**МЧС России**
**Федеральное государственное бюджетное образовательное учреждение высшего образования**
**«Санкт-Петербургский университет Государственной противопожарной службы Министерства Российской Федерации по делам гражданской обороны, чрезвычайным ситуациям и ликвидации последствий стихийных бедствий»**
Кафедра прикладной математики и безопасности информационных технологий

# Математическая статистика в психологии

**Сборник статей**

**Санкт-Петербург**
**2026**

УДК 519.22
ББК 22.172
ГРНТИ 27.43.17
М34

**Математическая статистика в психологии:** Сб. статей. / Под ред. Т.Н. Антошиной, А.А. Джафаровой, А.А. Кабанова, А.В. Матвеева, Ю.Г. Матвеевой. – СПб.: otvet-akab.ru, 2026. – 48 с.

ISBN 978-5-6055849-3-3

В сборнике кратко рассматриваются актуальные вопросы математической статистики в психологии, объёмом не более одной страницы на вопрос. В него вошли статьи студентов 2 курса очно-заочной формы обучения факультета обеспечения безопасности жизнедеятельности Санкт-Петербургского университета ГПС МЧС России набора 2024 года, обучающихся по специальности 37.05.02 «Психология служебной деятельности». Вступительная статья написана составителем сборника А.А. Кабановым, заключительная статья – Т.Н. Антошиной. Замечания и предложения по сборнику просим присылать по *e-mail:* ***akabanov@inbox.ru.***

*Кра́ткость – это не тала́нт,*
*Но – его сестри́ца.*
*У ребёнка я́ркий бант,*
*Со́лнце но́чью сни́тся!*

## Роль математической статистики в психологии

*А.А. Кабанов, доцент кафедры прикладной математики*
*и безопасности информационных технологий,*
*кандидат юридических наук, доцент*

Науки технические, физика, химия, астрономия и ряд других, активно используют математику для решения своих проблем. Науки, изучающие человека и общество, за исключением экономических, остаются по-прежнему на описательном уровне[1]. Математическая статистика в последние годы приобретает популярность. Этому процессу способствует самостоятельный поиск ответов на актуальные вопросы, имеющий место среди студентов, обучающихся по специальности «Психология служебной деятельности». Каждый студент самостоятельно выбирает тему исследования, ищет информацию, известную по этой теме; самостоятельно формулирует то, что, на его взгляд, представляет более важное. Такие навыки будут полезны студентам не только в математике, но и в профессиональной деятельности. Некоторые статьи данного сборника носят компилятивный характер. Но имеются в сборнике немало статей, представляющих творческое переосмысление известной информации. Навыки, приобретённые студентами при работе над ними, – важный шаг в формировании личности. В процессе написания статей студенты изложили свои мысли кратко, в пределах не более одной страницы, ясно и чётко, выделив главное и отбросив второстепенное.

[1] Синергетика и психология. Материалы круглого стола 10 марта 1997 г. Санкт-Петербург. Доклады / Отв. ред. М.А. Басин, С.В. Харитонов. – СПб.: СПбУВК, 1997. – С. 7.

## Математическая статистика и её основные задачи

*М.Ю. Феофанова, студентка группы ПСД 21.212*

*Математическая статистика* – это раздел математики, посвящённый математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов.

*Предметом математической статистики* является изучение случайных событий и случайных величин по результатам наблюдения. В основе этой дисциплины лежит понятие статистической совокупности (генеральной и выборочной). Статистическая совокупность – это совокупность предметов или явлений, объединённых каким-либо признаком. Результатом наблюдений над статистической совокупностью являются статистические данные – данные о количестве элементов в какой-либо совокупности, обладающих определённым свойством.

Обработка статистических данных методами математической статистики приводит к установлению определённых закономерностей, присущих массовым явлениям.

*Основные задачи математической статистики:*

I. Определение способов сбора (наблюдение, выборка) и группировки (типологические, структурные, аналитические) статистических данных.

II. Разработка методов анализа полученных данных, к которым относятся:

а) оценка неизвестной вероятности события; оценка неизвестной функции распределения; оценка параметров распределения, вид которых известен; оценка зависимости от других случайных величин и т.д.

б) проверка статистических гипотез о виде неизвестного распределения или о значениях параметров известного распределения.

***В психологии математическая статистика*** применяется для обработки данных психологических исследований, проверки гипотез и установления закономерностей между психологическими явлениями. Её используют для количественной оценки, повышения точности результатов и формализации психологических теорий. Математическая статистика в психологии помогает отсеять случайное и выявить закономерное.

## Генеральная и выборочная совокупности, их объёмы. Суть выборочного метода

*М.Ю. Феофанова, студентка группы ПСД 21.212*

В основе математической статистики лежит понятие *статистической совокупности*. Это совокупность предметов или явлений, объединённых каким-либо признаком. Результатом наблюдений над статистической совокупностью являются статистические данные – данные о количестве элементов в какой-либо совокупности, обладающих определённым свойством.

По *полноте охвата* статистические совокупности делятся на генеральные и выборочные. *Генеральная совокупность* – полное множество всех объектов или событий, над которыми проводятся исследования. *Выборочная совокупность (выборка)* – часть генеральной совокупности, отобранная для непосредственного наблюдения, эксперимента или опроса. При составлении выборки можно поступать двумя способами: после того, как объект отобран и над ним произведено наблюдение, он может быть возвращён или не возвращён в генеральную совокупность. В соответствии с этим выборки подразделяются на повторные и бесповторные. Повторные выборки – если объект возвращается в генеральную совокупность, бесповторные выборки – если объект не возвращается в генеральную совокупность. На практике обычно пользуются бесповторным случайным отбором.

*Объёмом совокупности* (генеральной и выборочной) называют число объектов этой совокупности. Объём генеральной совокупности обозначается буквой N, выборочной – буквой n. Например, если из 1000 человек отобрано для исследования 100, то объём генеральной совокупности $N = 1000$, а объём выборки $n = 100$.

*Выборочный метод* – это статистический метод исследования, при котором свойства генеральной совокупности определяются на основе изучения свойств выборки. *Суть выборочного метода* – изучение выборки из генеральной совокупности для получения информации о свойствах всей группы в целом. Ключевой принцип выборочного метода – репрезентативность. То есть выборка должна быть «микромоделью» генеральной совокупности, её структура и характеристики должны максимально точно отражать структуру и характеристики всей генеральной совокупности.

***В психологии выборочный метод*** позволяет делать выводы о больших группах людей (генеральной совокупности) на основе изучения небольшой, но репрезентативной части (выборки).

## Понятие закона больших чисел

*М.Ю. Феофанова, студентка группы ПСД 21.212*

*Закон больших чисел* – принцип, описывающий результат выполнения одного и того же эксперимента много раз. Закон больших чисел утверждает, что при многократном повторении одного и того же случайного эксперимента, среднее значение выборки из множества результатов стремится к математическому ожиданию этого эксперимента. То есть, чем больше проводится испытаний (например, бросков монеты), тем точнее относительная частота события (например, выпадение орла), приближается к его истинной вероятности (1/2), а среднее значение выборки приближается к теоретическому среднему. В более общем виде *закон больших чисел* говорит, что среднее арифметическое приближается к математическому ожиданию по мере роста количества наблюдений.

*Закон больших чисел* состоит из нескольких теорем, в которых доказывается приближение средних характеристик к некоторым постоянным значениям при соблюдении определённых условий, то есть при большом числе испытаний их средний результат перестает быть случайным и может быть предсказан с достаточной точностью.

*Основные теоремы закона больших чисел – теорема Бернулли* (в серии испытаний с одинаковой вероятностью появления события, частота появления события стремится к его вероятности при увеличении числа испытаний); *теорема Чебышёва* (устанавливает закон больших чисел для последовательности независимых случайных величин с конечными математическими ожиданиями и дисперсиями; также в закон больших чисел входят «Усиленный» закон больших чисел (теорема Колмогорова); теорема Маркова; теорема Пуассона (является частным случаем закона больших чисел, описывающим поведение редких событий в большом числе испытаний).

***В психологии закон больших чисел*** применяется для выявления общих закономерностей в массовых психологических явлениях, где при большом количестве наблюдений проявляются устойчивые тенденции.

Примеры применения:

- проведение опросов и анкетирования среди большой выборки для получения более надёжных данных;

- при проведении психологических тестов большой объём данных позволяет получить более точные выводы;

- в социологических исследованиях изучение поведения больших групп позволяет точнее выявить общие психологические тенденции.

# Теорема Чебышёва и её следствия

*Е.А. Михайлова, студентка группы ПСД 21.212*

*Теорема Чебышёва* свидетельствует о том, что случайная величина редко сильно отклоняется от своего среднего значения. Если мы знаем среднее значение и насколько сильно данные обычно разбросаны (дисперсия), теорема показывает, насколько вероятно получить очень большое отклонение от среднего. Использование данного заключения удобно, когда требуется оценить точность различных видов измерений или провести статистический анализ относительно устойчивости и воспроизводимости результатов.

Формула теоремы записывается так:

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\sigma}{\varepsilon^2},$$

где:

X — случайная величина, то есть переменная, которая может принимать разные значения;

$\mu$ — среднее значение величины, то есть математическое ожидание;

σ — дисперсия, показывающая, насколько сильно значения могут отклоняться от среднего;

ε — любое положительное число, указывающее, насколько далеко от среднего мы хотим проверить вероятность.

Главное следствие – усиление доверия к среднему арифметическому при большом числе независимых наблюдений: чем больше измерений, тем меньше вероятность большой ошибки в оценке среднего.

Например, если в группе студентов средняя оценка по тесту равна 70 баллам, а дисперсия результатов составляет 16, то вероятность того, что у студента будет оценка, отклоняющаяся от среднего на 4 балла или больше, не превышает $\frac{16}{4^2} = 1$. Это значит, что с подобной разницей могут столкнуться все или почти все, лучше посчитать для ε=8: тогда вероятность отклонения на 8 и больше баллов будет не более $\frac{16}{8^2} = 0{,}25$, то есть не более 25%.

***В психологии теорема Чебышёва*** полезна при анализе результатов исследований и тестирований. Она помогает понять, насколько выборочные данные устойчивы и какова вероятность, что средний показатель будет далёким от реального среднего по населению. Благодаря этой теореме исследователи могут делать выводы с уверенностью в статистической значимости, учитывая вариативность данных.

