Министерство сельского хозяйства Российской Федерации
Иркутский государственный аграрный университет
имени А.А. Ежевского
Институт управления природными ресурсами – факультет охотоведения
имени В.Н. Скалона

Г.В. Чудновская
В.О. Саловаров
А.П. Демидович

БИОМЕТРИЯ В ИХТИОЛОГИИ

УЧЕБНОЕ ПОСОБИЕ

Иркутск, 2018

УДК 574/577

Печатается по решению научно-методического совета Иркутского ГАУ им. А.А. Ежевского (протокол № 4 от 30.04.2018 г.).

Составители:

Г.В. Чудновская, заведующая кафедрой «Технологии в охотничьем и лесном хозяйстве», доцент, к.б.н.;

В.О. Саловаров, директор института Управления природными ресурсами – факультет охотоведения имени В.Н. Скалона, профессор, д.б.н.;

А.П. Демидович, заведующий кафедрой «Общей биологии и экологии», доцент, к.б.н.

Рецензенты:

Д.Ф. Леонтьев, профессор кафедры «Технологии в охотничьем и лесном хозяйстве», д.б.н.

Чудновская Г.В., Саловаров В.О., Демидович А.П. Биометрия в ихтиологии: Учебное пособие. – Иркутск: Издательство Иркутского ГАУ им. А.А. Ежевского, 2018. – 156 с.

В учебном пособии для студентов последовательно излагается порядок изучения дисциплины «Биометрия». Издание предназначено для обучающихся направления 35.03.08 - Водные биоресурсы и аквакультура.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ

Биометрия представляет собой своеобразный инструмент, способный выразить в числе и измерить значимость и надежность полученных результатов, заранее рассчитать и спланировать необходимую численность объектов для того или иного эксперимента, оценить достоверность проверяемой гипотезы.

Биометрия является разделом высшей математики – вариационной статистики, разрабатывающей методы изучения варьирующего признака, то есть проявляющего определенную закономерность в изменчивости своих значений на массовых материалах, и фиксирует совокупность математических методов из области статистики и теории вероятностей применительно к биологическим объектам. Однако отождествлять биометрию с математической статистикой и теорией вероятностей нельзя. Она имеет свою специфику, отличительные черты и занимает определенное место в системе биологических наук. Современная биометрия является инструментом планирования наблюдений и статистической обработки их результатов.

Для организации рационального промысла рыбных запасов и осуществления контроля и регулирования их состояния требуется оценка и прогнозирование возможных уловов. Поэтому исследования наиболее характерных признаков различных групп рыб требуют проведения статистического анализа полученных данных. В ходе ихтиологических исследований получают качественные и количественные характеристики, которые требуют объективного понимания и научного объяснения, для этого находят широкое применение разнообразные методы и приемы биометрии.

Игнорирование статистической обработки и математического анализа полученного материала может свести на нет результаты многих важных опытов, привести к необоснованным и даже ошибочным заключениям. Напротив, умелое использование биометрических методов увеличивает информационную ценность проведенного исследования, помогает правильно планировать их постановку, глубоко разобраться в полученных данных, объективно судить о результатах наблюдений, выявить скрытые закономерности и правильно их трактовать, что в конечном итоге делает биологию точной наукой. При этом следует иметь в виду, что сама по себе статистическая обработка данных, как бы ни была она совершенна с точки зрения математики, не может служить гарантией качественности выполнения биологом исследования и не способна обеспечить надежность полученных им результатов, если само исследование проведено неправильно или использованные данные ошибочны.

Применение математико-статистических методов в биологии, в том числе и прикладной, базируется на выборе определенной статистической модели и проверки ее соответствия экспериментальным данным, а также анализа полученных результатов, вытекающих из ее рассмотрения. Планирование эксперимента является самостоятельным разделом биометрии, описывающим методы эффективной постановки опыта (дисперсионный и последовательный анализы, планирование отсеивающих экспериментов и т.д.). Эти методы позволяют резко сократить объем эксперимента для получения того же количества информации.

ИСТОРИЯ БИОМЕТРИИ

В биологии использование математических приемов началось значительно позже, чем в других естественных науках, долгое время она развивались на основе только качественного понимания явлений. Необходимость количественной оценки стала осознаваться только в конце XIX века.

Бельгийский математик астроном, метеоролог и социолог Адольф Кетле (1796-1874 гг.) был первым, кто применил статистические методы в анализе биологических данных. В 1835 году в работе «Социальная физика, или опыт исследования о развитии человеческих способностей» он показал, что величина роста измеренных 10 тысяч человек, приблизительно приближается к нормальному распределению. Использовал выражение «средний человек», чтобы отразить тот факт, что большинство результатов физических измерений группируются вокруг их среднего значения или центра распределения, а количество остальных данных уменьшается по мере их отклонения от этой величины.

Из работ А. Кетле следовало, что задача статистики заключается не только в сборе и классификации данных, а в их анализе, целью выявления закономерностей, действующих в пределах многочисленных массовых случаев. Знание их и должно было превратить статистику в источник научного познания социальных и биологических явлений.

Исследования А. Кетле явились поворотным пунктом в истории статистической науки. Он убедительно показал, что случайности, наблюдаемые в живой природе и их периодичность, показывают направление, которое можно исследовать и описать точными математическими методами.

У истоков биометрии, как науки, стоял Фрэнсис Гальтон (1822-1911 гг.). Первоначально он готовился стать врачом, однако обучаясь в Кембриджском университете, увлекся естествознанием, метеорологией, антропологией, теориями наследственности и эволюции. В книге «Естественная наследственность» (1889 г.) Ф. Гальтон впервые ввел в употребление термин «biometry».

Затем в 1899 году Герман Дункер (Hermann Ludwig Rudolph Duncker (1874–1960 гг.)) предложил другой термин – «вариационная статистика», который отображал совокупность математико-статистических методов в биологических исследованиях. Так появились разные названия одного и того же предмета, хотя буквальный смысл их не одинаков. Понятие «биометрия» (от греч. «bios» – жизнь и «metron» – мера) означает измерение биологических объектов, а термин «вариационная статистика» (от лат. «variatio» – измерение, колебание и «status» – состояние, положение вещей), понимают, как статистическую обработку результатов измерений. Оба термина недостаточно точны. Поэтому А.В. Леонтович в книге «Применение метода Gaussa к оценке ошибок»(1909 г.), и много позже П.Ф. Рокицкий в учебнике «Биологическая статистика» (1973 г.), предлагали заменить их термином «биологическая статистика», который не прижился так как также не лишен недостатков. Ю.Л. Поморский в работе «Методы биометрических исследований» (1935 г.) пришел к выводу, что из всех предложенных терминов наиболее удачным следует считать термин «биометрия», как наиболее четко отражающий содержание этого предмета.

Работы Ф. Гальтона в области статистики привели к открытию одной из самых важных величин - корреляции, первое упоминание о которой появилось в 1888 году. Современные приемы определения обоснованности и надежности тестов так же, как и методы факторного анализа, напрямую связаны с этим открытием, которое стало результатом наблюдений за тем, как количественные характеристики наследственных признаков регрессируют (движутся) к своему среднему значению. К примеру, он отмечал, что сыновья очень высоких людей, в среднем, бывают ниже своих отцов, в то время как сыновья низкорослых мужчин оказываются, в среднем, выше своих отцов. Ф. Гальтон разработал графические способы для отражения основных свойств коэффициента корреляции и нашел формулу для его расчета.

Таким образом, Ф. Гальтон заложил основы новой науки и дал ей имя, однако в стройную научную дисциплину ее превратил его студент, математик Карл Пирсон (1857-1936 гг.). Он создал математический аппарат биометрии, развил учение о разных типах кривых распределения, разработал критерий «хи квадрат», ввёл в биометрию такие показатели, как стандартное отклонение, коэффициент вариации. При поддержке Ф. Гальтона, К. Пирсон вывел используемую и по сей день формулу определения коэффициента корреляции, получившего название «коэффициент корреляции Пирсона». В 1903 г. К. Пирсон создал основы теории сопряженности признаков, а в 1905 г. опубликовал основы нелинейного корреляционного анализа и метода нелинейной регрессии.

Следует отметить, что исследования Ф. Гальтона и К. Пирсона поначалу не получили признания у научной общественности, а их статьи отказывались печатать в ведущих научных изданиях. Поэтому в 1901 г. К. Пирсон организовал выпуск собственного журнала «Biometrika», который существует до сих пор и считается наиболее авторитетным изданием в своей области.

Одной из причин недоверия к первым биометрическим работам было то, что она делала упор на многочисленных рядах данных и фактически не занималась анализом «малых выборок». Поэтому большинство исследователей не могли использовать ее методики на практике. Данную проблему разрешил англичанин Уильям Сили Госсет (1876-1937 гг.), обосновав теорию малой выборки и представление о том, что даже для небольшого количества данных можно успешно использовать статистические приемы. У. Госсет в 1908 г. под псевдонимом Student (Стьюдент) опубликовал работу «Вероятная ошибка средней», где описал разработанный им t-критерий.

Первым, кто понял значение работ У. Госсета по оценке параметров малой выборки, и дал им дальнейшее развитие, был биолог и выдающийся английский статистик Рональд Фишер(1890-1962 гг.). Его научные работы по праву можно считать вершиной классической и фундаментом современной биометрии. Он является основателем дисперсионного анализа, статистической теории планирования экспериментов, метода максимального правдоподобия и многого другого, что составляет основу современной прикладной статистики.

Рассматривая историю биометрии, нельзя не отметить тот огромный вклад, который внесли в ее развитие российские ученые. Статистические методы специально для задач биологии в начале XX века начал развивать Александр Александрович Чупров (1874-1926 гг.), работавший в то время в Петербургском политехническом институте и уделявший особое внимание работам английских биометриков.

Представители естественных наук стали осваивать математические методы. Во многом этому способствовало издание работы Александра Васильевича Леонтовича (1869-1943 гг.) – советского физиолога и нейрогистолога «Элементарное пособие к применению методов Гаусса и Пирсона при оценке ошибок в статистике и биологии» (1909-1911гг.), первого русского научного издания по применению математики в биологии и медицине.

В последствие им были опубликованы учебники для студентов сельскохозяйственных вузов (1922 г.).

Хочется упомянуть наиболее выдающихся ученых советской и российской науки, способствовавших развитию и внедрению математических методов и биологию.

Труды Сергея Натановича Бернштейна (1880-1968 гг.) относятся к дифференциальным уравнениям в частных производных, теории вероятностей и конструктивной теории функций, создателем которой он является.

С именем Александра Яковлевича Хинчина (1894-1959 гг.) связаны фундаментальные достижения в области теории функций действительного переменного, теории чисел и теории вероятностей.

Евгений Евгеньевич Слуцкий (1880-1948 гг.) – один из создателей современной теории случайных функций (распределений в функциональных пространствах). Также он работал по оценке параметров корреляции и над составлением таблиц функций от нескольких переменных.

Борис Сергеевич Ястремский (1877-1962 гг.) – автор научных работ теоретического и прикладного характера в области математической, сельскохозяйственной статистики и демографии.

Андрей Николаевич Колмогорова (1903-1987 гг.) – величайший русский математик, создатель современной теории вероятностей, автор классических результатов в теории функций, в математической логике, топологии, теории дифференциальных уравнений, функциональном анализе, в теории турбулентности, теории гамильтоновых систем. Созданные им школы определили развитие этих направлений математики в XX столетии, получили мировое признание.

Сергей Сергеевич Четвериков (1880-1959 гг.) – биолог, генетик-эволюционист, сделавший первые шаги в направлении синтеза менделевской генетики и эволюционной теории Чарльза Дарвина. С.С. Четвериков, пользуясь несложными математическими методами, доказал, что мутации в природных популяциях животных не исчезают, могут накапливаться в скрытом состоянии и давать материал для изменчивости и естественного отбора.

Николай Александрович Плохинский (1899-1988 гг.) – биолог. Отечественная биология именно ему обязана возрождением такой области научного знания, как биологическая статистика. Николай Александрович первым из отечественных ученых начал приучать научных сотрудников к корректному представлению экспериментальных данных и заложил основы биометрической грамотности биологов в нашей стране. Первая его книга «Статистические методы в зоотехнии» была издана в 1937 году.

Юрий Александрович Филипченко (1882-1930 гг.) – биолог и генетик. Большое значение он придавал исследованию методов математической статистики в генетических исследованиях и одним из первых в СССР стал применять их в своих работах. На основе измерений и статистической обработки выявил признаки, характеризующиеся средними индексами краниологических особенностей, и показал, что они наследуются по законам Менделя. Ю.А. Филипченко не только умело применял биометрию в исследовательской работе, но и пропагандировал ее.

