	2,90	2,75	2,60	2,52	2,44	2,35	2,26	2,17	2,06
29	7,60	5,42	4,54	4,04	3,73	3,50	3,33	3,20	3,09	3,00	2,87	2,73	2,57	2,49	2,41	2,33	2,23	2,14	2,03
30	7,56	5,39	4,51	4,02	3,70	3,47	3,30	3,17	3,07	2,98	2,84	2,70	2,55	2,47	2,39	2,30	2,21	2,11	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	3,12	2,99	2,89	2,80	2,66	2,52	2,37	2,29	2,20	2,11	2,02	1,92	1,80
60	7,08	4,98	4,13	3,65	3,34	3,12	2,95	2,82	2,72	2,63	2,50	2,35	2,20	2,12	2,03	1,94	1,84	1,73	1,60
120	6,85	4,79	3,95	3,48	3,17	2,96	2,79	2,66	2,56	2,47	2,34	2,19	2,03	1,95	1,86	1,76	1,66	1,53	1,38
∞	6,63	4,61	3,78	3,32	3,02	2,80	2,64	2,51	2,41	2,32	2,18	2,04	1,88	1,79	1,70	1,59	1,47	1,32	1,00

k_1 — число степеней свободы для большой дисперсии, k_2 — для меньшей дисперсии.

Таблица G-критерия знаков

n	α			n	α			n	α			n	α		
	0,10	0,05	0,01		0,10	0,05	0,01		0,10	0,05	0,01		0,10	0,05	0,01
5	1	0	0	27	9	8	7	49	19	18	16	92	38	37	34
6	1	1	0	28	10	9	7	50	19	18	16	94	39	38	35
7	1	1	0	29	10	9	8	52	20	19	17	96	40	38	35
8	2	1	1	30	11	10	8	54	21	20	18	98	41	39	36
9	2	2	1	31	11	10	8	56	22	21	18	100	42	40	37
10	2	2	1	32	11	10	9	58	23	22	19	110	46	45	42
11	3	2	1	33	12	11	9	60	24	22	20	120	51	49	46
12	3	3	2	34	12	11	10	62	25	23	21	130	56	54	50
13	4	3	2	35	13	12	10	64	25	24	22	140	60	58	55
14	4	3	2	36	13	12	10	66	26	25	23	150	65	63	59
15	4	4	3	37	14	13	11	68	27	26	23	160	70	68	64
16	5	4	3	38	14	13	11	70	28	27	24	170	74	72	68
17	5	5	3	39	14	13	12	72	29	28	25	180	79	77	73
18	6	5	4	40	15	14	12	74	30	29	26	190	84	82	77
19	6	5	4	41	15	14	12	76	31	29	27	200	88	86	82
20	6	6	4	42	16	15	13	78	32	30	28	220	98	95	91
21	7	6	5	43	16	15	13	80	33	31	29	240	107	105	100
22	7	6	5	44	17	16	14	82	34	32	29	260	117	114	109
23	8	7	5	45	17	16	14	84	34	33	30	280	126	124	118
24	8	7	6	46	17	16	14	86	35	34	31	300	136	133	128
25	8	8	6	47	18	17	15	88	36	35	32				
26	9	8	7	48	18	17	15	90	37	36	33				

Таблица VIII

Критические значения критерия U Манна-Уитни для уровней статистической значимости $\rho \leq 0,05$ и $\rho \leq 0,01$ (по Гублеру Е.В., Генкину А.А., 1973)

Различия между двумя выборками можно считать значимыми ($\rho < 0,05$), если $U_{\text{эмп}}$ ниже или равен $U_{0,05}$, и тем более достоверными ($\rho < 0,01$), если $U_{\text{эмп}}$ ниже или равен $U_{0,01}$.