# Центральная предельная теорема Ляпунова

*Е.А. Михайлова, студентка группы ПСД 21.212*

*Центральная предельная теорема Ляпунова* утверждает следующее: если сложить вместе много разных, независимых случайных событий, то итоговый результат будет стремиться к нормальному распределению. *Важное условие* – ни одно из событий не должно вносить решающий, слишком большой вклад в общую сумму. Даже если каждое отдельное событие имеет своё, уникальное распределение, их сумма будет вести себя предсказуемо.

Итоговая оценка студента за семестр складывается из множества разных частей: десяти домашних заданий, пяти контрольных работ, двух проектов и одного экзамена. Каждая из оценок – отдельная случайная величина (студент мог лучше или хуже подготовиться, ему мог попасться лёгкий или сложный билет). Поскольку итоговая оценка – это сумма множества мелких, независимых друг от друга оценок, то распределение итоговых баллов у всей группы студентов будет очень похоже на нормальное.

Многие *психологические характеристики*, такие как уровень интеллекта или тревожности, формируются под влиянием огромного числа независимых факторов: генетика, воспитание, окружение, жизненный опыт и т.д. Каждый из факторов вносит свой небольшой вклад. Центральная предельная теорема позволяет предполагать, что распределение характеристик в большой группе людей будет близко к нормальному.

***Пример из психологии***: моделирование уровня тревоги.

Допустим, на индивида влияют множество независимых факторов: тревожность в разных контекстах (работа, семья, здоровье), сон, кофеин, физическая активность, социальная поддержка, оценка рисков и т.п. Каждое влияние вносит небольшой вклад в суммарный уровень тревоги. При достаточно большом числе факторов и при соблюдении условий Ляпунова (различные величины имеют конечные моменты и не доминируют над остальными) сумма этих вкладов может быть распределена приблизительно нормально. Это объясняет наблюдение, что общий уровень тревоги в группе часто распределён близко к нормальному, даже если каждое отдельное влияние распределено по-другому.

## Основные способы отбора элементов в выборку

*Е.А. Михайлова, студентка группы ПСД 21.212*

Существует несколько *основных способов отбора*:

- *простой случайный отбор*. Каждый элемент большой группы имеет абсолютно равные шансы попасть в выборку. Похоже на то, как вытаскивать имена из шляпы, не глядя;
- *систематический отбор*. Элементы отбираются через определённый, равный интервал. Сначала выбирается случайная отправная точка, а затем берётся, например, каждый десятый или каждый сотый элемент из общего списка, например: 7, 17, 27 и т.д.;
- *стратифицированный (районированный) отбор*. Вся большая группа сначала делится на небольшие однородные подгруппы (страты) по какому-то важному признаку (например, по возрасту, полу, уровню образования). Затем из каждой подгруппы берётся случайная выборка. Это гарантирует, что все ключевые подгруппы будут представлены;
- *кластерный (гнездовой) отбор*. Вся группа делится на множество мелких, обычно географических, групп (кластеров), например, по городам или районам. Затем случайным образом выбираются несколько кластеров, и в них опрашиваются все элементы.

*Пример:* представим, что исследователю нужно опросить 100 студентов из университета, в котором учится 1000 человек. При простом случайном отборе он присвоил бы каждому студенту номер от 1 до 1000 и с помощью генератора случайных чисел выбрал бы 100 номеров. При систематическом отборе он взял бы список всех студентов, отсортированный по алфавиту, и выбрал бы, например, каждого десятого (1000/100=10). При стратифицированном отборе он мог бы разделить всех студентов по курсам (1-й, 2-й, 3-й, 4-й) и затем из каждой группы (курса) случайным образом выбрать по 25 человек. При кластерном отборе он мог бы рассматривать каждую учебную группу как кластер, случайным образом выбрать 5 групп из 50 и опросить всех студентов в пяти группах.

***В психологии правильный отбор*** выборки является основой любого качественного исследования. Цель большинства психологических экспериментов – распространить выводы, полученные на небольшой группе людей, на всю интересующую популяцию (например, на всех подростков или всех людей с определённым расстройством).

# Понятие и виды вариационных рядов распределения частот и частостей

*В.А. Бабенко, студентка группы ПСД 21.212*

*Вариационный ряд* – ранжированное в порядке возрастания или убывания распределение совокупности по количественному признаку, состоит из двух элементов: *варианта и частота* (частость). *Варианта* – это отдельное значение признака, которое он принимает в вариационном ряду. *Частота* – численность отдельных вариант или каждой группы вариационного ряда, то есть это числа, показывающие, как часто встречаются те или иные варианты в ряду распределения. *Частости* – это частоты, выраженные в долях единицы или в процентах к итогу. Сумма частостей равна 1 или 100%. Сумма всех частот называется *объёмом совокупности.* Чтобы рассчитать *относительную частоту* $w_i$ нужно каждую из частот $n_i$ разделить на объём выборки n (сумма всех $w_i$ выборки равна 1). *Виды* вариационных рядов**:** *чётный и нечётный, большой* (больше 30 наблюдений) и *малый. Простой* – ряд, в котором каждая варианта встречается один раз, т.е. частоты равны единице. При увеличении числа наблюдений встречаются и повторяющиеся значения вариант. *Сгруппированный* – вариационный ряд, в котором варианты объединены в группы по их величине в пределах определённого интервала с указанием частоты повторяемости всех вариант, входящих в группу. *Дискретный* – основан на дискретных (прерывных) признаках, имеющих только целые значения, например, число детей в семье. *Интервальный* – основан на непрерывных признаках, когда отдельные значения отличаются друг от друга на сколь угодно малую величину, например, рост студентов. Его составляют для упрощения последующих вычислений при очень большом числе единиц наблюдения (больше 1000). *Графически* вариационные ряды могут быть представлены *в виде таблицы* (обозначения: варианты – $x_i$, частота варианта – $n_i$), либо *в виде полигона частостей* (для дискретного вариационного ряда, это ломаная, аналог многоугольника распределения из теории вероятностей), или *гистограммы частот* (для непрерывных значений).

***В психологии*** вариационные ряды распределения изучают для анализа данных и выявления закономерностей. Полученные в ходе психологического эксперимента данные представляют набор чисел, выявить закономерность их изменения (варьирования) сложно. Анализ вариационного ряда в психологии позволяет сделать выводы на основе полученных данных, например, о том, как стресс влияет на когнитивные функции (анализ результатов тестов, проведённых на группе испытуемых). Также можно оценить достоверность результатов: вариационный ряд помогает проверить, насколько полученный результат выборки отражает свойства генеральной совокупности.

# Построение интервальных рядов

*В.А. Бабенко, студентка группы ПСД 21.212*

Если изучаемая случайная величина непрерывная (её значения могут как угодно мало отличаться друг от друга), то следует строить интервальный вариационный ряд распределения. Для этого весь интервал разбивают на ряд частичных интервалов и подсчитывают частоту попадания значений в каждый частичный интервал. *Интервальным вариационным рядом* называют упорядоченную совокупность интервалов варьирования значений случайной величины с соответствующими частотами попаданий в каждый из них значений величины. *Алгоритм построения* интервального вариационного ряда следующий:

1. Определить *оптимальное количество интервалов k.* Число интервалов k зависит из объёма выборки, рассчитывается по формуле Стержеса и округляется до целого числа (так как количество групп не может быть дробным): $k = 1 + 3{,}32 \cdot \lg n$, где n − общее число изучаемых единиц совокупности. Также можно использовать таблицу:

| Объём выборки | 6-25 | 25-40 | 40-60 | 60-100 | 100-200 | более 200 |
|---|---|---|---|---|---|---|
| Число интервалов | 3-4 | 5-6 | 6-8 | 7-10 | 8-12 | 10-15 |

2. Вычислить разность между максимальным и минимальным значениями признака первичного ряда (определяется *размах вариации, R*), $R = x_{max} - x_{min}$.

3. Определить *ширину интервалов h,* для этого размах делим на оптимальное число интервалов, которое мы вычислили в п.1.

4. Установить для каждого интервала его *верхнюю и нижнюю границы*. Нижняя граница первого интервала находится по формуле: $x_{нижн} = x_{min} - 0{,}5 \cdot h$. *Промежуточные интервалы* получают, прибавляя к концу предыдущего интервала длину частичного интервала h. Построение шкалы интервалов продолжается, пока величина $x_{hi}$ удовлетворяет соотношению: $x_{hi} < x_{max} + 0{,}5 \cdot h$.

5. В соответствии со шкалой интервалов произвести *группировку результатов наблюдений.* Подсчёт частот для каждого интервала удобно проводить методом «конвертиков». Этот метод состоит в том, что попадание значения случайной величины в тот или иной интервал, отмечается точкой, и чёрточкой. В результате каждому десятку будет соответствовать фигура, похожая на конверт.


***В психологических исследованиях*** чаще всего используют равноинтервальные ряды для систематизации данных, полученных по непрерывному признаку (например, обработка баллов, полученных испытуемыми по шкале реактивной тревожности, или изучение уровня интеллектуального развития человека).

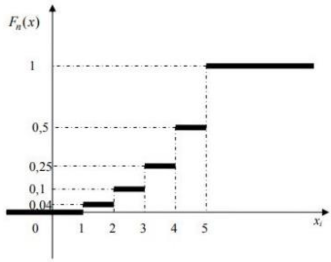
# Эмпирическая функция распределения и её связь с теоретической функцией распределения

*В.А. Бабенко, студентка группы ПСД 21.212*

*Эмпирической функцией распределения $F_n(x)$* называется относительная частота того, что признак (случайная величина X) примет значение, меньшее заданного x. Для данного x эмпирическая функция распределения представляет накопленную частость $W_x^{нак}$, т.е. $F_n(x) = W(X < x) = W_x^{нак}$. Эмпирическая функция распределения обладает следующими *свойствами*:

1. Значение эмпирической функции принадлежит отрезку [0; 1].
2. Эмпирическая функция является неубывающей.
3. Если $x_1$ – наименьшее значение варианты, а $x_k$ – наибольшее, то $F_n(x) = 0$ при $x \le x_1$ и $F_n(x)=1$ при $x > x_k$.