Павел Викторович Терентьев (1903-1970 гг.) – зоолог, биометрист. Внес большой вклад в развитие современной биометрии. В 1936 году защитил кандидатскую диссертацию на тему «Применение биометрических методов к изучению земноводных». Его с полным правом можно назвать новатором в области теоретической биогеографии, особенно в рамках применения в этой науки методов математической статистики и теории вероятностей.

Александр Александрович Любищев (1890-1972 гг.) - биолог и энтомолог. Известен благодаря своим работам по применению математических методов в биологии, среди них классические труды по дисперсионному анализу.

Владислав Иванович Василевич – биолог, в 1969 году опубликовал книгу «Статистические методы в геоботанике». В работе представлена обстоятельная трактовка математических методов применительно к геоботанике, дан детальный анализ всей имевшейся на тот момент мировой литературы в данной области, а также предложен ряд оригинальных разработок.

Лев Анатольевич Животовский – биолог, генетик, автор научных работ в области популяционной биометрии, математического моделирования селекционно-генетических процессов в популяциях сельскохозяйственных животных.

Возрастающая роль биометрии в исследовательской работе естественно сказалась на подготовке специалистов биологического профиля. Первый учебник по теории вероятностей «Основания математической теории вероятностей» был издан в России в 1846 году Виктором Яковлевичем

Буняковским (1804-1889 гг.). В этом обширном трактате, кроме теории, были даны объяснения относительно новых решений самых трудных и запутанных вопросов, указано много практических приложений, описана история возникновения и развития теории вероятностей.

Первым, кто еще в 1919 г. начал читать студентам Московского университета курс биометрии с основами генетики, был С.С. Четвериков. В 1924 г. он вел уже самостоятельный курс «Введение в биометрию» для студентов, специализирующихся на кафедре генетики. Его продолжателем по праву можно считать Н.А. Плохинского, который за годы своей научно-педагогической деятельности в МГУ, опубликовал несколько учебников по биометрии, в том числе «Биометрия», «Алгоритмы биометрии», «Дисперсионный анализ».

Также большой вклад в организацию преподавания основ статистического анализа данных биологических исследований, пропаганду необходимости статистической обработки научных материалов внес Петр Фомич Рокицкий (1903-1977 гг.). Работая на кафедрах генетики МГУ, разведения и генетики Московского пушно-мехового института, зоологии позвоночных в Белорусском университете, он читал общий курс «Биологическая статистика», издал учебные пособия «Основы вариационной статистики для биологов» (1961 г.) и «Биологическая статистика» (1964, 1967, 1973 гг.), ставшие на протяжении многих лет основными учебниками биологической статистики в ВУЗах СССР.

Основателем Ленинградской школы биометриков был Ю.А. Филипченко, организовавший при Ленинградском университете первую кафедру генетики. Написанное им руководство по биометрии «Изменчивость и методы ее изучения» еще при жизни автора выдержало четыре издания (1923, 1925, 1927, 1929 гг.).

После смерти Ю.А. Филипченко, курс биометрии в ЛГУ читал его ученик Адольф Иванович Зуйтин (1891–1941 гг.), погибший в блокадном Ленинграде. Затем биометрию в ЛГУ стал преподавать П.В. Терентьев, талантливый ученик С.С. Четверикова. Он первый ввел преподавание биометрии применительно к большому практикуму по зоологии позвоночных животных. П.В. Терентьев был одним из первых организаторов курсов повышения квалификации биологов по вопросам применения математических методов в биологии и четырех Всесоюзных совещаний (1958-1964 гг.), посвященных этому вопросу.

Немало издается в эти годы и учебной статистической литературы, ориентированной на агрономов, биологов, врачей, педагогов. Многие издания таких авторов, как Юрия Леонидовича Поморского, Валентина Павловича Левинского и Андрея Афанасьевича Сапегина, выдерживали многократные переиздания.

В 1929-33 годах развернулась острая дискуссия в биологии, особенно в генетике, вокруг проблемы наследования приобретенных признаков и реальности генов. Она достигла своего апогея на печально известной августовской сессии 1948 г. ВАСХНИЛ, завершившейся разгромом генетики. Выступая на этой сессии с заключительным словом, Трофим Денисович Лысенко (в это время президент ВАСХНИЛ) окончательно сформулировал тезис о том, что: «теория вероятностей и статистика нужны только менделистам-морганистам, а «мичуринской биологии» эти науки не нужны».

Генетики были причислены к так называемому «меньшевиствующему идеализму» - течению, которое осудил и окрестил этим термином И.В. Сталин. В это время из Москвы был выслан С.С. Четвериков - создатель школы экспериментальной и популяционной генетики. Это был первый ощутимый удар по российской биометрической школе.

23 августа 1948 г. министр высшего образования СССР С.В. Кафтанов издает приказ «О состоянии преподавания биологических дисциплин в университетах и о мерах по укреплению биологических факультетов квалифицированными кадрами биологов-мичуринцев», согласно которому в вузах создавались комиссии, которые должны были пересмотреть учебные программы и тематику кандидатских работ аспирантов и т.д. Из библиотек изымались учебники по генетике и селекции, в большинстве которых содержалось подробное изложение биометрии, поскольку для анализа законов генетики их авторы использовали ее методы. Это было следующим шагом в ликвидации российской биометрической школы. Другим приказом министра высшего образования СССР был снят с поста ректора Тимирязевской сельскохозяйственной академии крупнейший ученый в области статистики сельского хозяйства, отважный защитник генетики, Василий Сергеевич Немчинов (1894-1964гг.).

После смерти И.В. Сталина в печати стали появляться отдельные статьи с критикой «лысенковщины». Наибольшего подъема они достигли в 1955 году, в котором отмечалось 100-летие Мичурина. Большой вклад в это внес А.А.

Любищев, который в СССР был единственным на тот момент членом Международного биометрического общества.

Со второй половины 50-х годов прошлого столетия, благодаря усилиям Алексея Андреевич Ляпунов (1911-1973 гг.), Акселя Ивановича Берга (1893-1979 гг.)– основоположников биологической кибернетики, Раисы Львовны Берг (1913-2006 гг.) – генетика и популяризатора науки, Льва Семеновича Каминского (1889-1962 гг.) – крупнейшего специалиста в области санитарной и демографической статистики, П.Ф. Рокицкого, Н.А. Плохинского и многих других, применение биометрических методов стало, если не обязательным, то, по крайней мере, желательным элементом биологических исследований.

В 1957 году возобновилось преподавание курса «Биометрия» в сельскохозяйственных вузах и на биологических факультетах.

Начиная с 60-х годов XX века, широкое развитие получило математическое моделирование. Формирование данного направления связано с именами Юрия Михайловича Свирежева(1938-2007 гг.), Никиты Николаевич Моисеева(1917-2000 гг.), Владимира Петровича Пасекова и Бенциона Шимоновича Флейшмана (г. Москва), Ратмира Александровича Полуэктова (1930-2012 гг.) и Владимира Васильевича Меншуткина (г. Санкт-Петербург), Вадима Александровича Ратнера (1932-2002гг.) (г. Новосибирск), Рэма Григорьевича Хлебопроса (1931-2017 гг.), Ивана Александровича Терскова (1918-1989 гг.) и Николая Сергеевича Абросова (г. Красноярск), Альберта Макарьевич Молчанова (1928-2011 гг.) и Александра Дмитриевича Базыкина (1940-1994 гг.) (г. Пущино), Александра Борисовича Горстко (г. Ростов-на-Дону), Александра Павловича Шапиро (г. Владивосток), Валерия Ивановича Беляева(1931-1999 гг.) (г. Севастополь), Юрия Серафимовича Колесова (1939-2009 гг.) (г. Ярославль), Геннадия Самуиловича Розенберга (г. Уфа, г. Тольятти).

Не малую роль в развитии биометрии в России сыграли научные объединения и общества, семинары и конференции ученых. На общественных началах долгие годы действовали комиссия по математической геоботанике Всесоюзного ботанического общества (председатель В.И. Василевич), биометрический семинар под руководством Олега Михайловича Калинина и Александра Георгиевича Барта, организованного в 1963 году в Ленинградском государственной университете. В 1974 г. Александр Петрович Левич создал и возглавил при Московском государственном университете рабочую группу конструктивных разработок по теоретической биологии, которая работала на протяжении десяти лет.

Во второй половине прошлого – начале нашего века опубликовано значительное количество книг отечественных ученых и педагогов с изложением методов математической статистики в биологии: В.Ю. Урбах (1964), Н.А. Плохинский (1960, 1970, 1980), Г.Ф. Лакин (1968,1980, 1983,1990), А.А. Любищев (1986), Э.В. Ивантер (1969), Э.В. Ивантер, А.В. Коросов (2004, 2005, 2011, 2013) и применительно к отдельным ее отраслям: методики полевого опыта в сельском хозяйстве - В.Г. Вольф (1966), Б.А. Доспехов (1972, 1985), почвоведении- Е.А. Дмитриев (1972, 1995), ботаники - Г.Н. Зайцев (1984, 1990), геоботаники - В.И. Василевич (1969), фитопатологии – Н.И. Минкевич, Т.И. Захарова (1977), генетики – П.Ф. Рокицкий (1974), агрометеорологии – Е.С. Уланова, О.Д. Сиротенко (1968),Е.С. Уланова, В.Н. Забелин (1990), гидрологии- А.В. Рождественский, А.И. Чеботарев (1974), А.В. Рождественский, А.В. Ежов (1986), океанографии – В.И. Беляев (1973, 1987), лесном хозяйстве – А.В. Жигунов, И.А. Маркова, А.С. Бондаренко (2002), А.С. Бондаренко, А.В. Жигунов (2016).

Биометрия в своем историческом развитии прошла долгий и сложный путь – от чисто словесного описания биологических объектов к их измерений, статистических сводок и таблиц к статистическому анализу массовых явлений. В настоящее время роль методов математической статистики в биологических исследованиях существенно возросла, а в связи с компьютеризацией и разработкой компьютерных программ возможности биометрии многократно увеличились, а необходимость в ее изучении при подготовке специалистов биологического профиля стала насущной необходимостью и велением времени.

Методы биометрии находят применение в ихтиологии. В пятидесятых-шестидесятых годах прошлого столетия ихтиологи стали активно осваивать и применять разнообразные математические приемы в своих трудах, до этого исследования в этой области развивалась на основе качественного анализа явлений. Широко применял при изучении рыб биометрический метод Иван Федорович Правдин (1880-1963 гг.), во многих работах он не просто использовал методику биометрического описания рыб, но серьезно ее улучшил, внеся существенные изменения в сложные средства вариационной статистики и придав ей современный вид. Метод изучения изменчивости рыб изложен им в известном «Руководстве по изучению рыб», выдержавшем три издания с 1926 по 1939 гг. при его жизни. Четвертое издание (1966) значительно переработанное и дополненное, Иван Федорович не успел закончить в связи со скоропостижной кончиной, и завершение данного труда взяли на себя его товарищи по совместной работе и ученики. Данное учебное пособие и в настоящее время используют многие ихтиологи.

На современном этапе развития ихтиологии невозможно обойтись без применения методов математического моделирования. Они позволяют получать не только количественные оценки, по и выявлять взаимосвязи в популяциях рыб. Г.В. Никольский еще в 1963 году разработал биологические основы принципа математического моделирования в ихтиологии, В.В. Меншуткин (1964 г.) применил данный метод в исследования динамики численности рыб. В 1993 году им было предложена имитационная модель водных экологических систем, логико-математическое описание объекта, которое может быть использовано для экспериментирования на компьютере в целях проектирования, анализа и оценки функционирования объекта.

В последнее время значительно возросли роль и значение методов статистической математики в ихтиологических работах. Большое внимание уделяется рассмотрению основ математической статистики, обработке результатов наблюдений, опытов и исследований, группировке экспериментального материала, выявлению важнейших статистических показателей точности и критериев существенности, измерению сопряженности.

МЕТОДИКИ СБОРА ИХТИОЛОГИЧЕСКИХ МАТЕРИАЛОВ И ПРАВИЛА ИХ ОБРАБОТКИ

Схемы измерения рыб различных семейств

В течение довольно длительного времени в ихтиологии и, в частности в систематике, господствовал описательный метод диагностической характеристики рыб. В настоящее время он уступил свое лидерство биометрическому, основанному на проведении многих промеров тела рыб для установления качественных (пластических) и количественных (меристических) признаков. Он стал особенно необходим при изучении внутривидовых разностей, когда признаки приходится устанавливать с применением вариационной статистики.

Морфометрические признаки рыб являются одним из основных индикаторов состояния популяций рыб в водных экосистемах, причем на них существенное влияние оказывают факторы среды обитания. Строение тела рыб, особенно экстерьера и конституции прямо или косвенно зависит от влияния условий жизни.