n_1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
n_2	$\rho=0,05$																			
3	-	0																		
4	-	0	1																	
5	0	1	2	4																
6	0	2	3	5	7															
7	0	2	4	6	8	11														
8	1	3	5	8	10	13	15													
9	1	4	6	9	12	15	18	21												
10	1	4	7	11	14	17	20	24	27											
11	1	5	8	12	16	19	23	27	31	34										
12	2	5	9	13	17	21	26	30	34	38	42									
13	2	6	10	15	19	24	28	33	37	42	47	51								
14	3	7	11	16	21	26	31	36	41	46	51	56	61							
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72						
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83					
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96				
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109			
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123		
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138	
n_2	$\rho=0,01$																			
5	-	-	0	1																
6	-	-	1	2	3															
7	-	0	1	3	4	6														
8	-	0	2	4	6	7	9													
9	-	1	3	5	7	9	11	14												
10	-	1	3	6	8	11	13	16	19											
11	-	1	4	7	9	12	15	18	22	25										
12	-	2	5	8	11	14	17	21	24	28	31									
13	0	2	5	9	12	16	20	23	27	31	35	39								
14	0	2	6	10	13	17	22	26	30	34	38	43	47							
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56						
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66					
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77				
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88			
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101		
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114	

Учебно-методическое и информационное обеспечение дисциплины

1. Перечень основной и дополнительной литературы, необходимой для освоения дисциплины

1.1. Перечень основной литературы:

1. Гмурман В.Е. Теория вероятностей и математическая статистика. Юрайт. 2020. 480 с.
2. Гнеденко Б.В. Курс теории вероятностей. Едиториал УРСС. 2019. 456 с 8.1.2.

Перечень дополнительной литературы:

1. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. Физико-математическая библиотека инженера. 1988. 480 с.
2. Джафаров, К.А. Теория вероятностей и математическая статистика учебное пособие / К.А. Джафаров; Министерство образования и науки Российской Федерации, Новосибирский государственный технический университет. - Новосибирск: НГТУ, 2015. - 167 с.: схем. - Библиогр. в кн. - ISBN 978-5-7782-2720-0; то же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=438304>
3. Маталыцкий, М.А. Теория вероятностей и математическая статистика: учебник / М.А. Маталыцкий, Г.А. Хацкевич. - Минск: Вышэйшая школа, 2017. - 592 с. - Библиогр. в кн. - ISBN 978-985-06-2855-8

Перечень учебно-методического обеспечения самостоятельной работы обучающихся по дисциплине

1. Методические рекомендации к практическим занятиям по дисциплине «Теория вероятностей и математическая статистика».
 2. Методические рекомендации к самостоятельной работе по дисциплине «Теория вероятностей и математическая статистика».
 3. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины
1. <http://biblioclub.ru/index.php?page=book&id=477424>.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА
Методические указания для обучающихся по организации и проведению
самостоятельной работы

Направление подготовки 03.03.02 Физика
Направленность (профиль)
«Физика Земли и космоса»

Ставрополь
2026

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

ПЛАН-ГРАФИК ВЫПОЛНЕНИЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ИЗУЧЕНИЮ ТЕОРЕТИЧЕСКОГО МАТЕРИАЛА, ПОДГОТОВКЕ К ПРАКТИЧЕСКИМ ЗАНЯТИЯМ

ВВЕДЕНИЕ

Самостоятельная работа – планируемая учебная, учебно-исследовательская, научно-исследовательская работа студентов, выполняемая во внеаудиторное (аудиторное) время по заданию и при методическом руководстве преподавателя, но без его непосредственного участия (при частичном непосредственном участии преподавателя, оставляющем ведущую роль за работой студента).

Самостоятельная работа студента является важным видом учебной и научной деятельности студента. Федеральным государственным образовательным стандартом предусматривается значительный объем времени из общей трудоемкости дисциплины на самостоятельную работу. В связи с этим, обучение включает в себя две, практически одинаковые по объему и взаимовлиянию части – процесса обучения и процесса самообучения. Поэтому самостоятельная работа должна стать эффективной и целенаправленной работой студента.