В отличие от эмпирической функции распределения выборки $F_n(x)$ интегральную функцию F(x) распределения генеральной совокупности называют *теоретической функцией распределения. Различие между эмпирической и теоретической функциями* состоит в том, что теоретическая функция F(x) определяет вероятность события X < x, а эмпирическая функция $F_n(x)$ определяет относительную частоту этого же события. При больших n значения $F_n(x)$ и F(x) мало отличаются одно от другого. Следовательно, $F_n(x)$ может служить для оценки F(x).




В случае *дискретного* вариационного ряда Fn(x) представляет собой разрывную ступенчатую функцию.

Для интервального вариационного ряда имеем лишь значения функции Fn(x) на концах интервала, поэтому для графического изображения этой функции целесообразно её доопределить, соединив точки графика, соответствующие концам интервалов, отрезками прямой. В результате полученная ломаная совпадает с кумулятой.

Изучение эмпирической и теоретической функций распределения ***в психологии*** необходимо для понимания распределения данных, которое связывает значения переменной и частоты встречаемости, полученные на группе объектов. В психологических исследованиях чаще всего встречается нормальное (гауссовское) распределение значений признака по частоте. Оно возникает, когда на значения случайной переменной, например, на признак, характеризующий испытуемых (интеллект, тревожность), оказывают влияние множество разных воздействий.

## Понятие накопленной частоты и кумуляты распределения частот дискретного и интервального рядов распределения

*В.В. Чванова, студентка группы ПСД 21.212*

*1. Накопленная частота* (кумулятивная частота) – это количество вариантов со значением признака, меньшим определённого значения. Она определяется путём последовательного прибавления к частоте первого интервала частот последующих интервалов (суммирования всех предшествующих частот). То есть, накопленная частота показывает общее количество чего-то на каждый момент времени (или для каждого значения), считая все предыдущие значения.

*2. Кумулята* (кумулятивная кривая) – ломаная линия, составленная по накопленным частотам. Накопленные частоты наносятся в виде ординат; соединяя вершины отдельных ординат отрезками прямой, получают ломаную линию, имеющую неубывающий вид.

***Применение в психологии:***

•Описывают и сравнивают распределения данных.

•Определяют процентили (например, уровень, который превышают 10% испытуемых).

•Интерпретируют результаты тестов (сравнение с нормой).

•Оценивают эффективность вмешательств.

*Пример*:

Психолог измерил уровень депрессии у группы людей. Он хочет узнать, какая доля людей имеет лёгкую депрессию (значения 5 и меньше).

•Считает накопленную частоту для значения "5" (сумма частот значений 0, 1, 2, 3, 4 и 5).

•Делит накопленную частоту на общее число людей в группе.

•Получает кумуляту для значения "5".

*Результат:*

Кумулята показывает, какая доля людей в группе имеет уровень депрессии 5 или ниже (т.е. лёгкую депрессию). Это позволяет психологу сделать вывод о распространенности лёгкой депрессии в данной группе.

## Полигон распределения частот дискретного и интервального рядов распределения, его построение

*В.В. Чванова, студентка группы ПСД 21.212*

*Полигон распределения частот* – это ломаная линия, которая графически изображает вариационный ряд (статистическое распределение). Он может быть построен для дискретного и интервального рядов.

*Полигон (дискретный ряд):*

•Линейный график, показывающий частоту дискретных значений.

•Строится по точкам (значение, частота) и соединяется линиями.

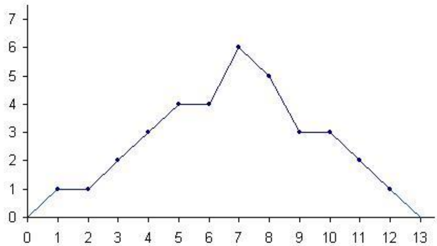



Рис. 1. Пример полигона (дискретный ряд)

*Гистограмма (интервальный ряд):*

•Столбчатый график, показывающий частоту значений в интервалах.

•Высота столбца соответствует частоте интервала.

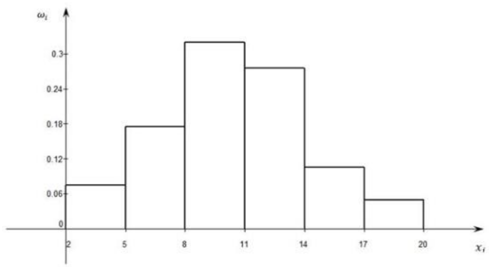



Рис. 2. Пример гистограммы (интервальный ряд)

***Применение в психологии:***

Полигон частот позволяет визуально оценить форму распределения психологических данных (например, нормальное, асимметричное), центральную тенденцию и изменчивость. Удобен для сравнения нескольких распределений на одном графике.

## Гистограмма распределения частот интервального ряда

*М.И. Кручай, студентка группы ПСД 21.212*

*Гистограмма* – это графическое представление распределения частот для интервального вариационного ряда. Она строится в виде прямоугольников, *основание* которых соответствует *интервалам разбиения, а высота – частоте или относительной частоте наблюдений.*

*Построение.* Пусть дан интервальный вариационный ряд:

| Интервал | $[x_0;x_1]$ | $[x_1;x_2]$ | … | $[x_{k-1};x_k)$ |
|---|---|---|---|---|
| Частота $n_1$ | $n_1$ | $n_2$ | … | $n_k$ |

- $x_0, x_1, \ldots, x_k$ – границы интервалов;
- $h_i = x_i - x_{i-1}$ – длина i-го интервала;
- $n_i$ – число элементов, попавших в интервал;
- $N = \sum^{k}_{i=1} n_i$ – общее число наблюдений;
- k – это количество интервалов (групп), на которые разбита вариация признака.

*Формула для высоты столбца*

Чтобы площади прямоугольников были пропорциональны частотам, высота $h_i^*$ выбирается так, чтобы выполнялось условие:

$$h_i^* = n_i / (N \cdot h_i),$$

тогда площадь прямоугольника: $S_i = h_i^* \cdot h_i = n_i / N$, что соответствует *относительной частоте интервала*.

*Интерпретация:*

*Ось X* – интервалы признака. *Ось Y* – частоты или плотности частот.

Сумма площадей всех прямоугольников = 1, если используется относительная частота.

*Назначение гистограммы* – наглядное представление распределения признака; оценка формы распределения (симметричность, скошенность, количество мод); приближённая оценка плотности распределения вероятности.

***Применение в психологии:***

Гистограммы широко используются в психологии для визуализации распределения психологических показателей – например, уровня тревожности, интеллекта, самооценки, экстраверсии и др. С их помощью можно определить, имеет ли распределение нормальный характер, есть ли перекос (например, в сторону высоких или низких значений), а также выявить наличие нескольких групп испытуемых (мод) в выборке. Это помогает психологу понять структуру исследуемой выборки и выбрать подходящие статистические методы для дальнейшего анализа.

# Средняя выборочная, её нахождение и сравнение с математическим ожиданием

*Д.А. Пестова, студентка группы ПСД 21.212*

*Выборочное среднее* (или среднее арифметическое выборки) – наиболее важная и часто используемая описательная статистика, которая характеризует центр расположения данных выборки.

*Выборочной средней* $\overline{x_в}$ называют среднее арифметическое значение признака выборочной совокупности.

*Математическое ожидание* – это теоретическая, числовая характеристика генеральной совокупности, а не выборки. Оно представляет собой среднее значение случайной величины в теории вероятностей.

*Математическое ожидание* – это среднее значение случайной величины, которое мы ожидали бы получить в пределе, если бы эксперимент повторялся бесконечное число раз. Это теоретический параметр генеральной совокупности, который часто бывает неизвестен на практике.

*Сравнение выборочного среднего и математического ожидания*

Несмотря на схожую роль, это принципиально разные понятия.

| **Признак для сравнения** | **Выборочное среднее ($\bar{x}$)** | **Математическое ожидание (μ, M[X])** |
|---|---|---|
| Природа | Статистика (функция от элементов выборки). Это случайная величина, так как зависит от конкретной реализации выборки. | Теоретический параметр генеральной совокупности. Это константа (фиксированное, неслучайное число) для данного распределения. |
| Источник данных | Рассчитывается на основе конечной выборки (n наблюдений) | Определяется на основе закона распределения всей генеральной совокупности (вероятностей $p_i$ или функции плотности f(x)). |
| Вычисление | Рассчитывается по конечной формуле $\bar{x} = (1/n) * \Sigma x_i$ на основе имеющихся данных. | Вычисляется теоретически через суммирование или интегрирование с использованием знаний о распределении. |
| Доступность | Всегда может быть вычислено, так как основывается на реальных данных. | Часто неизвестно на практике и является тем объектом, который мы хотим оценить по выборке. |

Средняя выборочная (среднее арифметическое выборки) позволяет оценить типичный уровень выраженности *психологического признака* в исследованной группе, например, средний уровень тревожности. Сравнивая её с теоретическим математическим ожиданием (например, с нормой по тесту), психолог может сделать вывод о том, насколько его выборка репрезентативна или отличается от генеральной совокупности, что критично для проверки гипотез.

## Понятие моды и её нахождение для дискретного и интервального рядов

*Д.А. Пестова, студентка группы ПСД 21.212*

Для характеристики структуры статистической совокупности применяются показатели, которые называют структурными средними. К ним относятся мода и медиана.

*Мода* ($M_o$) – чаще всего встречающийся вариант. Модой называется значение признака, которое соответствует максимальной точке теоретической кривой распределений.

Мода представляет наиболее часто встречающееся или типичное значение, применяется в коммерческой практике для изучения покупательского спроса и регистрации цен.

В дискретном ряду мода – это варианта с наибольшей частотой. В интервальном вариационном ряду модой считают центральный вариант интервала, который имеет наибольшую частоту (частность).

В пределах интервала надо найти то значение признака, которое является модой.

$$M_0 = x_0 + h \cdot \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})}$$

где $x_0$ – нижняя граница модального интервала;

$h$ – величина модального интервала;

$f_m$ – частота модального интервала;

$f_{m-1}$ – частота интервала, предшествующего модальному;

$f_{m+1}$ – частота интервала, следующего за модальным.

Мода зависит от величины групп, от точного положения границ групп.

*Мода* – число, которое в действительности встречается чаще всего (является величиной определённой), в практике имеет самое широкое применение (наиболее часто встречающийся тип покупателя).