Математический метод помогает приведению в порядок многих систематических групп рыб, а это требует подробных схем измерений, но все они должны быть однообразны как по линиям промеров, так и по терминологии, чтобы результаты, полученные разными исследователями, можно было сравнивать между собой.

В настоящее время создано много схем для измерений различных семейств рыб. Шведским ихтиологом Ф.А. Смиттом в 1886 году разработана схема измерений рыб семейства лососевых, немецким зоологом Ф. Гейнке (1898 г.) – для сельдей. Отечественные ученые также внесли свой вклад в данное направление. Н.Ю. Зограф в 1887 году дал схему измерений осетровых, Б.С. Ильин (1927 г.) – бычков, В.И. Травин (1951 г.) – морских окуней, И.Ф. Правдин (1966 г.) – карповых.

Измерение рыб – это определение размеров отдельных особей, различных частей их анатомии и соотношений между частями тела. К пластическим признакам относят размеры рыб или то, что можно измерить (длина, высота, объем, вес и т. д.). В связи с тем, что с возрастом эти показатели подвержены изменчивости кроме абсолютных величин, рассчитывают процентные отношения частей тела.

Для измерения рыб применяют различные приборы и приспособления: измерительные ленты, мерные доски, штангенциркули, линейки, сантиметровые рулетки (рис.1).

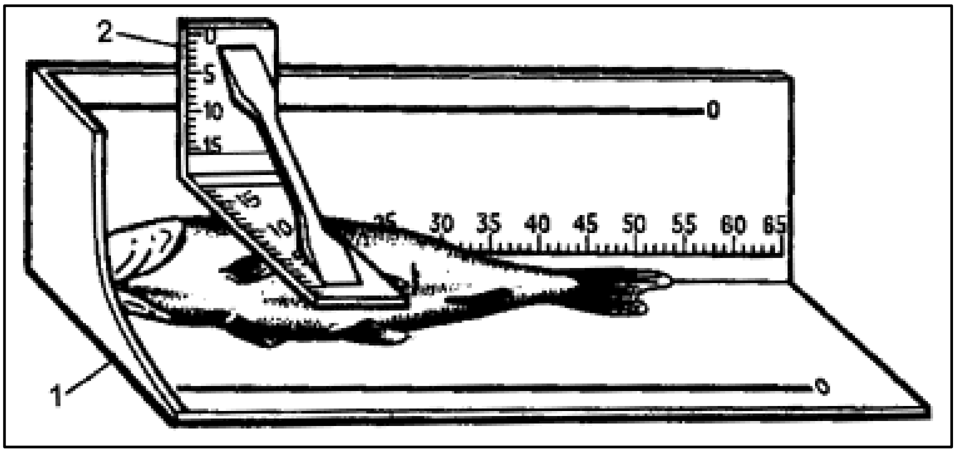


Рисунок 1 – Приспособления для измерения рыб:1 – бонитировочная доска;
2 – мерный треугольник

Линейные величины выражают в целых миллиметрах, малые (например, диаметр глаза) – с точностью до 0,5 мм.

Тело рыбы расчленено на голову, туловище и хвост. Границей между головой и туловищем считают наружную жаберную щель, между туловищем и хвостом - местоположение анального отверстия.

На голове рыб различают: рыло - пространство от начала головы до переднего края глаза, щеки - участки от глаз до заднего края правой и левой предкрышечных костей, лоб – промежуток между глазами и горло – пространство от жаберной перепонки до грудных плавников.

В головной части расположены: рот, ноздри – два мешочка, не сообщаются с полостью рта и глотки, глаза, жаберные крышки, прикрывающие жаберный аппарат, брызгальца – остатки нефункционирующей жаберной щели.

На голове и туловище рыб в большинстве случаев имеется боковая линия. Она отделяет спинную часть тела от брюшной. Это гидростатический орган рыб. Внешне она выглядит в виде ряда отверстий в чешуях.

Для регулирования положения тела в пространстве и движения рыбам служат парные и непарные плавники, образованные жесткими или мягкими лучами, обычно связанными кожной перепонкой. К непарным плавникам относятся вертикальные хвостовой, один или несколько спинных и анальный плавники. У некоторых рыб позади спинного плавника расположен мягкий вырост без лучей – жировой плавник. Парные грудные плавники лежат за жабрами. Парные брюшные плавники у разных видов могут располагаться впереди грудных, или под ними, или далеко позади них. Положение брюшных плавников является важным систематическим признаком разных групп рыб (рис.2).

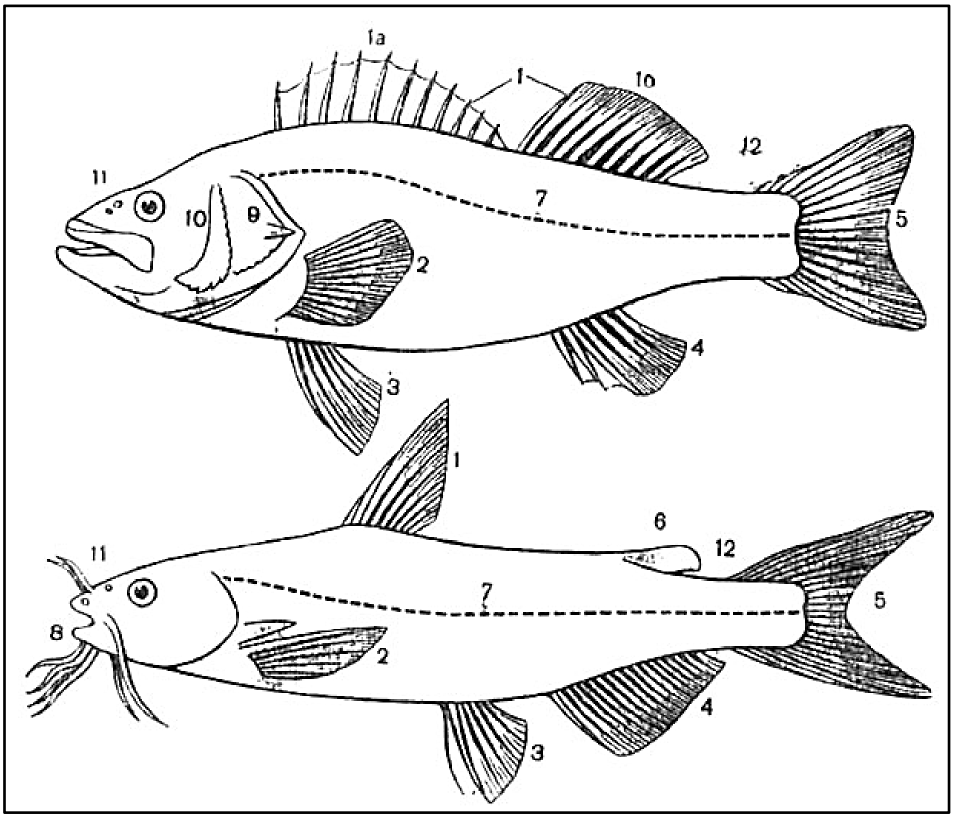


Рисунок 2 – Схема внешнего вида рыбы: 1 – спинной плавник (1а – колючая часть; 1б – мягкая часть); 2 – грудной плавник; 3 – брюшной плавник; 4 – анальный плавник; 5 – хвостовой плавник; 6 – жировой плавник; 7 – боковая линия; 8 – усики; 9 – жаберная крышка; 10 – предкрышка; 11 – ноздри; 12 – хвостовой стебель

При определении рыб иногда пользуются промерами частей их тела (рис.3).

Полная длина тела – расстояние от вершины рыла до вертикали конца наиболее длинной лопасти хвостового плавника при горизонтальном положении рыбы (до заднего края хвостового плавника).

Длина тела (промысловая длина тела) у всех рыб, кроме сельдевых и лососевых – расстояние от вершины рыла до конца чешуйчатого покрова у основания хвостового плавника (если чешуи нет, то до основания лучей хвостового плавника). У сельдевых и лососевых рыб – от вершины рыла до концов средних лучей хвостового плавника.

Длина головы – расстояние от вершины рыла до заднего конца жаберной крышки без учета окаймляющей ее перепонки.

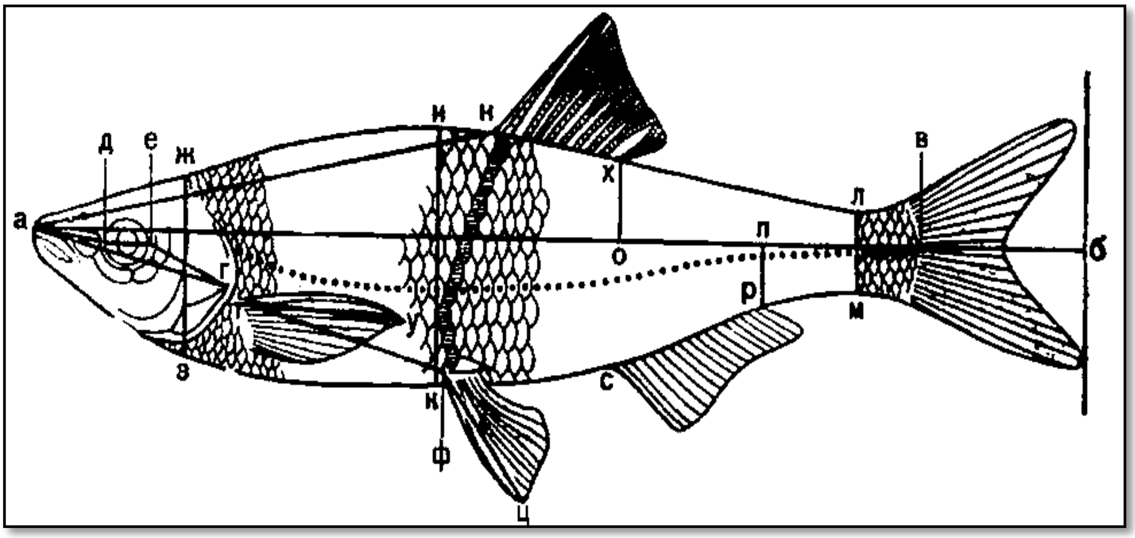


Рисунок 3 – Схема наиболее употребляемых стандартных промеров рыбы: аб – полная длина тела, ав – длина тела, аг – длина головы, ад – длина рыла, ам – антедорсальное расстояние, де –диаметр глаза, ег – заглазничный отдел головы, жз – высота головы, ик – наибольшая высота тела, лм –наименьшая высота тела, нх – длина спинного плавника, об – постдорсальное расстояние, пв – длина хвостового стебля, ср – длина анального плавника, ту – длина грудного плавника, фц – длина брюшного плавника, тф – расстояние между грудными и брюшными плавниками

Длина рыла или предглазничный отдел - расстояние от вершины рыла до переднего края глаза.

Антедорсальное – расстояние от вершины рыла до начала основания спинного плавника.

Диаметр глаза, если не оговорено особо, рассматривают горизонтальный диаметр. Измеряют диаметр роговицы, веки, если они есть, в расчет не принимают.

Заглазничный отдел головы - расстояние от заднего края глаза до заднего края жаберной крышки (без перепонки).

Высоту головы измеряют у затылка, место над прикреплением позвоночника к черепу или над задним краем верхнезатылочной кости.

Наибольшую высоту тела измеряют в том месте, где тело наиболее высокое, наименьшую- в наиболее низком месте тела, обычно близ основания хвостового плавника.

Длина спинного и анального плавников – расстояние от переднего края основания первого до заднего края основания последнего их луча.

Постдорсальное - расстояние от вертикали конца основания спинного плавника до основания хвостового плавника по средней линии тела.

Длина хвостового стебля - расстояние от вертикали конца основания анального плавника до основания хвостового плавника (или до конца чешуйчатого покрова) по средней боковой линии тела.

Длина грудных и брюшных плавников - расстояние от основания переднего края до вершины плавников.

Расстояние между грудными и брюшными плавниками измеряют между их основаниями, передняя часть брюха.

В схеме измерения рыб по И.Ф. Правдину добавлено еще ряд показателей (рис. 4).