Концепцией модернизации российского образования определены основные задачи профессионального образования – подготовка квалифицированного работника соответствующего уровня и профиля, конкурентоспособного на рынке труда, компетентного, ответственного, свободно владеющего своей профессией и ориентированного в смежных областях деятельности, способного к эффективной работе по специальности на уровне мировых стандартов, готового к постоянному профессиональному росту, социальной и профессиональной мобильности.

Решение этих задач невозможно без повышения роли самостоятельной работы студентов над учебным материалом, усиления ответственности преподавателей за развитие навыков самостоятельной работы, за стимулирование профессионального роста студентов, воспитание творческой активности и инициативы.

Самостоятельная работа студента предполагает, прежде всего, его научно-исследовательскую и практическую деятельность, поскольку именно эти виды учебной работы в первую очередь готовят к самостоятельному выполнению профессиональных задач. Конечной целью самостоятельной работы является организация самостоятельной познавательной деятельности, формирование умения самостоятельно анализировать информацию, выделять главное и второстепенное.

Основными формами самостоятельной работы студентов по дисциплине «Практическая физика» являются:

- подготовка к практическим занятиям;
- самостоятельное изучение литературы;

– подготовка к экзамену.

Практические занятия проводятся с целью закрепления усвоенной информации, приобретения навыков ее применения при решении практических задач в соответствующей предметной области.

В результате изучения дисциплины у обучающихся формируются следующие компетенции (в соответствии с образовательным стандартом):

ИД-2_{ОПК-1}

Обладает навыками работы с математическим аппаратом; владеет методами математического и численного моделирования для описания, анализа, теоретического и экспериментального исследования физических систем, явлений и процессов в своей профессиональной деятельности.

ПЛАН-ГРАФИК ВЫПОЛНЕНИЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ

Вид деятельности студентов	Итоговый продукт самостоятельной работы	Средства и технологии оценки	Время сдачи
Подготовка к занятию на тему: Элементы комбинаторики	Решение типовых заданий	Типовые задания	Практическое №1
Подготовка к занятию на тему: Основные понятия теории вероятностей	Решение типовых заданий	Типовые задания	Практическое №2
Подготовка к занятию на тему: Теоремы сложения и умножения вероятностей.	Решение типовых заданий	Типовые задания	
Подготовка к занятию на тему: Формула полной вероятности.	Решение типовых заданий	Типовые задания	Практическое №3
Подготовка к занятию на тему: Теорема гипотез и формула Бернулли	Решение типовых заданий	Типовые задания	Практическое №4
Подготовка к занятию на тему: Предельные теоремы	Решение типовых заданий	Типовые задания	Практическое №5
Подготовка к занятию на тему: Законы больших чисел	Решение типовых заданий	Типовые задания	Практическое №6
Подготовка к занятию на тему: Дискретная случайная величина	Решение типовых заданий	Типовые задания	Практическое №7
Подготовка к занятию на тему: Непрерывная случайная величина	Решение типовых заданий	Типовые задания	Практическое №9
Подготовка к занятию на тему: Основные законы распределения случайных величин	Решение типовых заданий	Типовые задания	Практическое №12
Подготовка к занятию на тему: Основные понятия математической статистики	Решение типовых заданий	Типовые задания	Практическое №13
Подготовка к занятию на тему: Оценка параметров генеральной совокупности по ее выборке	Решение типовых заданий	Типовые задания	Практическое №15
Подготовка к занятию на тему: Статистическая проверка гипотез	Решение типовых заданий	Типовые задания	Практическое №18

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ИЗУЧЕНИЮ ТЕОРЕТИЧЕСКОГО МАТЕРИАЛА, ПОДГОТОВКЕ К ПРАКТИЧЕ- СКИМ ЗАНЯТИЯМ

При подготовке к практическим занятиям необходимо:

- повторить материал соответствующих лекций;
- внимательно изучить описание практического занятия, изложенное в практикуме;
- Рассмотреть приведенные примеры решений типовых заданий
- Решить типовые задания, предложенные для самостоятельного решения

В результате выполнения заданий достигаются следующие результаты обучения:

- развитие мышления и творческих способностей студента,
- приобретение и закрепление навыков самостоятельной работы,
- приобретение и закрепление навыков оформления и составления письменных заданий.