***В психологии*** мода как наиболее часто встречающееся значение полезна для определения самого «популярного» результата, например, наиболее типичного числа ошибок в когнитивном тесте или предпочитаемого стиля конфликтного поведения в группе. В интервальном ряду (например, по шкале тревожности) модальный интервал показывает диапазон, в который попадает наибольшее число испытуемых, что помогает быстро выявить зону наибольшего «скопления» данных.

## Понятие медианы и её нахождение для дискретного и интервального рядов

*Д.А. Пестова, студентка группы ПСД 21.212*

*Медиана* ($M_e$) – это величина, которая делит численность упорядоченного вариационного ряда на две равные части: одна часть имеет значения варьирующего признака ме́ньшие, чем средний вариант, а другая – бо́льшие. *Медиана* – это элемент, который больше или равен и одновременно меньше или равен половине остальных элементов ряда распределения.

*Свойство* медианы заключается в том, что сумма абсолютных отклонений значений признака от медианы меньше, чем от любой другой величины.

*Применение медианы* позволяет получить более точные результаты, чем при использовании других форм средних.

*Порядок нахождения медианы в интервальном вариационном ряду* следующий: располагаем индивидуальные значения признака по ранжиру; определяем для данного ранжированного ряда накопленные частоты; по данным о накопленных частотах находим медианный интервал:

$$M_e = x_0 + i \times \frac{\left(\frac{\sum f}{2} - S_{m-1}\right)}{f_m},$$

где $x_0$ – нижняя граница медианного интервала;

$i$ – величина медианного интервала;

$\frac{\sum f}{2}$ – полусумма частот ряда;

$S_{m-1}$ – сумма накопленных частот, предшествующих медианному интервалу;

$f_m$ – частота медианного интервала.

Медиана делит численность ряда пополам, следовательно, она там, где накопленная частота составляет половину или больше половины всей суммы частот, а предыдущая (накопленная) частота меньше половины численности совокупности.

***В психологии** м*едиана, указывающая на центральное значение в упорядоченном ряду, особенно ценна в психологии для описания данных с асимметричным распределением или наличием выбросов, так как она менее чувствительна к ним, чем среднее арифметическое. Например, она позволяет определить «серединного» испытуемого по времени реакции, разделив выборку на две равные половины, что даёт более устойчивую оценку центральной тенденции, не искажённую единичными очень высокими или низкими показателями.

## Выборочная дисперсия и выборочное среднеквадратическое отклонение

*М.И. Кручай, студентка группы ПСД 21.212*

*Выборочная дисперсия и выборочное среднеквадратическое отклонение (СКО)* – это статистические характеристики, используемые для описания разброса данных в выборке. Выборочная дисперсия показывает, насколько элементы выборки отклоняются от выборочного среднего.

Формула несмещённой выборочной дисперсии:

$$s^2 = (1 / (n - 1)) \cdot \Sigma (x_i - \bar{x})^2,$$

где:

$x_i$ – i-й элемент выборки

$\bar{x}$ – выборочное среднее ($\bar{x} = (1 / n) \cdot \Sigma x_i$)

n – объём выборки

$s^2$ – выборочная дисперсия.

*Выборочное среднеквадратическое отклонение (СКО)* – это квадратный корень из выборочной дисперсии.

$$s = \sqrt{s^2}$$

СКО измеряется в тех же единицах, что и сами данные, и показывает среднее отклонение значений выборки от её среднего.

*Пример:*

Пусть выборка: 5, 7, 8. Выборочное среднее: $\bar{x} = (5 + 7 + 8)/3 = 6.67$.

Дисперсия: $s^2 = [(5 - 6.67)^2 + (7 - 6.67)^2 + (8 - 6.67)^2] / (3 - 1) \approx 2.33$.

СКО: $s = \sqrt{2.33} \approx 1.53$.

*Назначение*

- дисперсия и СКО – меры разброса данных;
- используются для оценки вариабельности выборки;
- являются основой для статистических гипотез, доверительных интервалов и других методов анализа.

***Применение в психологии:***

Выборочная дисперсия и среднеквадратическое отклонение применяются в психологии для оценки индивидуальных различий и вариативности психологических характеристик в группе — например, различий в уровне мотивации, когнитивных способностей или эмоциональной устойчивости. Эти показатели позволяют судить о степени однородности выборки и о том, насколько сильно испытуемые отличаются друг от друга. Они также необходимы при построении доверительных интервалов, проверке статистических гипотез и оценке надёжности психометрических тестов.

# Коэффициент асимметрии вариационного ряда и его смысловое значение

*А.В. Бычкова, студентка группы ПСД 21.212*

*Коэффициент асимметрии* (или коэффициент асимметрии распределения) – это статистическая мера, которая описывает степень и направление асимметрии распределения данных. Он показывает, насколько распределение отклоняется от симметричного относительно среднего значения. *Положительный* (наклон вправо), *Отрицательный* (наклон влево), *Ноль* (симметричный).

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

где $\mu_3$ – третий центральный момент;

$\sigma^3$ – стандартное отклонение вариационного ряда.

*Типы асимметрии:*

1. *Симметричное распределение*: если коэффициент асимметрии равен 0, то распределение симметрично (нормальное распределение).
2. *Положительная асимметрия*: если As > 0, то распределение имеет правостороннюю асимметрию (длинный «хвост» справа). Это означает, что более высокая часть данных сконцентрирована слева от среднего значения.
3. *Отрицательная асимметрия*: если As < 0, то распределение имеет левостороннюю асимметрию (длинный «хвост» слева). Это указывает на то, что большая часть данных находится справа от среднего значения.

***В психологических исследованиях*** коэффициент асимметрии вариационного ряда может применяться для анализа распределения различных психологических показателей (например, уровней личностных качеств, результатов тестов, показателей психофизиологических измерений и др.).

Асимметрия помогает выявить наличие перекоса в данных, что может говорить о преобладании высоких или низких значений; оценить однородность выборки; принять решение о применимости статистических методов, ориентированных на симметричные распределения.

Коэффициент асимметрии вариационного ряда – это важный статистический показатель формы распределения в психологических данных, который позволяет оценить скошенность распределения признака относительно его среднего значения и принять соответствующие методологические решения в исследовании и интерпретации результатов.

# Коэффициент эксцесса вариационного ряда и его смысловое значение

*А.В. Бычкова, студентка группы ПСД 21.212*

*Коэффициент эксцесса* – это статистическая мера, которая описывает «выпуклость» или «плоскость» распределения данных относительно нормального распределения. Он показывает, насколько *распределение* данных отличается от нормального в плане наличия тяжёлых хвостов или пиков. Лептокуртическое (тяжёлохвостое), мезокуртическое (нормальное), Платикуртическое (светлохвостое).

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

где $\mu_4$ – центральный момент четвертого порядка;
$\sigma^4$ – стандартное отклонение вариационного ряда.

*Типы эксцесса:*

1. *Нормальное распределение*: если $Ex = 0$, то распределение является нормальным (или близким к нормальному).
2. *Положительный эксцесс*: если $Ex > 0$, то распределение имеет высокий пик (много значений около среднего) и более тяжёлые хвосты. Это указывает на наличие экстремальных значений и может свидетельствовать о том, что данные имеют больше выбросов.
3. *Отрицательный эксцесс*: если $Ex < 0$, то распределение имеет более плоскую форму и более лёгкие хвосты. Это означает, что данные менее концентрированы вокруг среднего значения и имеют меньше выбросов.

*Коэффициент эксцесса* важен для:

- анализа данных: помогает понять, как данные распределены и насколько они отклоняются от нормального распределения;
- выбор статистических методов: некоторые статистические методы и тесты предполагают нормальность данных. Знание эксцесса может помочь выбрать правильный метод анализа;
- оценка рисков: в финансовом анализе, например, высокий эксцесс может указывать на повышенные риски из-за возможных экстремальных значений.

***В психологии*** этот показатель важен для анализа характера распределения признаков: например, при анализе асимметрии и эксцесса вариационного ряда можно выявлять особенности психологических данных и изменения признаков, понять природу и структуру исследуемых психологических данных, а также выявить отклонения от нормального поведения признаков.

## Понятие точечной оценки и её интерпретация как случайной величины

*Н.В. Метушевская, студентка группы ПСД 21.212*

*Точечная оценка* – это единственное значение (статистика), вычисленное на основе выборочных данных (n), которое используется для оценки неизвестного параметра генеральной совокупности (θ). Например, выборочное среднее (M) является точечной оценкой генерального среднего (μ), а выборочная дисперсия ($s^2$) – точечной оценкой генеральной дисперсии ($\sigma^2$).

Если говорить просто, то *точечная оценка* – это одно число, которое мы вычисляем на основе наших данных (выборки), чтобы угадать неизвестный нам параметр всей большой группы (генеральной совокупности).

*Например, я психолог,* и хочу узнать средний уровень тревожности всех студентов в моём университете.

Опросить всех – невозможно. Поэтому я опрашиваю случайную группу из 30 человек, вычисляю их средний балл тревожности, например, 58 баллов.

Это число (58) и есть *точечная оценка* среднего уровня тревожности для всего университета.

Теперь, почему эта оценка – *случайная величина*? Потому что, если бы я, как исследователь, набрала другую случайную группу из 30 студентов, их средний балл, скорее всего, был бы другим – например, 55 или 61. То есть значение оценки зависит от случая – от того, какие именно люди попали в мою выборку. Каждую новую выборку можно считать новым «испытанием», которое даёт новое числовое значение оценки.

Таким образом, *до того,* как мы провели эксперимент и собрали данные, точечная оценка – это потенциальная, переменная величина, которая может принимать разные значения. Именно это и делает её *случайной величиной*. А *после того*, как данные собраны и число вычислено, мы получаем одно конкретное значение этой оценки – точку на числовой прямой, отсюда и название «точечная».

*Коротко:* Точечная оценка – это наше лучшее предположение о параметре, основанное на выборке. А так как выборка каждый раз случайна, то и наше предположение (оценка) – это случайная величина, которая может меняться от исследования к исследованию.

# Основные свойства точечных оценок

*Н.В. Метушевская, студентка группы ПСД 21.212*

Чтобы наша «догадка» о параметре всей группы была хорошей и нам можно было ей доверять, она должна обладать тремя главными свойствами. Объясню их на примере из психологии.

Представьте, что мы измеряем самооценку у группы подростков с помощью нашего теста.