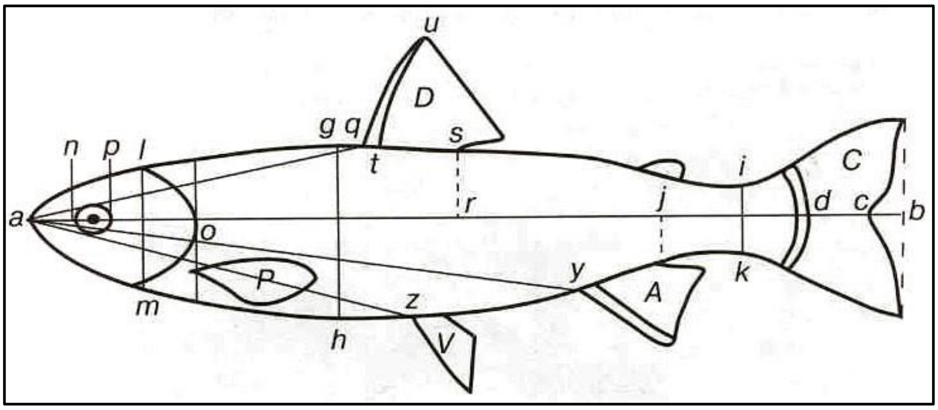


Рисунок 4 – Схема измерения рыб (Правдин, 1966)
D – спинной плавник, C – хвостовой плавник, A – анальный плавник, V – брюшной плавник, P – грудной плавник.

Ось тела – прямая линия, которая начинается от вершины рта и заканчивается у корней средних лучей хвостового плавника.

В наших водоемах наиболее многочисленными являются рыбы семейства карповые, поэтому приводим схему их измерений (рис. 5).

Кроме выше перечисленных стандартных параметров, при измерениях рыб семейства карповые применяют дополнительные промеры.

Длина всей рыбы (общая или абсолютная длина) – тоже, что полная длина тела (см. выше).

Длина тела по Смитту – расстояние от вершины рыла до конца средних лучей хвостового плавника.

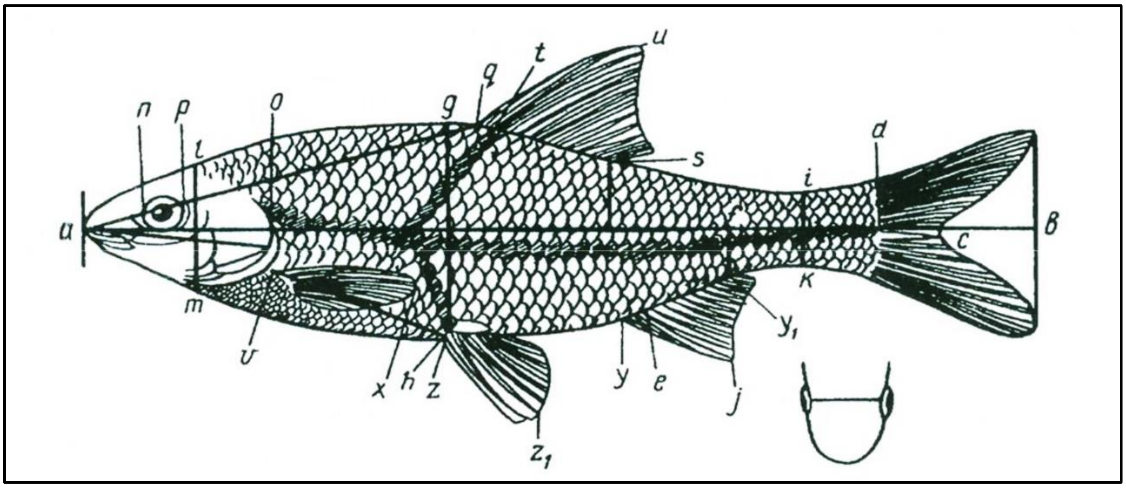


Рисунок 5 – Схема измерения рыб семейства карповые:
ab – длина всей рыбы; ac – длина по Смитту; ad – длина без С; od – длина туловища; an – длина рыла; np – диаметра глаза (горизонтальный); po – заглазничный отдел головы; ao – длина головы; lm – высота головы у затылка; gh – наибольшая высота тела; ik – наименьшая высота тела; aq– антедорсальное расстояние; rd – постдорсальное расстояние; fd – длина хвостового стебля; qs – длина основания D; tu – наибольшая высота D; yy_1 – длина основания А; ej – наибольшая высота А; vx – длина Р; zz_1 – длина V; yz – расстояние между Р и V; zy – расстояние между V и А.

Длина без С – расстояние от вершины рыла до конца чешуйного покрова.

Длина туловища - расстояние от заднего края жаберной крышки до конца чешуйного покрова.

Длина основания D – тоже, что длина спинного плавника.

Наибольшая высота D – высота наибольшего луча спинного плавника.

Длина основания А – тоже, что длина анального плавника.

Наибольшая высота А – высота наибольшего луча анального плавника.

Длина Р – расстояние от передней линии прикрепления грудного плавника до вершины наиболее длинного луча.

Длина V – расстояние от передней линии прикрепления брюшного плавника до вершины наиболее длинного луча.

Расстояние между Р и V – тоже, что расстояние между грудными и брюшными плавниками (см. выше)

Расстояние между V и А - расстояние между брюшным и анальным плавниками, задняя часть брюха.

Наибольший обхват тела измеряют сантиметровой лентой в месте наибольшей толщины тела, не беря в расчет плавники.

Наибольшая толщина тела – наибольшее расстояние между боками.

К меристическими, или количественным относят признаки, выражаемые числом элементов какого-либо органа такие как: число чешуй, позвонков, лучей в плавнике, глоточных зубов, количество жаберных тычинок и все другие отличия, определяемые в результате подсчета.

Определение возраста рыбы по чешуе

Рыба растет в течение всей жизни, поэтому при ихтиологическом исследовании важно знать ее возраст. Интенсивность роста рыб зависит от годовой цикличности физиологических процессов и закономерной их смены в связи с изменением образа жизни. Для определения возраста используют чешую, отолиты и кости скелета. Принцип определения возраста и роста рыб основан на свойстве чешуи и костей образовывать наслоения в виде чередующихся колец.

Чешуя у рыб представляет собой пластинку, на которой множество колечек и несколько резко выделяющихся широких кругов. Их число говорит о количестве прожитых рыбой лет. В осенне-зимний период рост рыбы замедлен в длину и расстояния между кольцами будут небольшими (темная зона). Весной и летом эти расстояния значительно шире (светлая зона). Каждая светлая зона к периферии постепенно переходит в темную. Полосы светлых (широких летних) и темных (узких зимних) зон, которые образуются в течение одного года, составляют годовое кольцо роста. Сколько на чешуе таких годовых колец, столько рыбе лет.

Часто между годовыми кольцами просматриваются добавочные, происхождение которых связывают с нерестом (лососи, сельди) или изменением интенсивности питания (карповые). Как правило, они выражены не по всей длине (рис.6).

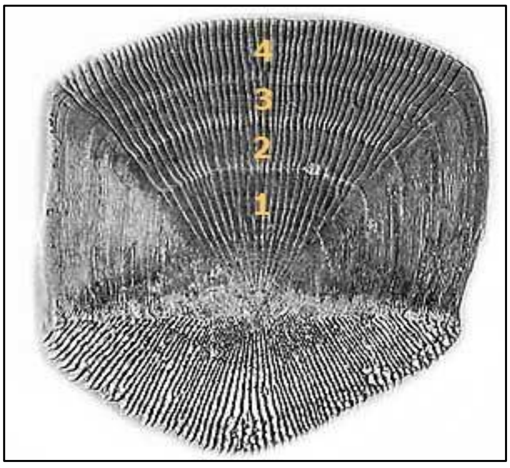


Рисунок 6 – Годичные кольца на чешуе рыбы

Чешуя отбирается с помощью скальпеля или ножа выше боковой линии, но недалеко от нее, в районе наибольшей высоты тела. Если у рыбы нет боковой линии, то чешую берут с середины бока под спинным плавником. В этом месте чаще всего чешуя крупная, с отчетливой скульптурой верхнего слоя и правильной формы.

Боковая или латеральная линия – орган чувств, воспринимающий движение и вибрации окружающей воды. У рыб она выглядит в виде ряда или нескольких рядов чешуи с отверстиями, напоминающими черточки. У большинства рыб боковая линия проходит в виде прямой линии по бокам тела от головы до хвостового плавника. У сельдевых, бычковых и некоторых других рыб боковой линии нет. Функцию ее выполняет сильно развитая система сенсорных каналов на голове или мелкие поры.

С каждой рыбы берут по 5-10 чешуй и кладут их в конверт или в особую книжку (размером 5-10 см) сделанную из писчей бумаги. Собранные чешуи хранят в сухом месте. При определении возраста их промывают в разведенном нашатырном спирте или в воде и очищают мягкой щеткой от покрывающей их слизи.

Возраст определяют обычно по передней части чешуи, ее просматривают под лупой или при малом увеличении микроскопа и подсчитывают количество годовых колец.

Определение степени упитанности рыб

Упитанность – универсальный показатель, который характеризует как содержание жира в организме, так и физиологическое состояние рыбы, и ее потребительскую ценность.

Для определения степени упитанности рыб широко пользуются коэффициентом Т. Фультона, предложенным им в 1902 году, при помощи которого можно выявить половые, сезонные и возрастные изменения:

Коэффициент Фультона:

$$Q = \frac{W \cdot 100}{L^3},$$

где Q – коэффициент упитанности;

W – вес рыбы, г;

L – длина тела рыбы от начала рыла до конца чешуйного покрова, см.

При вычислении упитанности следует пользоваться длиной туловища, то есть промысловой длиной, так как именно эта часть наиболее полно характеризует данный показатель у рыбы. Причем промысловую длину возводят в третью степень, потому что приращение веса идет пропорционально увеличению объема рыбы.

Однако в этом случае не удается устранить влияния веса гонад (семенники и яичники - парные лентовидные или мешковидные образования, подвешенные на складках брюшины: брыжейке, в полости тела рыбы), которые в ряде случаев могут превышать 15% веса тела. Кроме этого существенное влияние может оказывать вес пищи, особенно в период интенсивного нагула, когда масса пищевого комка в отдельных случаях достигает 30 % веса рыбы. Для исключения влияния этих факторов Ф. Кларком, в 1928 г. было предложено использовать при расчетах вес тушки без внутренностей.

Коэффициент Кларка:

$$K = \frac{m - m_1}{L^3} \cdot 100,$$

где K – коэффициент упитанности;

m – вес рыбы, г;

m_1 – вес внутренностей рыбы, г;

L – длина тела рыбы от начала рыла до конца чешуйного покрова, см.

Однако следует помнить, что устранение из тела внутренностей сопровождается удалением внутриполостного жира, количество которого также в значительной мере связано с упитанностью. Поэтому при изучении упитанности рыб лучше пользоваться одновременно обоими показателями.

МАТЕМАТИЧЕСКАЯ ОБРАБОТКА ИХТИОЛОГИЧЕСКИХ МАТЕРИАЛОВ

Методы и приемы, используемые в биометрии, предназначены для планирования и обработки экспериментов и наблюдений, при этом исследователь всегда имеет дело с количественными и качественными значениями различных признаков и свойств, проявляющихся в различных вариациях. Биометрия представляет собой своеобразный инструмент, способный выразить в числе и измерить значимость и надежность полученных результатов, заранее рассчитать и спланировать необходимую численность объектов для того или иного эксперимента. Она позволяет определить пределы случайных колебаний изучаемой величины и установить, достоверны или случайны наблюдаемые разницы между вариантами. Это дает возможность интерпретировать полученные результаты и решить вопрос о мере их достоверности при оценке.

При обработке данных экспериментов и наблюдений возникают три основные статистические задачи: оценка параметров распределения, сравнение их в различных выборочных совокупностях и выявление статистических связей, которые решают различными приемами и методами. В каждом отдельном случае выбор того или иного метода анализа определяется природой биологического явления.

При исследовании биологических объектов чаще всего приходится иметь дело с какой-то их частью, ан их объединением в целом. В связи с этим, любое количество значений или объектов, отличных друг от друга, или наоборот сходных по каким-либо признакам составляет совокупность, которую разделяют на генеральную и выборочную.

Генеральная – это вся подлежащая изучению совокупность объектов, состоящая из бесконечно большого количества отдельных единиц. Предположим, мы поставили цель исследовать изменчивость массы тела конкретного вида рыбы, тогда генеральной совокупностью будут данные о массе всех без исключения особях этого вида. Понятно, что при изучении природных популяций в подавляющем большинстве случаев мы не сможем иметь дело с генеральной совокупностью в ее полном объеме, а будем располагать лишь какой-то, как правило, очень небольшой, ее частью.