Примеры типовых заданий

1. В туристической группе 14 женщин и 9 мужчин. Среди них разыгрываются 6 билетов на бесплатное посещение театра. Какова вероятность того, что среди обладателей билетов окажутся три женщины и трое мужчин?

2. Кодовый замок содержит 5 цифр, которыми могут быть числа от 0 до 9. Замок открывается при наборе только одной единственной комбинации из пяти различных цифр. Какова вероятность открыть этот замок, набрав случайным образом пять цифр?

3. В аптеку поступили лекарственные средства от 3 фирм производителей. Первый производитель поставил 200 ампул, второй – 200, а третий – 400. Известно, что процент бракованной продукции составил: у первого – 2%, у второго – 3%, у третьего – 1%. Для контроля взята одна ампула. Найти вероятность того, что она оказалась бракованной.

4. Детали, изготавливаемые цехом завода, попадают для проверки их на стандартность к одному из двух контролеров. Вероятность того, что деталь попадет к первому – 0,65, ко второму – 0,35. Вероятность того, что годная деталь при проверке будет признана стандартной первым контролером, равна 0,94, а вторым – 0,98. Годная деталь при проверке была признана стандартной. Найти вероятность того, что деталь проверил первый контролер.

5. Случайная величина задана плотностью распределения

$$f(x) = \begin{cases} 0, & x \leq 0, \\ \frac{\sin x}{2}, & 0 < x \leq \pi, \\ 0, & x > \pi. \end{cases}$$

Найти: а) функцию распределения; б) вероятность того, что в результате испытания случайная величина примет значение, заключенное в интервале $(0, \frac{\pi}{4})$.

6. Независимые случайные величины X и Y заданы следующими законами распределения:

X_i	7	2	1
P_i	0,6	0,1	0,4

Y_i	1	4	6	7
P_i	0,1	0,4	0,3	0,5

Найти $M(X + Y)$ и $M(X \cdot Y)$

Вопросы для подготовки к экзамену

1. Предмет и основные определения теории вероятностей.
2. Комбинаторика: размещения, сочетания, перестановки. Размещения, сочетания и перестановки с повторениями. Примеры.
3. Классическое определение вероятности. Свойства вероятности, вытекающие из классического определения. Примеры.
4. Статистическое определение вероятности, его особенности и связь с классическим определением.
5. Полная группа несовместных событий, противоположные события, свойства их вероятностей.
6. Зависимые и независимые события. Условные и безусловные вероятности.
7. Теоремы умножения вероятностей.
8. Теоремы сложения вероятностей.
9. Формула полной вероятности. Формула Байеса.
10. Случайные величины и случайные события. Дискретные и непрерывные случайные величины. Закон распределения случайной величины и способы его задания.
11. Числовые характеристики случайных величин. Начальные и центральные моменты. Асимметрия и эксцесс.
12. Математическое ожидание случайной величины. Его смысл и примеры. Свойства математического ожидания.
13. Дисперсия и среднее квадратическое отклонение случайной величины. Их смысл и примеры вычисления. Формулы для вычисления дисперсии. Свойства дисперсии.
14. Математическое ожидание, дисперсия и среднее квадратическое отклонение частоты и частоты.
15. Формула Бернулли. Биномиальное распределение. Наивероятнейшее число появлений события.
16. Формула Пуассона. Закон распределения вероятностей редких событий.
17. Гипергеометрическое распределение.
18. Равномерное распределение.
19. Непрерывные случайные величины. Дифференциальная и интегральная функции их распределения, их смысл и связь между ними.
20. Вероятность попадания случайной величины в заданный интервал. Вероятность того, что непрерывная случайная величина примет точно заданное значение.
21. Нормальное распределение. Плотность нормального распределения и ее свойства. Функция распределения нормально распределенной случайной величины.
22. Нормированное (стандартное) нормальное распределение.
23. Функция Лапласа: график, свойства, таблицы.
24. Вероятность попадания нормально распределенной случайной величины в заданный интервал.
25. Вероятность заданного отклонения нормальной случайной величины от своего математического ожидания. Правило трех сигм.
26. Локальная теорема Лапласа.
27. Интегральная теорема Лапласа.
28. Закон больших чисел. Понятие о теореме Чебышева (общий случай). Значение теоремы Чебышева.
29. Закон больших чисел. Теорема Бернулли.
30. Понятие о центральной предельной теореме Ляпунова.
31. Предмет и основные задачи математической статистики.
32. Вариационные ряды. Виды вариации. Границы интервалов в вариационных рядах, величина интервала. Накопленные частоты.
33. Графическое изображение вариационных рядов.