1. *Состоятельность*. Это свойство означает, что чем больше людей мы опросим, тем точнее будет наш результат. Если мы опросим 10 человек, наша оценка может быть довольно случайной. Но если мы будем увеличивать выборку – 50, 100, 1000 человек – то наше вычисленное среднее значение самооценки будет всё ближе и ближе подбираться к *истинному* среднему значению самооценки всех подростков. Хорошая оценка «учится на опыте»: больше данных – лучше результат.

2. *Несмещённость*. Это значит, что наш метод оценки не должен постоянно завышать или занижать результат. Допустим, наш тест самооценки составлен так, что люди всегда отвечают чуть более позитивно, чем есть на самом деле. Тогда средний балл по выборке будет *систематически* завышать истинный уровень самооценки. Это пример *смещённой* оценки. *Несмещённая* же оценка не имеет такой системной ошибки; она может иногда давать результат чуть выше, иногда чуть ниже, но в среднем, за много таких исследований, она будет попадать точно в цель. Классический пример – мы знаем, что выборочная дисперсия (мера разброса) является смещённой оценкой, и чтобы это исправить, мы делим не на количество человек в выборке (n), а на (n-1).

3. *Эффективность.* Это свойство отвечает на вопрос: «А насколько точно наша оценка бьёт в цель при одном и том же объёме данных?». Допустим, у нас есть два способа оценить среднюю самооценку: по среднему арифметическому и по медиане. Обе оценки могут быть *состоятельными и несмещёнными*. Но если значения среднего арифметического от выборки к выборке колеблются меньше, чем значения медианы, то среднее арифметическое является *более эффективной* оценкой. Эффективная оценка – это «снайпер»: у неё самый маленький разброс вокруг истинного значения параметра.

*Итог:* Идеальная точечная оценка – это та, которая с ростом выборки стремится к истинному значению (*состоятельность*), не имеет систематической ошибки (*несмещённость*) и имеет при этом наименьший возможный разброс (*эффективность*).

## Точечная оценка для генеральной средней

*Н.В. Метушевская, студентка группы ПСД 21.212*

В своей работе мы постоянно имеем дело с выборками – группой клиентов, участников исследования – и на основе этих данных нам нужно сделать вывод о более широкой популяции (например, о всех людях с определённым расстройством). *Генеральная средняя* – это и есть тот самый средний показатель (уровень тревожности, интеллекта, агрессии и т.д.) для всей этой большой, зачастую недоступной для полного изучения, группы. Это наш главный, но неизвестный, параметр, который мы хотим узнать.

*Точечная оценка для генеральной средней* – это наше лучшее предположение, основанное на конкретных данных, которое мы можем вычислить. Этой оценкой является *выборочное среднее арифметическое*.

Если у нас есть выборка объёма n, представленная значениями $x_1, x_2, ..., x_n$, то выборочное среднее вычисляется по формуле:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Представьте, что мы хотим узнать *средний уровень страха перед публичными выступлениями* у всех студентов нашего вуза (это *генеральная средняя*). Опросить всех (несколько тысяч) студентов невозможно.

Мы берём *выборку* – случайным образом отбираем, например, 50 студентов ($n=50$) и измеряем у каждого уровень страха по шкале от 0 до 100. После опроса у нас есть 50 чисел ($x_1, x_2, ..., x_{50}$).

Мы складываем все 50 результатов и делим на 50. Полученное число (например, 68 баллов) – это и есть *выборочное среднее* $\bar{x}=68$.

Это число *68* и является *точечной оценкой*. Мы говорим: «На основании нашего исследования, мы предполагаем, что средний уровень страха у всех студентов университета составляет примерно 68 баллов». Мы заменили неизвестный параметр большой группы (m) на одно конкретное число ($\bar{x}$), рассчитанное по малой, доступной нам части этой группы.

*Точечной оценкой для генеральной средней является выборочное среднее арифметическое*.

Это главный и наилучший статистический «заменитель», который мы вычисляем по данным нашей выборки, чтобы сделать обоснованное предположение о среднем значении признака во всей генеральной совокупности.

## Точечная оценка для генеральной дисперсии

*М.А. Воронин, студент группы ПСД 21.212*

*Точечной оценкой* называется число, которое используют для оценки параметра генеральной совокупности.

*Выборочная* дисперсия:

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

*Генеральная* дисперсия:

$$\sum_{i=1}^{N} \frac{(x_i - \overline{x_\text{г}})^2}{N}$$

*Теорема.* Выборочная дисперсия повторной и бесповторной выборок есть смещённая и состоятельная оценка генеральной дисперсии $\sigma^2$. Вычисленная по этой формуле дисперсия для разных выборок в среднем занижает генеральную дисперсию.

*Исправленная выборочная дисперсия* вычисляется по формуле:

$$\sigma_\text{в}^2 = \sum_{i=1}^{n} \frac{(x_i - \overline{x})^2}{n-1}$$

В *психологии* точечные оценки могут быть связаны с различными измерениями, например,

- *тесты интеллекта.* С их помощью получают числовое значение IQ и сравнивают его с нормативным значением;
- *тесты личности.* На основе полученных чисел описывают психологические особенности человека;
- *тесты достижений.* С их помощью выясняют, насколько хорошо был усвоен учебный материал;
- *подсчёт количества определённых актов поведения* в ходе наблюдения за испытуемыми;
- *подсчёт площади штриховки* в проективных рисунках;
- *подсчёт количества ошибок* в корректурной пробе.

## Доверительная вероятность и интервальная оценка параметра

*М.А. Воронин, студент группы ПСД 21.212*

*Доверительная вероятность и интервальная оценка параметра* – понятия в математической статистике, связанные с оценкой характеристик генеральной совокупности на основании выборочных данных.

*Доверительная вероятность* (надёжность оценки) – это вероятность того, что истинное (генеральное) значение характеристики попадёт в заданный интервал.

Часто используют доверительную вероятность 0,95.

В некоторых случаях принимают 0,99 или 0,999.

Она связана с величиной, которая называется *уровнем значимости* (α = 1 − доверительная вероятность) – это вероятность того, что истинное значение оцениваемого параметра находится за пределами доверительного интервала. Чем больше доверительная вероятность, тем больше точность оценки доверительного интервала.

*Интервальной оценкой* называется диапазон значений, используемый для оценки параметра генеральной совокупности.

При использовании интервальной оценки наше утверждение относительно истинного значения заключается в том, что оно попадает в некоторый диапазон значений.

Например, утверждения, что средний возраст студентов составляет $19{,}9 < m < 20{,}7$ лет или $m = 20{,}3 \pm 0{,}4$ года, являются интервальными оценками.

Выбор доверительной вероятности для *психологического исследования* зависит от целей и условий проведения.

*Обычно в исследованиях используют уровень доверия 95%.* Он означает, что если повторять аналогичное исследование много раз, независимо друг от друга, то в 95% случаев истинное значение параметра будет попадать в доверительный интервал.

Однако *для небольших выборок*, когда высокая точность не нужна, можно понизить уровень доверительной вероятности до 90% и даже до 85%. Важно учесть это в процессе анализа и в выводах.

Также есть мнение, что *если вероятность повторения результата в генеральной совокупности находится в диапазоне между 95% и 100%*, то можно говорить о статистически достоверном (значимом) результате. То есть исследователю предоставлена возможность совершить ошибку в пределах 5% случаев.

# Доверительный интервал для генеральной средней при известном выборочном среднеквадратическом отклонении

*М.А. Воронин, студент группы ПСД 21.212*

*Доверительный интервал для генеральной средней при известном выборочном среднеквадратическом отклонении* – это диапазон значений, в котором с определённой вероятностью находится истинное значение генеральной средней, оценённое по выборочной средней.

Ширина интервала зависит от объёма выборки, изменчивости данных и выбранного уровня доверия (обычно 95% или 99%). Уровень доверия показывает, как часто при многократном повторении эксперимента истинное значение параметра будет попадать в рассчитанный интервал.

*Формула для расчёта доверительного интервала среднего значения при известном стандартном отклонении генеральной совокупности*: *$\bar{X} \pm Z(\sigma/\sqrt{n})$**,

где: $\bar{X}$ – выборочное среднее;

$Z$ – критическое значение для выбранного уровня доверия (например, 1,96 для 95%);

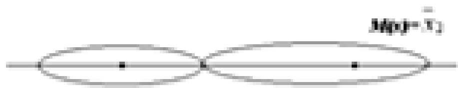
σ – стандартное отклонение генеральной совокупности;

n – объём выборки.

*Генеральная средняя* – среднее арифметическое значений генеральной совокупности $\bar{x}_2 = \frac{1}{N}\sum_{i=1}^{n} x_i$

$\bar{x}_г = (x_1 + x_2 + \cdots + x_N)/N$ – с повторениями.

Генеральная средняя есть среднее взвешенное (СВ) значений генеральной совокупности с их весами, равными соответствующим частотам. Если рассматривать *x* генеральной совокупности как СВ,



$M(X) = \bar{x}_г$

В психологии доверительный интервал для генеральной средней может использоваться в ***психологических исследованиях,*** направленных на выявление среднестатистического результата в ходе психодиагностики и сопутствующих дисциплин.

Данный интервал для генеральной средней *позволяет выявить закономерности*, например, в поведении выбранной группы индивидов и их средние по шкале показатели.

# Доверительный интервал для генеральной средней при неизвестном выборочном среднеквадратическом отклонении

*В.А. Кузык, студентка группы ПСД 21.212*

В *психологических исследованиях* мы часто сталкиваемся с ситуацией, когда необходимо сделать вывод не только о тех людях, которых мы обследовали, но и о более широкой группе – так называемой генеральной совокупности. Например, мы можем изучать уровень тревожности у 30 студентов, но хотим понять, каков этот уровень в среднем у всех студентов города или университета.

Так как мы не можем измерить каждого человека, мы используем данные выборки и стараемся оценить среднее значение для всей совокупности. Однако выборка всегда даёт некоторую погрешность – ведь если бы мы взяли других участников, результаты могли бы немного отличаться. Чтобы показать, в каком диапазоне с определённой уверенностью находится истинное среднее значение, мы строим доверительный интервал.

Если бы мы точно знали, насколько результаты колеблются в популяции (то есть знали стандартное отклонение σ), расчёт был бы проще. Но на практике это почти никогда неизвестно. Поэтому мы используем выборочное стандартное отклонение (s) – показатель того, насколько различаются результаты внутри исследуемой группы.