Выборочная совокупность (выборка) – часть генеральной, группа особей или объектов, отобранная методом случайного отбора для проведения на ней исследований, способная с определенной степенью достоверности характеризовать генеральную совокупность. Для того чтобы выборочная совокупность более точно отражала генеральную, необходимо учитывать следующие основные условия:

- выборка должна быть вполне представительной, то есть иметь определенное количество наиболее типичных особей генеральной совокупности;
- выборка должна быть объективной, то есть сформированной по принципу случайного отбора без субъективных влияний на ее состав;
- выборка должна быть качественно однородной (выделенные для опыта группы должны быть аналогами по видовым, возрастным, физиологическим и другим факторам);
- выборка должна быть достаточной по объему, чтобы более точно выразить особенности генеральной совокупности.

По объему выборки делятся на малочисленные, содержащие до 30 особей, и многочисленные.

Следовательно, выборочная совокупность представляет часть, а генеральная совокупность является целым, о котором мы судим по величинам, получаемым при обработке выборочной совокупности.

Оценку и образование статистических совокупностей производят по конкретным признакам, которые позволяют сравнивать и различать исследуемые значения между собой. По своей природе признаки делят на качественные или атрибутивные и количественные.

Качественные признаки не имеют количественного выражения, то есть они неизмеряемые (цвет, вкус, запах). Количественными являются признаки, варианты которых имеют числовое выражение и отражают размеры, масштабы изучаемого объекта или явления. Они могут быть счетными и метрическими. Счетные признаки учитывают путем подсчета, а мерные – измеряют, и те, и другие имеют числовые значения, называемые вариантами, а их изменение варьированием. По характеру варьирования их подразделяют на дискретные и непрерывные. Дискретные могут принимать только вполне определенные показатели, между которыми не бывает промежуточных и их варианты выражают в виде целых чисел. Непрерывные имеют любые значения, как целые, так и дробные.

По способу измерения признаки делят на первичные и вторичные. Первичные выражают единицу совокупности в целом, то есть абсолютные величины. Вторичные непосредственно не измеряют, а рассчитывают. То есть первичные признаки лежат в основе наблюдения статистической совокупности, а вторичные определяют в процессе обработки и анализе данных, и они представляют собой соотношение первичных признаков.

Расстановку вариант в порядке возрастания (или убывания) называют ранжированием (ранжированный ряд). Группу чисел, сгруппированных в классы в зависимости от величины изучаемого признака, называют вариационным рядом. Существующие между биологическими признаками связи, при которых определенному значению одного признака соответствует несколько значений другого признака, варьирующих около своей средней величины, называют корреляцией.

Биологические признаки, если они выражены при помощи счета или меры, приобретают значение математических величин: средняя арифметическая, средняя квадратическая, коэффициент изменчивости, коэффициент корреляции и ряд других.

Средние величины

Для того чтобы получить характеристики не отдельных объектов, а всей группы в целом, определяют среднюю величину признака. В зависимости от исследуемых объектов и от поставленных целей среднюю величину вычисляют различными способами. Средние статистические величины позволяют отвлечься от индивидуального варьирования признака и служат выражением типичного для всей совокупности значения.

Статистическая обработка методом средних величин заключается в замене индивидуальных показателей варьирующего признака некоторой уравновешенной средней величиной.

Средние величины делятся на два больших класса: степенные средние и структурные средние:

- Степенные средние:
    - M – средняя арифметическая;
    - G – средняя геометрическая;

- S – средняя квадратическая;
- H – средняя гармоническая.

➢ Структурные средние:

- Mo – мода;
- Me – медиана.

Для вычисления степенных средних используют все имеющиеся значения признака. Моду и медиану определяют в связи со структурой распределения и используют как среднюю характеристику в тех совокупностях, где расчет средних степенных невозможен или нецелесообразен.

Средняя арифметическая М

Средняя арифметическая (М) – наиболее распространенный и широко применяемый статистический показатель среднего значения варьирующего признака при количественном его выражении. Она представляет то значение признака, которое имела бы каждая единица совокупности, если бы общий итог всех показателей был распределен равномерно между всеми ее единицами, при этом сумма отрицательных и положительных отклонений от нее равна нулю.

В простых и наиболее общих случаях вычисление данной величины сводится к суммированию всех значений варьирующего признака и делению полученной суммы на общее количество единиц совокупности, то есть объема выборки:

Формула простой средней арифметической выглядит следующим образом:

$$M = \frac{\sum v}{n},$$

где М – средняя арифметическая;

$\sum$ - символ суммирования;

v – результат измерения признака у каждого объекта;

n – число особей в группе или число наблюдений.

Пример: средняя для пяти значений 1,2,3,4,5 равна:

$$M = \frac{1 + 2 + 3 + 4 + 5}{5} = 3$$

Простая средняя арифметическая – самый распространенный, но не единственный обобщающий показатель для характеристики варьирующих явлений. Часто для практических и научных целей необходимо объединить простые средние, полученные для однородного материала и на этой основе найти общее среднее, характеризующее весь изученный материал с учетом частоты повторяемости варианта. Такую величину называют средней взвешенной. Для ее расчета необходимо каждое значение признака (каждую вычисленную среднюю арифметическую) помножить на его «вес» (частоту встречаемости), все эти произведения сложить и сумму разделить на сумму весов.

Взвешенную среднюю арифметическую рассчитывают по следующей формуле:

$$M_{взв.} = \frac{\sum v \cdot p}{\sum p} = \frac{v_1 \cdot p_1 + v_2 \cdot p_2 + \dots + v_n \cdot p_n}{p_1 + p_2 + \dots + p_n},$$

где v– значение признака;

p – математический «вес» усредняемого значения.

Пример: определить среднюю полную длину тела тихоокеанской трески Берингова моря исходя из данных улова.

Возраст рыб	Количество рыб (p), шт.	Средняя полная длина тела (v), см
0+	3	16,5
1+	1	37,0
2+	4	52,4
3+	4	61,9
4+	5	70,4
5+	2	76,4
6+	5	75,8
7+	3	82,9
8+	1	83,9

$$Mвзв. = \frac{16,5 \cdot 3 + 37,0 \cdot 1 + 52,4 \cdot 4 + 61,9 \cdot 4 + 70,4 \cdot 5 + 76,4 \cdot 2 + 75,8 \cdot 5 + 82,9 \cdot 3 + 83,9 \cdot 1}{3 + 1 + 4 + 4 + 5 + 2 + 5 + 3 + 1} =$$

$$\frac{49,5 + 37,0 + 209,6 + 247,6 + 352,0 + 152,8 + 379,0 + 248,7 + 83,9}{28} =$$

$$\frac{1760,1}{28} = 62,86(см)$$

Значение средней арифметической характеризует целую группу однородных единиц наблюдения одним числом, является центром вариационного распределения, вокруг которого группируются отдельные значения выборочной совокупности, взаимопогашаются и отметаются случайные колебания от центральной тенденции. Она позволяет легко и быстро производить сравнительный анализ выборок разного объема.

Приведенные способы вычисления простой и взвешенной средней арифметической удобны в тех случаях, когда число наблюдений (n) в выборочной совокупности, подвергнутой обработке меньше 30, и выборка изображается в виде простого статистического ряда, образованного выписанными подряд значениями варьирующего признака.

В биологических материалах чаще всего приходится иметь дело с большим числом наблюдений. Для таких выборок существуют другие приемы вычисления средних, а также и других статистических величин.

При большом числе наблюдений в выборке вычислительной обработке подвергают не простой статистический ряд, а ряд, составленный из классов варьирующего признака, по которым произведена разноска данных, образующих значения частот (p).

Обработку большой выборки осуществляют, прежде всего, путем составления вариационного ряда, то есть расположения данных в определенном порядке (ранжировка). Ступени, на которые разбивается весь вариационный ряд, называют вариациями или классами.

В практической работе приходится иметь дело с такими признаками, величина которых колеблется в огромных пределах, и отдельные особи отличаются друг от друга на самые разнообразные величины. В этих случаях принято объединять в каждый класс особей с определенной величиной изменчивости, установленной заранее (например: 6-10, 11-15, 16-20 сми т.д.). Количество классов берут произвольно, при очень точных расчетах – до 15-20, но чаще 8-10.

<u>Порядок составления вариационного ряда:</u>

1. Найти лимит (Lim), то есть минимальную и максимальную величины, и определить размах колебаний изучаемого признака (промежуток между man. и min., путем вычитания min. из max.);
2. Определить число классов вариационного ряда;
3. Найти величину интервала К(допустимо округление), для чего нужно разделить полученную при вычитании min.от max. разницу на намеченное число

классов $K = \frac{max.-min}{число\ классов}$;

4. Установить границы классов (начало и конец);
5. Вычислить среднюю величину признака в каждом классе;
6. Произвести разноску материала в соответствующие классы.

Пример построения вариационного ряда: получены следующие данные длины туловища (промысловой длины) омуля байкальского возраста 5+ оз. Байкал (см), n=35: 25,9; 21,5; 23,9; 22,8; 21,6; 24,9; 26,0; 25,1; 23,6; 25,5; 22,9; 23,8; 26,5; 24,7; 26,3; 25,6; 25,9; 24,8; 24,2; 23,6; 24,7; 25,6; 25,2; 23,9; 25,4; 22,5; 21,7; 26,1; 23,0; 25,4; 23,6; 24,3; 24,0; 22,9; 22,1.

Находим минимальный и максимальный показатели и убеждаемся, что средняя длина колеблется от 21,5 до 26,5 см.

Размах колебаний признака – Lim= 25,5-21,5=5,0 (см).

Делим эту величину на количество классов (5:10=0,5) с интервалом К=0,5, устанавливаем границы классов, находим среднюю величину промысловой длины туловища омуля в каждом классе, проводим разноску полученных данных в соответствующие классы и подсчитывает частоты в каждом классе:

Классы	Средняя величина класса	Частоты, р
21,5-21,9	21,7	··· 3
22,0-22,4	22,2	· 1
22,5-22,9	22,7	···· 4
23,0-23,4	23,2	· 1
23,5-23,9	23,7	······ 6
24,0-24,4	24,2	··· 3
24,5-24,9	24,7	···· 4
25,0-25,4	25,2	···· 4
25,5-25,9	25,7	····· 5
26,0-26,4	26,2	··· 3
26,5-26,9	26,7	· 1
		$\sum p=35$

Данные вариационного ряда подвергают дальнейшей статистической обработке, как для получения значения средней арифметической (М), так и других статистических показателей.

Обработку вариационных рядов можно осуществлять приемом, называемым «Метод условных отклонений с применением способа произведений», который используют и при способе сумм. Формула средней арифметической для большой выборки по способу произведений выражается следующим образом:

$$M = A + K \cdot \frac{\sum p \cdot a}{n},$$

где М – средняя арифметическая;

А – условная средняя, приближающаяся по своему значению к средней арифметической;

К – величина класса;

р – частоты вариационного ряда;

n = ∑р – число наблюдений в данном вариационном ряду или объем данной совокупности;

а – условное отклонение каждого класса от класса, в котором находится условная средняя (а), выраженное числом классов.

По этой формуле произведем вычисления по данным вариационного ряда:

Классы	Средняя величина класса	Частоты, р
21,5-21,9	21,7	··· 3
22,0-22,4	22,2	· 1
22,5-22,9	22,7	···· 4
23,0-23,4	23,2	· 1
23,5-23,9	А=23,7 Условно средний класс	······ 6
24,0-24,4	24,2	··· 3
24,5-24,9	24,7	···· 4
25,0-25,4	25,2	···· 4
25,5-25,9	25,7	····· 5
26,0-26,4	26,2	··· 3
26,5-26,9	26,7	· 1
		∑р=35

После того как написан ряд из значений классов (v) и ряд частот (р), производят выделения класса, серединой которого является значение условной средней (А). В качестве условного среднего класса берут тот, который занимает центральное место и имеет наибольшее число наблюдений, то есть наибольшее значение частот (р) по сравнению с другими классами. В нашем примере это класс с интервалом 23,5-23,9 (5-й сверху). Значение А представляет собой середину нулевого класса. Выделив класс с условной средней (А), отделяем его от остальных и принимает за нулевой от которого по порядку нумеруем остальные классы, что и будет выражать условное отклонение (а) каждого класса от нулевого. В сторону уменьшения признака (в нашем примере от нулевого класса вверх) условные отклонения (а) для каждого класса будут иметь знак минус (-а), а в сторону увеличения признака (вниз от нулевого класса) - знак плюс (+а). Запишем в таблице графу из значений (а).