34. Числовые характеристики вариационного ряда. Средняя арифметическая и ее свойства, мода и медиана.
35. Показатели колебаний: вариационный размах, дисперсия, стандартное отклонение, коэффициент вариации.
36. Основные положения теории выборочного метода. Генеральная совокупность и выборка.
37. Законы распределения, применяемые в математической статистике: распределения 2χ , Стюдента, Пирсона.
38. Статистические оценки параметров распределения (сущность теории оценивания): несмещенность, состоятельность, эффективность оценок.
39. Точечные оценки: выборочная средняя, дисперсия, среднее квадратическое отклонение.
40. Точечная оценка генеральной средней по выборочной средней.
41. Точечная оценка генеральной дисперсии. “Исправленные” выборочная дисперсия и среднее квадратическое отклонение.
42. Интервальные оценки. Точность оценки. Доверительная вероятность.
43. Доверительные интервалы для оценки неизвестного значения генеральной средней и генеральной доли.
44. Статистическая проверка гипотез. Статистическая гипотеза: параметрическая и непараметрическая; нулевая и альтернативная. Ошибки I и II рода. Уровень значимости и мощность критерия.
45. Статистический критерий проверки нулевой гипотезы. Наблюдаемое значение критерия. Критическая область. Область принятия гипотезы. Критические точки. Отыскание односторонней и двусторонней критических областей.
46. Основные этапы проверки статистических гипотез.
47. Проверка гипотезы о равенстве выборочной средней и гипотетической генеральной средней нормальной совокупности при известной и неизвестной генеральной дисперсии.
48. Проверка гипотезы о равенстве наблюдаемой относительной частоты и гипотетической вероятности появления события.
49. Проверка гипотезы о равенстве долей признака в двух совокупностях.
50. Проверка гипотезы о законе распределения случайной величины. Критерий согласия Пирсона.

Перечень основной литературы:

1. Гмурман В.Е. Теория вероятностей и математическая статистика. Юрайт. 2020. 480 с.
2. Гнеденко Б.В. Курс теории вероятностей. Едиториал УРСС. 2019. 456 с 8.1.2.

Перечень дополнительной литературы:

1. Вентцель Е.С., Овчаров Л.А. Теория вероятностей и ее инженерные приложения. Физико-математическая библиотека инженера. 1988. 480 с.
2. Джафаров, К.А. Теория вероятностей и математическая статистика учебное пособие / К.А. Джафаров; Министерство образования и науки Российской Федерации, Новосибирский государственный технический университет. - Новосибирск: НГТУ, 2015. - 167 с.: схем. - Библиогр. в кн. - ISBN 978-5-7782-2720-0; то же [Электронный ресурс]. - URL: <http://biblioclub.ru/index.php?page=book&id=438304>
3. Маталыцкий, М.А. Теория вероятностей и математическая статистика: учебник / М.А. Маталыцкий, Г.А. Хацкевич. - Минск: Вышэйшая школа, 2017. - 592 с. - Библиогр. в кн. - ISBN 978-985-06-2855-8