В этом случае применяется t-распределение Стьюдента, которое позволяет учитывать дополнительную неопределённость, связанную с тем, что мы оцениваем разброс лишь по выборке.

Формула доверительного интервала: $\bar{x} \pm t\alpha/2;\ n-1 \times (s / \sqrt{n}),$
где $\bar{x}$ – среднее значение по выборке, s – выборочное стандартное отклонение, n – количество испытуемых, $t\alpha/2;\ n-1$ – значение из таблицы распределения Стьюдента, зависящее от уровня доверия (обычно 95%) и числа степеней свободы (n–1).

Если мы берём средний результат и добавляем к нему некоторую величину, которая отражает возможную ошибку. Например, если средний уровень тревожности составил 42 балла, а доверительный интервал получился от 38 до 46 баллов, то это означает, что с вероятностью 95% истинное среднее значение находится в этом диапазоне. Таким образом, *доверительный интервал* – это не просто математическая формула. Для психолога это инструмент, который помогает понимать надёжность полученных данных и не делать категоричных выводов по ограниченной выборке. Он показывает, насколько точно мы можем судить о результатах всей группы людей, основываясь на результатах конкретного исследования.

## Понятие статистической гипотезы, виды гипотез

*В.А. Кузык, студентка группы ПСД 21.212*

В психологических исследованиях, как и в других науках, нам часто нужно проверить, есть ли реальные различия между группами людей или влияние какого-то фактора. Например, отличается ли уровень тревожности у мужчин и женщин, или помогает ли новая методика снизить стресс. Чтобы это проверить, используется понятие статистической гипотезы. ***Статистическая гипотеза*** с это предположение о параметрах или распределении данных в генеральной совокупности. То есть это научное предположение, которое можно проверить с помощью статистических методов. Когда мы начинаем анализ, всегда формулируются *две гипотезы*: *Нулевая (основная) гипотеза* ($H_0$) – утверждает, что никаких различий или связей нет, а наблюдаемые результаты – это просто случайность. Пример: «Между уровнем тревожности юношей и девушек нет статистически значимых различий» и *Альтернативная гипотеза* ($H_1$) предполагает, что различия существуют. Пример: «У девушек более высокий уровень тревожности». Другая альтернативная гипотеза: «У девушек более низкий уровень тревожности».

Проверка гипотезы заключается в том, чтобы оценить вероятность получить такие данные, если нулевая гипотеза на самом деле верна. Если вероятность мала (например, меньше 5%), нулевая гипотеза отвергается, и мы принимаем альтернативную.

Существуют ***два основных вида*** статистических гипотез:
*Простая гипотеза* – полностью определяет распределение исследуемой величины, включая все параметры (например, что среднее значение тревожности = 40 баллов), и *Сложная гипотеза* – предполагает, что параметры неизвестны, но находятся в определённых пределах (например, средний уровень тревожности выше 40 баллов, но точное значение неизвестно).

Кроме того, гипотезы делятся ***по направленности***: *Односторонние (направленные)* – когда исследователь предполагает конкретное направление различий (например, «уровень тревожности у девушек выше») и *Двусторонние (ненаправленные)* – когда предполагается лишь факт различия, но не указывается направление (например, «уровень тревожности у юношей и девушек различается»).

Таким образом, *статистическая гипотеза* – это основа любого количественного исследования. Она помогает перейти от наблюдений к научно обоснованным выводам, показывая, насколько полученные результаты случайны, а насколько – действительно отражают реальные различия между людьми или явлениями.

## Ошибки первого и второго рода

*В.А. Кузык, студентка группы ПСД 21.212*

В процессе проверки статистических гипотез психолог сталкивается с тем, что выводы делаются на основе вероятности, а не абсолютной уверенности. Мы никогда не можем знать наверняка, верна ли гипотеза – можем лишь оценить риск ошибки. Именно поэтому в статистике выделяют ошибки первого и второго рода.

*Представим ситуацию*: мы проверяем гипотезу, помогает ли новая программа по развитию эмоционального интеллекта снижать тревожность у подростков.

*Ошибка первого рода (**α**-ошибка).* Это ситуация, когда исследователь отклоняет нулевую гипотезу, хотя на самом деле она верна. Проще говоря, мы делаем ложный вывод о существовании эффекта, которого на самом деле нет. *Пример:* на самом деле новая программа не снижает тревожность, но в исследовании получилось, что эффект есть. Мы «поверили» в результат, которого в действительности нет. Вероятность такой ошибки обозначается греческой буквой **α** (альфа) – обычно её устанавливают на уровне 0,05, то есть допускается риск 5% ошибиться в сторону ложного открытия. *Аналогия:* ложная тревога – мы приняли невинного за виновного.

*Ошибка второго рода (**β**-ошибка).* Возникает, когда исследователь не отвергает нулевую гипотезу, хотя на самом деле она ложна. То есть эффект действительно существует, но мы его не заметили. *Пример:* программа действительно снижает тревожность, но из-за малого объёма выборки или слабого дизайна исследования мы не нашли статистически значимого различия. Вероятность такой ошибки обозначается **β** (бета). Её противоположность – мощность критерия (**1 – β**), то есть способность метода обнаружить реальный эффект, если он есть. *Аналогия:* оправдать виновного – эффект есть, но мы его не увидели.

*Вывод:* Ошибки первого и второго рода – это две стороны одной медали. *Ошибка первого рода* связана с излишней уверенностью, когда мы видим эффект там, где его нет. *Ошибка второго рода* – с чрезмерной осторожностью, когда мы не замечаем реальных различий.

***Для психолога*** важно понимать, что обе ошибки искажают научные выводы. Поэтому при планировании исследований нужно выбирать оптимальный уровень значимости (**α**) и достаточный размер выборки, чтобы снизить вероятность обеих ошибок и сделать результаты максимально надёжными**.**

## Понятие статистического критерия и критической области

*А.А. Калашникова, студентка группы ПСД 21.212*

*Статистический критерий* — это правило, по которому принимается или отвергается нулевая статистическая гипотеза. Он основан на вычислении определённого показателя (статистики) по выборке и сравнении его с критическими значениями.

*Основные характеристики статистического критерия:*

1. Основан на случайной величине.
2. Имеет известную статистическую функцию распределения.
3. Позволяет оценить вероятность получения наблюдаемых результатов.
4. Помогает принять решение о гипотезе.

*Критическая область* – это множество значений статистического критерия, при которых нулевую гипотезу отвергают. Она играет ключевую роль в процессе проверки гипотез.

*Виды критических областей:*

Двусторонняя критическая область – используется при проверке гипотез о равенстве параметра некоторому значению.

Правосторонняя критическая область – применяется при проверке гипотез о превышении параметром некоторого значения.

Левосторонняя критическая область – используется при проверке гипотез о том, что параметр меньше некоторого значения.

*Процесс определения критической области включает:*

- выбор уровня значимости;
- определение типа критической области;
- нахождение критических точек по таблицам распределения;
- вычисление наблюдаемого значения критерия;
- сравнение наблюдаемого значения с критической точкой.

*В психологии* статистические критерии используются для решения различных задач, таких как: выявление различий в уровне исследуемого признака между различными группами; анализ взаимосвязей между психологическими характеристиками; оценка эффективности определённых программ и вмешательств; проверка гипотез о наличии значимых различий.

# Общая схема проверки статистической гипотезы

*А.А. Калашникова, студентка группы ПСД 21.212*

Общая схема проверки статистической гипотезы включает следующие основные этапы:

*1. Формулировка гипотез*

• Нулевая гипотеза ($H_0$): Предполагаемое утверждение, которое требуется проверить.

• Альтернативная гипотеза ($H_1$): Утверждение, противоположное $H_0$.

*2. Выбор уровня значимости (α)*

• Это вероятность ошибки первого рода (ложного отклонения верной гипотезы $H_0$), обычно принимается $\alpha$=0,05 или $\alpha$=0,01.

*3. Выбор статистического критерия*

• Критерий – это правило, по которому принимается или отвергается $H_0$. Он основан на выборочной статистике.

*4. Вычисление наблюдаемого значения критерия*

• Подсчёт значения критерия на основе выборочных данных.

*5. Определение критической области*

• Это область значений критерия, при попадании в которую $H_0$ отвергается. Она зависит от уровня значимости $\alpha$ и вида альтернативной гипотезы.

*6. Принятие решения*

• Если наблюдаемое значение критерия попадает в критическую область – $H_0$ отвергается в пользу $H_1$.

• Если не попадает – $H_0$ не отвергается (но не обязательно принимается как истинная).

*7. Интерпретация результата*

• Формулировка выводов на основе принятого решения с учётом уровня значимости и возможных ошибок.

*В **психологических исследованиях*** проверка гипотез позволяет: объективно оценивать эффективность методик; проверять теоретические предположения; устанавливать статистическую значимость полученных результатов и избегать субъективных интерпретаций данных.

# Решение задачи по проверке статистической гипотезы о генеральной средней при известном среднеквадратическом отклонении

*А.А. Калашникова, студентка группы ПСД 21.212*

*Шаг 1: Формулировка гипотез*

1. Нулевая гипотеза ($H_0$): $\mu=\mu_0$ (генеральная средняя равна некоторому заданному значению $\mu_0$).

2. Альтернативная гипотеза ($H_1$): двусторонний тест: $\mu\neq\mu_0$. Односторонний тест: $\mu>\mu_0$.

*Шаг 2: Выбор уровня значимости (α)*

Обычно $\alpha$=0,05 (5%), но может быть другим в зависимости от требований задачи.

*Шаг 3: Расчёт тестовой статистики*

Используем Z-критерий: $Z = \frac{\bar{X}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$, где:

$\bar{X}$ – выборочная средняя, $n$ – размер выборки.

*Шаг 4: Определение критической области*

1. Двусторонний тест.

2. Односторонний тест ($\mu>\mu_0$).

*Шаг 5: Принятие решения*

Если тестовая статистика попадает в критическую область – отклоняем $H_0$. Иначе – не отклоняем $H_0$.

***В психологии*** *статистические гипотезы* используются для проверки научных предположений о свойствах психики, условиях её формирования и развития. Например:

- проверка эффективности новых методов психотерапии;
- исследование влияния факторов среды на когнитивные способности;
- анализ различий между группами испытуемых и др.