Согласно формуле средней арифметической, для каждого класса требуется найти произведение р·а, что и записываем в четвертом столбике таблицы. После этого находим сумму р·а:

Классы	Частоты, р	Условное отклонение, а	р·а
21,5-21,9	3	-4	-12
22,0-22,4	1	-3	-3
22,5-22,9	4	-2	-8
23,0-23,4	1	-1	-1
23,5-23,9 А=23,7	6	0	0
24,0-24,4	3	1	3
24,5-24,9	4	2	8
25,0-25,4	4	3	12
25,5-25,9	5	4	20
26,0-26,4	3	5	15
26,5-26,9	1	6	6
	∑р=35		∑р·а=40

Подставим полученные значения в формулу М:

$$\mathrm{M} = \mathrm{A} + \mathrm{K} \cdot \frac{\sum \mathrm{p} \cdot \mathrm{a}}{\mathrm{n}} = 23{,}7 + 0{,}5 \cdot \frac{40}{35} = 24{,}27\ (\text{см})$$

Свойства средней арифметической

В средней арифметической происходит как бы устранение варьирования признака и установление его обобщающего абстрактного среднего уровня, присущего для данной совокупности при конкретной изменчивости признака. Таким образом, средняя арифметическая является обобщенным статистическим параметром или статистической характеристикой среднего уровня варьирующего признака.

Средняя арифметическая величина абстрактная, так как при ее вычислении можно получать такие дробные значения, которые в действительности не могут иметь место в связи с природой самого признака. В тоже время, средняя величина имеет конкретное выражение, показывая величину признака в том же именовании, в котором он измерялся.

Основное свойство средней арифметической состоит в том, что сумма отклонений каждого варианта от средней арифметической данной совокупности всегда равно нулю:

$$\sum (v\text{-}M) = 0$$

То есть если произвести отклонение каждого члена выборки по значению его варьирующего признака от значения средней арифметической, то такая сумма всегда будет равна нулю.

Следующее свойство средней арифметической заключается в том, что сумма отклонений варьирующего признака от любой другой величины, например, от условной средней (А), не равной ей, будет всегда больше нуля:

$$\sum (v\text{-}A) \neq 0$$

Это выражение используется при вычислении средней арифметической при большом числе наблюдений.

Сумма квадратов отклонений от суммы средней арифметической всегда меньше суммы квадратов отклонений, взятых от любого другого числа, отличающегося от средней арифметической:

$$\sum(v\text{-}M)^2 < \sum(v\text{-}A)^2$$

Произведение средней на сумму частот всегда равно сумме произведений вариант (отдельных значений) на частоты.

Если от каждой варианты отнять (прибавить) какое-либо произвольное число, то новая средняя уменьшится (увеличится) на то же число.

Если каждую варианту умножить (разделить) на какое-то произвольное число, то новая средняя увеличится (уменьшится) во столько же раз.

Если все частоты (веса) разделить или умножить на какое-либо число, то средняя арифметическая от этого не изменится.

Кроме указанных основных свойств, средняя арифметическая имеет и другие особенности, которые используют в формулах различных статистических величин.

Средняя геометрическая G

Средняя геометрическая – величина, которая выявляет средний прирост (или среднее уменьшение) какого-либо показателя за определенный период времени. Она дает возможность сохранять в неизменном виде не сумму, а произведение индивидуальных значений.

Средняя геометрическая необходима для определения среднего значения признака, если от характеризует темп роста, темп увеличения численности популяции. Особенно она удобна в тех случаях, если признак выражен в долях единицы или в процентах и изменяется во времени и по периодам.

Формула средней геометрической имеет следующий вид:

$$\mathrm{G} = \sqrt[\mathrm{n}]{\mathrm{v}_1 \cdot \mathrm{v}_2 \cdot \cdots \cdot \mathrm{v}_\mathrm{n}},$$

где v – значение варьирующего признака;

n – число наблюдений в выборке.

В формуле средней геометрической под корнем стоит произведение вариантов, число которых равно величине наблюдений. Степень корня также, соответствует числу наблюдений.

Для упрощения расчетов в тех случаях, когда степень корня больше двух, производят логарифмирование формулы, а затем по полученному логарифму средней геометрической находят ее абсолютное значение.

Логарифм корня равен логарифму подкорневой величины, деленной на показатель степени корня, а логарифм произведения – сумме логарифмов, взятых для каждого члена произведения. Следовательно, логарифмирование формулы средней геометрической даст следующее выражение:

$$\mathrm{lgG} = \frac{\mathrm{lgv}_1 + \mathrm{lgv}_2 + \cdots + \mathrm{lgv}_\mathrm{n}}{\mathrm{n}} = \frac{\sum \mathrm{lgv}}{\mathrm{n}}$$

Так, при определения средней геометрической значений варьирующего признака (v) величин: 5, 8, 10, 12, 13 логарифмированием, вычисления выглядят следующим образом:

$$G = \sqrt[5]{5 \cdot 8 \cdot 10 \cdot 12 \cdot 13}$$

$$\lg G = \frac{\lg 5 + \lg 8 + \lg 10 + \lg 12 + \lg 13}{5}$$

Используя таблицу десятичных логарифмов(табл.1, 2 приложения) находим логарифм средней геометрической.

Таблица десятичных логарифмов (фрагмент)

00 -∞	01 0	02 0,30103	03 0,47712	04 0,60206	05 0,69897	06 0,77815	07 0,8451	08 0,90309	09 0,95424
10 1	11 1,04139	12 1,07918	13 1,11394	14 1,14613	15 1,17609	16 1,20412	17 1,23045	18 1,25527	19 1,27875
20 1,30103	21 1,32222	22 1,34242	23 1,36173	24 1,38021	25 1,39794	26 1,41497	27 1,43136	28 1,44716	29 1,4624

lg5=0,69897≈0,6990

lg8=0,90309≈0,9031

lg10=1,00000=1,0000

lg12=1,07918≈1,0792

lg13=1,11394≈1,1139.

$$\lg G = \frac{\lg 5 + \lg 8 + \lg 10 + \lg 12 + \lg 13}{5}$$
$$= \frac{0{,}6990 + 0{,}9031 + 1{,}000 + 1{,}0792 + 1{,}1139}{5} = \frac{4{,}7952}{5}$$
$$= 0{,}959$$

По lgG находим по таблицам логарифмов абсолютное значение средней геометрической.

Малая таблица десятичных логарифмов

Число	Логарифм	Число	Логарифм	Число	Логарифм	Число	Логарифм
1,0	0,0000	3,3	0,5185	5,6	0,7482	7,8	0,8921
1,1	0,0414	3,4	0,5315	5,7	0,7559	7,9	0,8976
1,2	0,0792	3,5	0,5441	5,8	0,7634	8,0	0,9031
1,3	0,1139	3,6	0,5563	5,9	0,7709	8,1	0,9085
1,4	0,1461	3,7	0,5682	6,0	0,7782	8,2	0,9138
1,5	0,1761	3,8	0,5798	6,1	0,7853	8,3	0,9191
1,6	0,2041	3,9	0,5911	6,2	0,7924	8,4	0,9243
1,7	0,2304	4,0	0,6021	6,3	0,7993	8,5	0,9294
1,8	0,2553	4,1	0,6128	6,4	0,8062	8,6	0,9345
1,9	0,2788	4,2	0,6232	6,5	0,8129	8,7	0,9395

2,0	0,3010	4,3	0,6335	6,6	0,8195	8,8	0,9445
2,1	0,3222	4,4	0,6435	6,7	0,8261	8,9	0,9494
2,2	0,3424	4,5	0,6532	6,8	0,8325	9,0	0,9542
2,3	0,3617	4,6	0,6628	6,9	0,8388	9,1	0,9590
2,4	0,3802	4,7	0,6721	7,0	0,8451	9,2	0,9638
2,5	0,3879	4,8	0,6812	7,1	0,8513	9,3	0,9685
2,6	0,4150	4,9	0,6902	7,2	0,8573	9,4	0,9731
2,7	0,4314	5,0	0,6990	7,3	0,8633	9,5	0,9777
2,8	0,4472	5,1	0,7076	7,4	0,8692	9,6	0,9823
2,9	0,4624	5,2	0,7160	7,5	0,8751	9,7	0,9868
3,0	0,4771	5,3	0,7243	7,6	0,8808	9,8	0,9912
3,1	0,4914	5,4	0,7324	7,7	0,8865	9,9	0,9956
3,2	0,5051	5,5	0,7404				

$$\lg G = 0{,}959 = G = 9{,}1$$

Средняя геометрическая имеет ряд особенных свойств:

1. Произведение чисел ряда, для которого определяют среднюю геометрическую, всегда равно произведению, полученному от возведения средней геометрической в степень, равную числу членов ряда, то есть:

$$G^n = v_1 \cdot v_2 \cdot \ldots v_n$$

Это свойство средней геометрической используют для проверки правильности ее вычисления.

В нашем примере это свойство выражается следующим образом:

$$G^5 = 9{,}1 \cdot 9{,}1 \cdot 9{,}1 \cdot 9{,}1 \cdot 9{,}1 = 62403$$

$$G^5 = 5 \cdot 8 \cdot 10 \cdot 12 \cdot 13 = 62400$$

Поскольку оба конечных произведения практически равны (разница образовалась в результате округлений), то средняя геометрическая вычислена правильно.

2. Среднюю геометрическую следует применять в тех случаях, когда показателями служат отношения, выраженные в долях единицы или в процентах.

3. Средняя геометрическая пригодна для таких вариационных рядов, у которых имеется асимметричное распределение частот.

4. Произведение отношений средней геометрической к числам, которые меньше ее, всегда равно произведению отношений чисел ряда, превышающих ее по своему значению. По данным нашего примера, это правило дает такое

равенство:

Числа ряда превышающие G

$$\frac{9,1}{5}\cdot\frac{9,1}{8}=\frac{10}{9,1}\cdot\frac{12}{9,1}\cdot\frac{13}{9,1}$$

Числа ряда меньше G

$$\frac{82,81}{40}=\frac{1560}{753,571}, \text{то есть } 2,07=2,07$$

Это свойство средней геометрической наиболее важное и оно указывает на то, что она представляет собой среднюю из отношений и может служить удобной величиной для вычисления среднего прироста различных показателей во времени.

Для определения прироста какого-либо показателя за определенный период времени, на основании данных о приросте по частным периодам, составляющих общий период времени пользуются формулой среднего прироста:

$$x=G-1=\sqrt[n]{(1+a_1)\cdot(1+a_2)\cdot\ldots\cdot(1+a_n)}$$

где x – средний прирост за n периодов равной длительности;

a - фактический (или плановый) прирост за каждый период, выраженный в долях (%/100);

n - число периодов, за которые определяют прирост;

G – средняя геометрическая.

Для вычисления прироста (x) прежде требуется определить среднюю геометрическую (G), используя при этом логарифмирование.

Разберем на примере вычисление среднего прироста за общий период по данным процентного прироста за частные и более короткие, но равные сроки.

Пример: определить среднегодовой увеличение массы байкальского омуля оз. Байкал за двенадцать лет, по средним данным возрастных групп

Находим величины показателей прироста по периодам вычитая из определяемого значения предыдущее, суммируем их, принимаем их за 100%, находим процент прироста по периодам, долю прироста для каждой возрастной группы, к полученным значениям прибавляем1 и вычисляем (используя таблицу (табл.3 приложения)) десятичные логарифмы рассчитанных величин.

Возрастная группа	Средняя масса, г	Величина прироста по периодам, г	Процент прироста по периодам, %	Прирост, выраженный в долях (%/100) а	1+а	lg (1+а)
1+	13	13	2,23	0,022	1,022	0,0095
2+	39	26	4,45	0,044	1,044	0,0187
3+	72	33	5,65	0,056	1,056	0,0237
4+	130	58	9,93	0,099	1,099	0,0410
5+	176	46	7,88	0,079	1,079	0,0330
6+	245	69	11,82	0,118	1,118	0,0484
7+	329	84	14,38	0,144	1,144	0,0584
8+	419	90	15,41	0,154	1,154	0,0622
9+	513	94	16,10	0,161	1,161	0,0649
10+	555	42	7,19	0,072	1,072	0,0302
11+	570	15	2,57	0,026	1,026	0,0111
12+	584	14	2,40	0,024	1,024	0,0103
		∑584-100%				∑lg (1+а) = 0,4114

Расчет дробных частей десятичных логарифмов: $10^{x} = 10^{\{x\}} \cdot 10^{[x]}$, где {x} — дробная часть x, а [x] — целая часть x.

Пример: lg1,022=0,0086+0,0009=0,0095

Мантиссы (дробные части) десятичных логарифмов (фрагмент)

N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
10	0	43									4	9	13	17	22	26	30	35	39
			86	128	170						4	9	13	17	21	25	30	34	38
						212	253				4	8	12	16	21	25	29	33	37
								294	334	374	4	8	12	16	20	24	28	32	36
11	414	453	492								4	8	12	16	20	24	27	31	35
				531	569	607					4	8	11	15	19	23	27	30	34
							645	682	719	755	4	7	11	15	18	22	26	29	33
12	792	898	864	899	934						3	7	11	14	18	21	25	28	32

$$\lg G = \frac{\sum \lg(1 + a)}{n} = \frac{0{,}4114}{12} = 0{,}0343$$

Используя таблицу десятичных антилогарифмов (табл.4 приложения) находим величину средней геометрической.