*Ошибки первого и второго рода*

- ошибка первого рода (ложное отклонение верной гипотезы) может привести к отказу от потенциально эффективного метода лечения.
- ошибка второго рода (ложное принятие неверной гипотезы) опасна тем, что неэффективный метод может быть внедрён в практику.

*Статистические критерии*

- параметрические применяются для сравнения средних значений в экспериментальных и контрольных группах.
- непараметрические критерии (используются для анализа категориальных данных, таких как распределение ответов на тесты.

# Проверка статистической гипотезы о генеральной средней при неизвестном среднеквадратическом отклонении

*У.Д. Молинова, студентка группы ПСД 21.212*

*Статистическая гипотеза* – это некоторое предположение о свойствах и характеристиках исследуемых генеральных совокупностей. *Нулевая гипотеза* $H_0$: $\mu = \mu_0$ – это гипотеза, которой мы придерживаемся, пока наблюдения не заставят признать обратное, где $\mu_0$ – заданное значение средней. *Альтернативная гипотеза* $H_1$: $\mu \neq \mu_0$, (двусторонняя), или $\mu > \mu_0$ / $\mu < \mu_0$ (односторонняя), в зависимости от исследовательской задачи.

Для проверки статистической гипотезы о генеральной средней при неизвестном среднеквадратическом отклонении используется t-тест для одного среднего. Поскольку дисперсия (или стандартное отклонение) генеральной совокупности неизвестна, в статистику теста вместо неизвестного σ подставляется выборочное среднеквадратическое отклонение s. Тогда критерий проверки нулевой гипотезы строится на основе t-статистики: $t = \frac{\bar{X}-\mu_0}{s/\sqrt{n}}$, где $\bar{X}$ – выборочное среднее, $s$ – выборочное стандартное отклонение, $n$ – объём выборки. Распределение этой статистики при верности гипотезы $H_0$ и нормальности генеральной совокупности соответствует t-распределению Стьюдента с $n - 1$ степенью свободы. Далее по значению расчётного t и критическим значениям из таблиц t-распределения принимается решение о принятии или отклонении нулевой гипотезы. Этот тест надёжен при нормальном распределении совокупности и полезен для малых выборок и неизвестной дисперсии.

*Пример использования в психологии:* предположим, психолог хочет проверить, изменилась ли средняя степень тревожности у группы испытуемых после прохождения психологической тренировки.

*Гипотеза:* $H_0$: $\mu = 50$ (средняя тревожность до тренировки равна 50 баллам). Допустим, после тренировки была выбрана выборка из $n = 10$ человек, где: выборочное среднее $\bar{X} = 45$; выборочное стандартное отклонение $s = 8$. Вычислим t-статистику: $t = \frac{45-50}{8/\sqrt{10}} \approx -1{,}98$. Степени свободы: $n - 1$. По таблице t-распределения для уровня значимости $\alpha = 0{,}05$ двусторонний критический критерий равен 2,2622. Поскольку рассчитанное значение $t = -1{,}98$ по модулю меньше критического 2,2622 гипотеза $H_0$ не отвергается на уровне 5%.

Вывод: статистически значимых оснований полагать, что средняя тревожность после тренировки отличается от 50 баллов, нет.

# Проверка статистической гипотезы о генеральной дисперсии

*У.Д. Молинова, студентка группы ПСД 21.212*

*Проверка статистической гипотезы о генеральной дисперсии* обычно проводится с помощью критерия хи-квадрат, если генеральная совокупность нормально распределена.

Основные шаги проверки гипотезы:

1. Формулируется нулевая гипотеза $H_0: \sigma^2 = \sigma_0^2$ , где $\sigma_0^2$ – утверждаемое значение дисперсии генеральной совокупности.

2. Вычисляется выборочная дисперсия $S^2$ по имеющейся выборке.

3. Статистика критерия $\mathcal{X}^2 = (n-1) \cdot S^2/\sigma_0^2$, где $n$ – объём выборки.

4. Статистика $\mathcal{X}^2$, с $n-1$ степенями свободы.

5. Определяется критическая область в зависимости от типа альтернативной гипотезы (двусторонняя, лево- или правосторонняя).

6. На основании сравнения вычисленного $\mathcal{X}^2$ с критическим значением принимается решение: если $\mathcal{X}^2$ *не попадает* в критическую область, гипотеза $H_0$ не отвергается; если $\mathcal{X}^2$ *попадает* в критическую область, гипотеза $H_0$ отвергается.

*Пример расчета проверки гипотезы о генеральной дисперсии в психологии.* Допустим, психолог исследует уровень тревожности группы из 10 человек по некоторой шкале. Известно, что в популяции (генеральной совокупности) дисперсия уровня тревожности равна $\sigma_0^2 = 16$. Проводится выборка, и получены значения тревожности: 20, 18, 22, 17, 19, 21, 23, 20, 18, 22. Основные шаги проверки гипотезы о значении генеральной дисперсии:

1. Формулируем гипотезы: $H_0: \sigma^2 = 16, \quad H_1: \sigma^2 \neq 16$

2. Рассчитаем выборочную дисперсию $S^2$. Находим среднее выборки $x^2 = \frac{20+18+22+17+19+21+23+20+18+22}{10}$ =19,8. Вычисляем квадраты отклонений каждого значения от среднего и суммируем: (20-19.8)$^2$ + (18-19.8) $^2$+ … + (22-19.8) $^2$= 39,6. Выборочная дисперсия: $S^2 = 39{,}6/(10-1) = 4{,}4$.

3. Рассчитываем статистику критерия хи-квадрат:

$$\mathcal{X}^2 = 9 \times 4.4/16 = 2{,}475.$$

4. Определяем критические значения для уровня значимости $\alpha = 0{,}05$ и степеней свободы $n-1 = 9$. Нижняя критическая граница: $\mathcal{X}^2$=2,7. Верхняя критическая граница: $\mathcal{X}^2$= 19,0.

5. Сравниваем: 2.475 < 2.7 – не входит в область принятия $H_0$.

Вывод: на уровне значимости 0.05 нет оснований отвергать гипотезу о том, что дисперсия генеральной совокупности равна 16, т.е. изменчивость не является статистически значимо иной от ожидаемой.

## Понятие критериев согласия

*А.А. Мухамутдинова, студентка группы ПСД 21.212*

*Критерии согласия* предназначены для проверки гипотез о соответствии эмпирического распределения, построенного по выборке, извлекаемой из генеральной совокупности, некоторому теоретическому закону.

Такие критерии разделяются на два класса:

1. *Общие критерии согласия*. Применяются к общей формулировке гипотезы, а именно к гипотезе о согласии наблюдаемых результатов с любым априорно предполагаемым распределением вероятностей.
2. *Специальные критерии согласия*. Предполагают специальные нулевые гипотезы, формулирующие согласие с определённой формой распределения вероятностей.

Критерии согласия ***в психологии*** обычно относятся к статистическим методам, которые используются для проверки соответствия эмпирических данных теоретическим моделям или гипотезам. Примером критерия согласия является критерий хи-квадрат Пирсона, который проверяет, насколько наблюдаемые распределения данных соответствуют ожидаемым распределениям по гипотезе. В психологическом эксперименте это может проявляться, например, в проверке того, что выборы испытуемых распределены равномерно между условиями эксперимента или что ответная частота на тест соответствует нормальному распределению. Критерии согласия принимают или отвергают гипотезу, анализируя величину отклонения эмпирических данных от теории с использованием определённых статистических функций отклонения.

***Пример в психологии***: в эксперименте испытуемому предлагают выбрать задания с двух разных столов, между которыми гипотетически нет разницы. Если из 150 человек 98 выбрали стол справа, а 52 – слева, с помощью критерия хи-квадрат оценивается, является ли такое распределение случайным или показывает предпочтение. Если вероятность получения таких данных по случайности мала, гипотеза о равновероятности выбора отвергается, и делается вывод о наличии предпочтения.

## Критерий Пирсона проверки статистической гипотезы о нормальном распределении признака X

*А.А. Мухамутдинова, студентка группы ПСД 21.212*

*Суть метода*: сравнение эмпирических частот с теоретическими частотами, которые должны наблюдаться при нормальном распределении.

Существуют гипотезы, где: $H_0$: Признак X распределён нормально; $H_1$: Признак X распределён не нормально.

Как проверить гипотезу?

Для начала устанавливается допустимая вероятность ошибки первого рода. Далее по выборке вычисляется выборочное среднее математическое ожидание и выборочная дисперсия, которые являются оценками параметров нормального распределения. Выборочные данные группируются в интервалы, и для каждого интервала подсчитывается его наблюдаемая частота (число элементов выборки, которые попали в данный интервал). Следом, для каждого интервала определяется теоретическая вероятность попадания в этот интервал. Затем эти вероятности умножаются на общий объем выборки, чтобы получить теоретические частоты. Вычисляется статистика, которая отражает расхождение между наблюдаемыми ($n_i$) и ожидаемыми ($n_i'$) частотами по формуле: $\chi^2 = \Sigma\ [(n_i - n_i')^2 / n_i']$

После расчёта определяются чи́сла степеней свободы. Число степеней свободы (v) рассчитывается как $v = k-1-r$, где k – число интервалов, а r – число оцененных параметров (для нормального распределения $r=2$). Сравниваются статистики с критическим значением. Полученная статистика $\chi^2$ сравнивается с критическим значением $\chi^2\ (\alpha, \nu)$ из таблицы распределения хи-квадрат при заданном уровне значимости $\alpha$ и вычисленных степенях свободы.

Если $\chi^2 < \chi^2(\alpha, \nu)$, то нулевая гипотеза $H_0$ принимается, т.е. можно считать, что признак X имеет нормальное распределение.

Если $\chi^2 \geq \chi^2(\alpha, \nu)$, то нулевая гипотеза $H_0$ отвергается в пользу альтернативной, и делается вывод, что эмпирическое распределение не соответствует нормальному.

***Пример***: В психологическом исследовании измеряется уровень тревожности у 200 испытуемых. Данные разбиваются на 7 интервалов уровней. Вычислены теоретические частоты по нормальному распределению с параметрами выборки. Рассчитано значение $\chi 2 = 10.5$, критическое $\chi 2$ при 4 степенях свободы и уровне 0.05 равно 9.49. Поскольку $10.5 > 9.49$, гипотеза о нормальном распределении отвергается на данном уровне значимости.

## Нахождение теоретических частот по интервальному вариационному ряду в предложении, что признак X распределён нормально

*А.А. Мухамутдинова, студентка группы ПСД 21.212*

*Суть метода*: предположим, наша выборочная совокупность данных (интервальный ряд) взята из генеральной совокупности, имеющей нормальное распределение. Нам необходимо вычислить, какими были бы частоты (количество наблюдений) в каждом интервале, если бы данные действительно точно подчинялись нормальному закону с параметрами, оцененными по выборке.

Чтобы найти *теоретические частоты* для интервального вариационного ряда при нормальном распределении, необходимо:

1. Рассчитать среднее значение (M) и стандартное отклонение (σ) по наблюдаемым данным выборки. Эти два показателя будут характеризовать нормальное распределение.
2. Выбрать значение интервала, соответствующее центральной точке (x_i). Например, для интервала от 30 до 40 середина будет (30+40)/2=35.
3. Используя формулу, рассчитать Z-значение для каждого интервала: Z = (x_i – M) / σ. Z-значение показывает, насколько стандартных отклонений середина интервала отстоит от среднего значения.
4. Найти вероятность попадания в каждый интервал, используя Z- значения (по таблице нормального распределения или с помощью статического программного обеспечения). Вероятность попадания в интервал находится как разность значений из таблицы для верхней и нижней границы интервала.
5. Рассчитать теоретическую частоту (f_i) по формуле: f_i = P(интервал) * n, где P(интервал) – вероятность попадания в данный интервал, n – общее число наблюдений в выборке.

***В психологии*** применяется при валидизации методик, в подборе статистических методов, при нормировании тестов. Это позволяет сравнить эмпирические частоты с теоретическими частотами нормального распределения, проверить гипотезу о нормальности и использовать статистические критерии для оценки соответствия.

Такой подход типичен для психологических исследований, где признаки (например, тестовые баллы) часто предполагаются распределёнными нормально, и это применяется для построения гистограмм и оценки распределения данных по интервалам.

## Методы математической статистики при обработке результатов ретестовых исследований

*О.Г. Юренков, кандидат социологических наук*

Ретестовые исследования являются важным инструментом в психологии, позволяющим оценивать стабильность и надёжность измеряемых характеристик. Для корректной интерпретации результатов исследований необходима качественная обработка данных, что в свою очередь требует применения методов математической статистики.

Ретестовые исследования (или повторные тестирования) предполагают многократное измерение одной и той же переменной у одних и тех же испытуемых в разные временные моменты. Это позволяет исследователям оценить надёжность и валидность используемых тестов, а также выявить изменения в психических состояниях или характеристиках испытуемых.

Основные задачи, которые решаются с помощью математической статистики в контексте ретестовых исследований:

- *оценка надёжности*: Ключевым аспектом является проверка стабильности результатов. Для этого часто применяются коэффициенты надёжности, такие как коэффициент альфа Кронбаха, тесты на согласованность (например, корреляция Пирсона или Спирмена);

- *сравнение групп:* В ретестовых исследованиях может возникнуть необходимость в сравнении результатов между различными группами (например, до и после вмешательства). Используются методы, позволяющие оценить статистическую значимость различий;

- *анализ изменений:* важно не только установить наличие различий, но и оценить саму величину изменений.

При обработке данных ретестовых исследований можно использовать различные статистические методы:

- *корреляционный анализ:* позволяет оценить степень взаимосвязи между результатами различных тестирований. Корреляция выше 0,7 обычно считается высокой и указывает на хорошую надёжность.

- *регрессионный анализ:* может быть использован для предсказания значений одного показателя на основе другого. Это полезно, например, когда нужно оценить, как изменение одного из параметров влияет на результаты тестирования.

- *модели роста:* если ретестовые данные собираются в течение длительного времени, модели роста могут помочь понять динамику изменений в измеряемых переменных.

Правильный выбор и применение методов способствуют более глубокому пониманию сущности и динамики психических процессов.

# Математическая статистика в современной психологии

*Т.Н. Антошина, доцент кафедры прикладной математики*
*и безопасности информационных технологий,*
*кандидат педагогических наук*

Математическая статистика в современной психологии – это не просто инструмент для «подтверждения гипотез», а способ мышления и язык строгого научного диалога.

Дисциплина «Математика и математическая статистика» является фундаментом эмпирической психологии, превращая её из области умозрительных рассуждений в науку, основанную на данных.

Целью изучения дисциплины является формирование целостного представления о математических методах, применяемых в психологии.

Содержание сборника направлено на изучение методов статистики, их роли в решении психологических задач на любом предприятии, приобретение знаний и развитие компетенций, необходимых для работы с информационными процессами. Структура сборника сформирована с учётом всех аспектов и тенденций развития.

Несмотря на научный характер, данный сборник изложен очень доступным языком; он не требует высокого уровня предварительных знаний, и будет интересен не только специалистам, но и обучающимся, которые только начинают свой путь в освоении математической статистики.

Некоторые статьи носят компилятивный характер, а некоторые – вполне творческий, оригинальный.

## Содержание

**Роль математической статистики в психологии**
*Кабанов А.А.* 3

**Математическая статистика и её основные задачи**
*Феофанова М.Ю.* 4

**Генеральная и выборочная совокупности, их объёмы. Суть выборочного метода**
*Феофанова М.Ю.* 5

**Понятие закона больших чисел**
*Феофанова М.Ю.* 6

**Теорема Чебышёва и её следствия**
*Михайлова Е.А.* 7

**Центральная предельная теорема Ляпунова**
*Михайлова Е.А.* 8

**Основные способы отбора элементов в выборку**
*Михайлова Е.А.* 9

**Понятие и виды вариационных рядов распределения частот и частостей**
*Бабенко В.А.* 10

**Построение интервальных рядов**
*Бабенко В.А.* 11

**Эмпирическая функция распределения и её связь с теоретической функцией распределения**
*Бабенко В.А.* 12

**Понятие накопленной частоты и кумуляты распределения частот дискретного и интервального рядов распределения**
*Чванова В.В.* 13

**Полигон распределения частот дискретного
и интервального рядов распределения, его построение**
*Чванова В.В.* 14

**Гистограмма распределения частот интервального ряда**
*Кручай М.И.*
15

**Средняя выборочная, её нахождение и сравнение
с математическим ожиданием**
*Пестова Д.А.* 16

**Понятие моды и её нахождение для дискретного
и интервального рядов**
*Пестова Д.А.* 17

**Понятие медианы и её нахождение для дискретного
и интервального рядов**
*Пестова Д.А.* 18

**Выборочная дисперсия и выборочное
среднеквадратическое отклонение**
*Кручай М.И.* 19

**Коэффициент асимметрии вариационного ряда и его
смысловое значение**
*Бычкова А.В.* 20

**Коэффициент эксцесса вариационного ряда и его
смысловое значение**
*Бычкова А.В.* 21

**Понятие точечной оценки и её интерпретация как
случайной величины**
*Метушевская Н.В.* 22

**Основные свойства точечных оценок**
*Метушевская Н.В.* 23

**Точечная оценка для генеральной средней**
*Метушевская Н.В.* 24

**Точечная оценка для генеральной дисперсии**
*Воронин М.А.* 25

**Доверительная вероятность и интервальная оценка параметра**
*Воронин М.А.* 26

**Доверительный интервал для генеральной средней при известном выборочном среднеквадратическом отклонении**
*Воронин М.А.* 27

**Доверительный интервал для генеральной средней при неизвестном выборочном среднеквадратическом отклонении**
*Кузык В.А.* 28

**Понятие статистической гипотезы, виды гипотез**
*Кузык В.А.* 29

**Ошибки первого и второго рода**
*Кузык В.А.* 30

**Понятие статистического критерия и критической области**
*Калашникова А.А.* 31

**Общая схема проверки статистической гипотезы**
*Калашникова А.А.* 32

**Решение задачи по проверке статистической гипотезы о генеральной средней при известном среднеквадратическом отклонении**
*Калашникова А.А.* 33

**Проверка статистической гипотезы о генеральной средней при неизвестном среднеквадратическом отклонении**
*Молинова У.Д.* 34

**Проверка статистической гипотезы о генеральной дисперсии**
*Молинова У.Д.* 35

**Понятие критериев согласия**
*Мухамутдинова А.А.* 36

**Критерий Пирсона проверки статистической гипотезы о нормальном распределении признака X**
*А.А. Мухамутдинова* 37

**Нахождение теоретических частот по интервальному вариационному ряду в предложении, что признак X распределён нормально**
*А.А. Мухамутдинова* 38

**Методы математической статистики при обработке результатов ретестовых исследований**
*Юренков О.Г.* 39

**Математическая статистика в современной психологии**
*Антошина Т.Н.* 39

**Авторский коллектив:**

*Антошина Татьяна Николаевна, Бабенко Вероника Алексеевна, Бычкова Анастасия Валерьевна, Воронин Максим Андреевич, Джафарова Анастасия Алексеевна, Каба́нов Андрей Александрович, Калашникова Анастасия Алексеевна, Кручай Мария Ильинична, Кузык Вероника Андреевна, Лабинский Александр Юрьевич, Матвеев Александр Владимирович, Матвеева Юлия Геннадьевна, Метушевская Нина Васильевна, Михайлова Екатерина Алексеевна, Молинова Ульяна Дмитриевна, Мухамутдинова Азалия Азатовна. Пестова Диана Андреевна, Феофанова Маргарита Юрьевна, Чванова Валерия Васильевна*

*Для заметок*

*Для заметок*

Составление, вступительная статья и компьютерная вёрстка:
**КабановАндрей Александрович**,
кандидат юридических наук, доцент,
e-mail:***akabanov@inbox.ru***

сайт: ***otvet-akab.ru***

# Математическая статистика в психологии

## Сборник статей

ISBN 978-5-6055849-3-3

Редакционная коллегия:
Т.Н. Антошина, А.А. Джафарова, А.А. Кабанов, А.В. Матвеев, Ю.Г. Матвеева
Компьютерная верстка: А.А. Кабанов
Печатается в авторской редакции
Не для продажи

Подписано в печать и свет 20.02.2026. Формат 60×84 1/16
Печать офсетная. Объём 3 п.л. Тираж 100 экз.

Отпечатано в типографии ООО «Р-КОПИ»
Санкт-Петербург, вн. тер. г. муниципальный округ Коломна,
Пер. Дровяной, д. 5, литера А, помещ. 1-Н