Десятичные антилогарифмы (фрагмент)

m	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
,00	1000	1002	1005	1007	1009	1012	1014	1016	1019	1021	0	0	1	1	1	1	2	2	2
,01	1023	1026	1028	1030	1033	1035	1038	1040	1042	1045	0	0	1	1	1	1	2	2	2
,02	1047	1050	1052	1054	1057	1059	1062	1064	1067	1069	0	0	1	1	1	1	2	2	2
,03	1072	1074	1076	1079	1081	1084	1086	1089	1091	1094	0	0	1	1	1	1	2	2	2
,04	1096	1099	1102	1104	1107	1109	1112	1114	1117	1119	0	1	1	1	1	2	2	2	2
,05	1122	1125	1127	1130	1132	1135	1138	1140	1143	1146	0	1	1	1	1	2	2	2	2
,06	1148	1151	1153	1156	1159	1161	1164	1167	1169	1172	0	1	1	1	1	2	2	2	2
,07	1175	1178	1180	1183	1186	1189	1191	1194	1197	1199	0	1	1	1	1	2	2	2	2
,08	1202	1205	1208	1211	1213	1216	1219	1222	1225	1227	0	1	1	1	1	2	2	2	3
,09	1230	1233	1236	1239	1242	1245	1247	1250	1253	1256	0	1	1	1	1	2	2	2	3

$$\lg 0{,}0344 = 1{,}081 + 0{,}0001 = 1{,}082$$

$$x = (G-1)\cdot 100\% = (1{,}082-1)\cdot 100\% = 8{,}2\%$$

Переведем найденный среднегодовой процент в величину прироста (в г):

$$x = 584\cdot 8{,}2/100 = 47{,}89 \text{ (г)}$$

Таким образом, зная среднюю геометрическую, которая равна 1,082, можно определить среднегодовое увеличение массы байкальского омуля оз. Байкал за двенадцать лет. Средний прирост оказался равным 8,2% или 47,89 г.

Средняя квадратическая S

Среднюю квадратическую используют для признаков, которые характеризуются площадью круга, и для ее получения измеряют величину диаметра, или средних величин сторон и квадратных участков.

К их числу могут быть отнесены диаметры мышечного волокна на поперечном срезе, икринок, глаз, альвеолярных пузырьков молочной или щитовидной желез, колоний микробов, вакуолей у инфузорий, отдельных клеток или их ядер и т п.

Когда необходимо сохранить неизменной сумму квадратов исходных величин при их замене на среднюю, то средняя является квадратической и характеризует изменение показателей по сравнению с их базисными, взятых в их усредненной величине.

Простую среднюю квадратическую вычисляют по формуле:

$$S = \sqrt{\frac{\sum v^2}{n}},$$

где v^2 - значение варьирующего признака, взятое в квадрате;

n – число наблюдений.

То есть она равна корню квадратному из суммы квадратов признаков, деленному на их число.

Пример: определить средний диаметр икринок речного окуня (мм), если получены следующие показатели: 0,77; 0,84; 1,00; 0,94; 0,70; 0,88; 0,33; 0,36; 0,59.

$$S = \sqrt{\frac{0{,}77^2 + 0{,}84^2 + 1^2 + 0{,}94^2 + 0{,}70^2 + 0{,}88^2 + 0{,}33^2 + 0{,}36^2 + 0{,}59^2}{9}}$$

$$= \sqrt{\frac{0{,}59 + 0{,}71 + 1 + 0{,}88 + 0{,}49 + 0{,}77 + 0{,}11 + 0{,}13 + 0{,}35}{9}}$$

$$= \sqrt{\frac{5{,}03}{9}} = \sqrt{0{,}559} = 0{,}75 \text{ (мм)}$$

Также как для средней арифметической в случае получения обширного неоднородного материала, рассчитывают среднюю квадратическую взвешенную с учетом частоты повторяемости варианта. Для ее определения каждый квадратный корень признака умножают на его «вес» (частоту встречаемости), произведения складывают, сумму делят на сумму весов и из полученного результата извлекают квадратный корень.

Формула средней квадратической взвешенной:

$$S\text{взв.} = \sqrt{\frac{\sum v^2 \cdot p}{\sum p}},$$

где v^2 – значение варьирующего признака, взятое в квадрате;

n – число наблюдений;

p – математический «вес» усредняемого значения.

Средняя гармоническая H

Средняя гармоническая необходима для вычисления средних значений, получаемых во времени. Для этих процессов характерно, что при увеличении одного показателя другой изменяется в обратном направлении, то есть уменьшается. В связи с этим, средняя гармоническая позволяет обрабатывать такие совокупности, у которых значение варьирующего признака находится в обратном соотношении по отношению к суммарному результату.

Эти особенности можно показать на следующем примере: чем быстрее идет человек, тем меньше он тратит времени на прохождение пути.

Среднюю гармоническую рассчитывают по формуле:

$$\mathrm{H} = \frac{\mathrm{n}}{\sum \frac{1}{\mathrm{v}}},$$

где n – число значений признака;

v – величина варьирующего признака.

Пример: 1-йстудент для решения задачи тратит 10 мин., 2-й - 12 мин., 3-й - 15 мин., 4-й - 11 мин., 5-й - 13 мин. Определить среднее время выполнения задания.

Если для решения задачи использовать показатель средней арифметической, то получим:

$$\mathrm{M} = \frac{\sum \mathrm{v}}{\mathrm{n}} = \frac{10+12+15+11+13}{5} = 12{,}2 \text{ (мин.)}$$

Данная средняя определена не верно. Так как, если 1-й студент на выполнение одной задачи расходует 10 мин., то за одно занятие, длительностью 90 минут он может решить 9 задач (90:10=9), 2-й – 7,5 задач (90:12=7,5), 3-й – 6 задач (90:15=6), 4-й – 8,18 задач (90:11=8,18), 5-й – 6,92 задач (90:13=6,92). Вместе они решат ∑9+7,5+6+8,18+6,92=37,6 задач и затратят 450 минут. Если поделить общие трудозатраты на общее возможное выполнение задач, то будет получено другое значение общей арифметической:

$$\mathrm{M} = \frac{\sum \mathrm{v}}{\mathrm{n}} = \frac{450}{37{,}6} = 11{,}97 \text{ (мин.)}$$

Поэтому, среднее время на выполнение задач, вычисляют по формуле средней гармонической, так как индивидуальные значения частот встречаются по 1 разу.

$$\mathrm{H} = \frac{5}{\frac{1}{10} + \frac{1}{12} + \frac{1}{11} + \frac{1}{15} + \frac{1}{13}} = 11{,}9 \text{ (мин.)}$$

Средняя гармоническая - показатель, обратный средней арифметической.

Мода Мо

Модой, или модальным вариантом, называют наиболее часто встречающиеся значения. Он может быть выражен как качественным, так и количественным признаками, не зависит от крайних значений вариант и применяется для характеристики центра в рядах распределения с неопределенными границами.

Для количественных признаков модальной считают ту величину признака (веса, размера и т.п.), которой будет обладать большее число объектов (членов) генеральной совокупности в случайной выборке.

В дискретном вариационном ряду моду определяют визуально, и она равна варианте с наибольшей частотой или частостью.

Величину моды рассчитывают по следующим формулам:

$$\mathrm{Mo} = \mathrm{V_{Mo}} + \mathrm{K} \cdot \frac{\mathrm{p_2 - p_1}}{\mathrm{2p_2 - p_1 - p_3}},$$

$$\mathrm{Mo} = \mathrm{V_{Mo}} + \mathrm{K} \cdot \frac{\mathrm{p_2 - p_1}}{\mathrm{(p_2 - p_1) + (p_2 - p_3)}},$$

где V_{Mo} – начало модального класса;

К – величина класса;

p_1 – частота класса, предшествующая модальному;

p_2 – частота модального класса;

p_3 – частота класса, следующего за модальным.

Вычислим величину моды для вариационного ряда массы байкальского омуля возраста 7+ если отдельные измерения дали следующие результаты (г): 279, 376, 400, 271, 355, 342, 314, 235, 277, 290, 314, 337, 312, 356, 332, 241, 345, 290, 336, 325, 328, 381, 347, 357.

Находим минимальный и максимальный показатели и убеждаемся, что масса колеблется от 235 до 400 г.

Размах колебаний признака – Lim 400-235= 165 (г).

Делим эту величину на количество классов (165:10=16,5 г), округляем до целого значения ≈17 г, то есть интервал К=17. Разбиваем вариационный ряд на классы и разносим по ним имеющиеся значения массы байкальского омуля:

Классы варьирующего признака	235-251	252-268	269-286	287-303	304-320	321-337	338-354	355-371	372-388	389-405
Частота, р	··2	0	···3	··2	···3	·····5	···3	···3	··2	·1

В данном примере модальным классом является класс, имеющий частоту р=5, начало этого класса V_{Mo}=321, p_1 – частота класса, предшествующего модальному=3, p_2 – частота модального класса=5, p_3 – частота класса, следующего за модальным=3.

Произведем вычисление используя обе формулы:

$$\text{Mo} = V_{\text{Mo}} + \text{K} \cdot \frac{p_2 - p_1}{2p_2 - p_1 - p_3} = 321 + 17 \cdot \frac{5-3}{2 \cdot 5 - 3 - 3} = 321 + 17 \cdot \frac{2}{4}$$
$$= 321 + 17 \cdot 0{,}5 = 321 + 8{,}5 = 329{,}5 \ (\text{г})$$

$$\text{Mo} = V_{\text{Mo}} + \text{K} \cdot \frac{p_2 - p_1}{(p_2 - p_1) + (p_2 - p_3)} = 321 + 17 \cdot \frac{5-3}{(5-3) + (5-3)}$$
$$= 321 + 17 \cdot \frac{2}{4} = 321 + 17 \cdot 0{,}5 = 321 + 8{,}5 = 329{,}5 \ (\text{г})$$

Величина модального варианта может совпадать или отличаться от значения средней арифметической.

Модальная величина особенно удобна для характеристики качественных признаков, что имеет распространение при изучение генетических особенностей альтернативных признаков. Например, модальными будут доминантные признаки.

Медиана Ме

Медианой называют вариант, значение которого делит всю совокупность наблюдений или данных на две равные части. Одна половина объектов будет иметь значения варьирующего признака меньше, а другая половина - больше ее.

При определении медианы для количественных признаков при малом числе наблюдений члены выборки записывают подряд в возрастающем или убывающем порядке. Средний член такого ранжированного ряда будет служить ее показателем.

Например: имеем ряд, состоящий из нечетного количества чисел: 5, 4, 3, 3, 3, 2, 2, 2, 2. В этом случае медиана будет соответствовать числу, находящемуся в середине, то есть пятому по порядку: Ме=3.

Есть ряд имеет четное количество значений: 5, 4, 3, 3, 3, 2, 2, 2, 2, 1, величину медианы вычисляют как среднее из двух вариант, составляющих середину ряда (5-я и 6-я варианты):

$$\text{Me} = \frac{3+2}{2} = 2{,}5$$

При большом числе наблюдений величину медианы вычисляют по формуле:

$$Me = V_{Me} + K\frac{i_1 - i_2}{p_{Me}},$$

где V_{Me} – начало класса, в котором находится медиана;

K – величина класса;

i_1 – число вариантов или сумма накоплений частот, соответствующих половине всех наблюдений(n/2);

p_{Me} – частота медианного класса;

i_2 – число вариантов, или сумма накоплений частот по всем классам, предшествующим медианному классу.

Обработку вариационного ряда для вычисления медианы осуществляют приемом, называемым методом «накопленных частот». Ряд накопленных частот составляют путем последовательного сложения частоты каждого последующего класса с частотами предыдущего.

Разберем вычисление медианы на примере, приведенному при определении моды вариационного ряда массы байкальского омуля возраста 7+:

Классы варьи-рующего признака	235-251	252-268	269-286	287-303	304-320	321-337	338-354	355-371	372-388	389-405	Ито-го
Частота, p	··2	0	···3	··2	···3	·····5	···3	···3	··2	·1	n=24
Накопи-тельные частоты	2	2	5	7	10	15	18	21	23	24	-

В первом классе ряда частота равна 2. В следующем классе она составлена из суммирования частот первого и второго классов (2+0=2), накопленные частоты третьего класса получают из их суммирования для второго и третьего (2+0+3=5) и т.д. до последнего класса.

Сумма накопленных частот в последнем классе должна быть равна общему числу наблюдений данного ряда $\sum p$ или n. В нашем примере в последнем классе эта сумма будет равна 24.

Исходя из ряда накопленных частот, можно найти данные, необходимые для формулы медианы.

Медианным является класс, в котором накопленные частоты составляют половину всех наблюдений (либо наиболее близкий к этому значению, но не меньше i_1).

$$i_1 = \frac{n}{2} = \frac{24}{2} = 12,$$

Следовательно, в таблице медианным классом будет класс (шестой слева), имеющий 15 накопительных частот с границами признака 321-337. V_{Me}, соответствующее нижней границе, или началу медианного класса, будет равно 321.

Значение i_2равно сумме накопительных частот по классам, предшествующему медианному, и соответствует накопительным частотам, проставленным в пятом классе, то есть 10. P_{Me} – частота медианного класса, равна 15. Величина класса - К=17

Подставим в формулу медианы все необходимые данные и вычислим ее значение:

$$Me = V_{Me} + K\frac{i_1 - i_2}{p_{Me}} = 321 + 17 \cdot \frac{12 + 10}{15} = 321 + 17 \cdot \frac{22}{15}$$
$$= 321 + 17 \cdot 1{,}467 = 321 + 24{,}93 = 345{,}93 \text{ (г)}$$

Медиану используют вместо средней арифметической, когда крайние варианты ранжированного ряда (наименьшая и наибольшая) по сравнению с остальными оказываются чрезмерно большими или чрезмерно малыми.

Показатели разнообразия (определение степени изменчивости варьирующего признака)

Первой характеристикой совокупности объектов, вошедших в выборку, как уже известно, служат средние величины. Но этих показателей совершенно недостаточно для суждения о свойствах совокупности по изучаемому признаку, так как всякая группа состоит из неодинаковых особей, отличающихся друг от друга.

Вторым существенным показателем любой совокупности является мера изменчивости (вариабельности) значений данного признака между особями, составляющими совокупность. Вариабельность демонстрирует степень многообразия значений, вариантов характеристик исследуемых объектов и может определяться как в целом для выборки, так и для какой-либо ее части.

Выявление степени изменчивости признаков между членами совокупности, а также установление особенностей в характере распределений особей в вариационном ряду достигают особыми методами, разработанными вариационной статистикой.

Основными показателями вариации служат следующие статистические величины:

- Лимит – Lim;
- Дисперсия (варианса) – δ^2;
- Среднее квадратичное отклонение – δ;
- Нормированное отклонение – x или t;
- Коэффициент изменчивости – C_V или V.

Лимит Lim

Лимит характеризует разнообразие изучаемого признака только по двум крайним вариантам без учета их распределения в совокупности, то есть игнорируя ее внутреннюю структуру. Эта характеристика является неточной и применяется для быстрой, ориентировочной оценки. Осуществляется определением изменчивости в сопоставлении максимального и минимального значения варьирующего признака у членов совокупности, то есть $v_{макс}$и $v_{мин}$. Чем больше величина лимита, то есть разница между максимальными и минимальными значениями, тем больше вариабельность признака. Лимиты показывают размах значений и тем самым характеризуют разнообразие признака в группе.

Например: имеется два вариационных ряда характеризующих массу окуня возраста 5+(г) с равным количеством частот (n=12):

1-й ряд: 48,67, 98, 111, 118, 129, 144, 153, 155, 160, 202, 206

$$M_1 = \frac{\sum v}{n} = \frac{1591}{12} = 132{,}6 \text{ (г)}$$

2-й ряд: 46, 69, 89, 94, 95,97, 153, 166, 173, 180, 212, 217

$$M_2 = \frac{\sum v}{n} = \frac{1591}{12} = 132{,}6 \text{ (г)}$$

Средняя масса окуня у обоих рядов одинакова, но в первом разнообразие по этому признаку гораздо меньше, чем на втором.

В приведенном примере:

$$Lim_1= 48\text{-}206$$

$$Lim_2 = 46\text{-}217$$

Величину лимита определяют всегда при обработке выборки, несмотря на то, что он является упрощенным показателем изменчивости. Она может быть использована для статистического анализа даже при отсутствие конкретного вариационного ряда. Но лимиты могут служить только грубым показателем изменчивости и не выявляют ее вполне правильно. Могут быть две совокупности, у которых лимит, и средняя арифметическая будут одинаковы, а истинная изменчивость, выраженная более точными методами, окажется различной.

Дисперсия, или варианса δ^2

В переводе с латинского языка дисперсия переводится как «рассеяние», что можно трактовать как небольшое отклонение, разброс от среднего уровня. Она указывает на степень разнообразия показателей у членов совокупности, то есть это мера разброса значений случайной величины относительно ее математического ожидания.

Если представить совокупность, состоящую из одновозрастных (возраста 2+) особей плотвы, то изменчивость их по массе может быть выражена путем сопоставления массы каждой рыбины с величиной, характеризующей среднюю арифметическую этого показателя в данной совокупности, то есть изменчивость выражается отклонением конкретного значения (v)от средней арифметической (M).

Из данных, характеризующих массу шести экземпляров плотвы, можно составить простой вариационный ряд, который даст значения вариантов и их отклонений от средней арифметической:

Показатели	Плотва						
	№1	№2	№3	№4	№5	№6	
Варианты массы рядаv (г)	44	51	33	21	64	56	$M = \frac{\sum v}{n} = \frac{269}{6} = 44{,}8 \approx 45$
v-M (г)	44-45 =-1	51-45 =6	33-45 =-12	21-45 =-24	64-45 =19	56-45 =11	∑v-M =0
$(v-M)^2$	1	36	144	576	361	121	$\sum(v-M)^2$ =1239

Так как ∑(v-M) всегда рана нулю, что является свойством средней арифметической, то для определения степени изменчивости у членов совокупности нашего примера необходимо возвести каждое значение (v-M) в квадрат, а затем просуммировать и получить величину $\sum(v-M)^2$, что и сделано в последнем столбце таблицы.

По второму свойству средней арифметической эта величина будет наименьшей по сравнению с любой другой величиной, взятой вместо средней арифметической (М) в этом выражении. Поэтому ее и используют для измерения изменчивости. Если затем разделить $\sum(v\text{-}M)^2$ на число наблюдений (n), то получим величину дисперсии, выражающую изменчивость признака в данной совокупности. Таким образом, формула дисперсии будет выглядеть следующим образом:

$$\delta^2 = \frac{\sum(v\text{-}M)^2}{n},$$

где δ^2 – дисперсия;

М – средняя арифметическая;

v– результат измерения признака у каждого объекта;

n – число особей в группе или число наблюдений.

Это означает, что дисперсия, измеряющая изменчивость признака, выражает его через средний квадрат отклонения каждого члена совокупности от средней арифметической данного признака.

В нашем примере:

$$\delta^2 = \frac{\sum(v\text{-}M)^2}{n} = \frac{1239}{6} = 206{,}5 \text{ (г)}$$

Дисперсия имеет большое значение в работах по углубленному анализу изменчивости признака с помощью статистического метода «Дисперсионный анализ».

Среднее квадратичное отклонение δ

Среднее квадратичное отклонение служит основным способом измерения изменчивости. Оно может быть получено из значения дисперсии, если из нее извлечь квадратный корень.

Формула среднего квадратичного отклонения:

$$\delta = \sqrt{\frac{\sum(v\text{-}M)^2}{n}},$$

где δ – среднее квадратичное отклонение;

М – средняя арифметическая;

v – результат измерения признака у каждого объекта;

n – число особей в группе или число наблюдений.

Для определения среднего квадратичного отклонения необходимо извлечь квадратный корень из суммы квадратов отклонений исследуемых величин и их среднего арифметического поделенных на число вариантов в выборке.

При малом числе наблюдений (n<30), эта формула несколько изменяется в знаменателе и приобретает следующий вид:

$$\delta = \sqrt{\frac{\sum(v-M)^2}{n-1}}$$

Среднее квадратичное отклонение – наиболее распространенная статистическая величина для измерения изменчивости как количественных, так и качественных признаков членов совокупности. Показывает, насколько в среднем каждый вариант отклоняется от средней арифметической, вычисленной для данной совокупности. Чем больше значение δ, тем больше изменчивость данного признака в совокупности.

Среднее квадратичное отклонение – величина именованная, она имеет то же именование, что и единица измерения изучаемого признака (кг, см, шт. и т.п.).

В так называемых нормальных вариационных рядах весь размах изменчивости, ограниченный максимальным и минимальным значением варьирующего признака, заключает в себе шестикратную величину среднего квадратичного отклонения.

При этом максимальный вариант отстоит от средней арифметической на значение +3δ, а минимальный вариант – на значение -3δ.Поэтому принято весь размах изменчивости выражать записью:M=±3δ.На основании этой особенности изменчивости в нормальных рядах можно осуществлять некоторые расчеты. Например, по показателю средней арифметической и значению δможно рассчитать каковы будут значения $v_{макс}$ и $v_{мин}$ в данной совокупности.

Пример: средний диаметр икринок леща составил 1,11 мм, а изменчивость выражается δ=0,22. Исходя из этих данных, можно предположить, что наибольший размер составляет:

$$M+3\delta = 1{,}11+3\cdot0{,}22= 1{,}77 \text{ (мм)}$$

Наименьший соответственно:

$$M-3\delta = 1{,}11-3\cdot0{,}22= 0{,}45 \text{ (мм)}$$

Следовательно, весь размах изменчивости выражается:

$$\text{Lim}0{,}45-1{,}77 = 1{,}32 \text{ (мм)}$$

По размаху изменчивости можно произвести обратное действие и рассчитать приблизительное значение среднего квадратичного отклонения.

Зная, что в пределах лимита содержится шесть δ, вычисляется значение одной δ. В нашем примере это даст следующее:

$$\delta = \frac{\mathrm{Lim}}{6} = \frac{1{,}77 - 0{,}45}{6} = 0{,}22\ (\text{мм})$$

Такую систему расчета для грубого определения среднего квадратичного отклонения целесообразно делать в тех случаях, когда никаких данных о вариационном ряде нет, а исследователь предполагает возможный размах изменчивости признака, и ему необходимо для ряда статистических вычислений иметь хотя бы ориентировочное значение δ.

Вычисление среднего квадратичного отклонения (δ) для больших выборок осуществляется путем обработки вариационного ряда, разбитого на классы.

Рабочая формула среднего квадратичного отклонения при большом числе наблюдений:

$$\delta = \mathrm{K} \cdot \sqrt{\frac{\sum \mathrm{p} \cdot \mathrm{a}^2}{\mathrm{n}} - (\frac{\sum \mathrm{p} \cdot \mathrm{a}}{\mathrm{n}})^2},$$

где δ – среднее квадратичное отклонение;

К – величина класса;

р – частоты вариационного ряда;

n = ∑р – число наблюдений в данном вариационном ряду или объем данной совокупности;

а – условное отклонение каждого класса от класса, в котором находится условная средняя (а), выраженное числом классов.

Для получения величин, входящих в эту формулу, применяют обработку вариационного ряда методом произведений.

Разберем технику вычисления δ на примере которых уже был использован при вычислении средней арифметической длины туловища (промысловой длины) омуля байкальского возраста 5+ оз. Байкал (см).

Приведем данные таблицы, в которой имелись столбцы со значениями классов (v), частот (р), условных отклонений (а), произведения частот на соответствующее условное отклонение (р·а). Дополним обработку данных введением столбца, в котором проставлены для каждого класса значения $р·а^2$, чтобы его получить для каждого класса следует уже имеющиеся значения р·а умножить на условное отклонение (а) этого класса:

Классы	Частоты, p	Условное отклонение, a	p·a	p·a²
21,5-21,9	3	-4	-12	48
22,0-22,4	1	-3	-3	9
22,5-22,9	4	-2	-8	16
23,0-23,4	1	-1	-1	1
23,5-23,9	6	0	0	0
24,0-24,4	3	1	3	3
24,5-24,9	4	2	8	16
25,0-25,4	4	3	12	36
25,5-25,9	5	4	20	80
26,0-26,4	3	5	15	75
26,5-26,9	1	6	6	36
	∑p=35		∑p·a=40	∑p·a²=320

Полученные данные подставим в формулу:

$$\delta = K \cdot \sqrt{\frac{\sum p \cdot a^2}{n} - \left(\frac{\sum p \cdot a}{n}\right)^2} = 0{,}5 \cdot \sqrt{\frac{320}{35} - \left(\frac{40}{35}\right)^2} = 0{,}5 \cdot \sqrt{9{,}14 - 1{,}30}$$

$$= 0{,}5 \cdot \sqrt{7{,}84} = 0{,}5 \cdot 2{,}80 = 1{,}40 \text{ (см)}$$

Среднее квадратичное отклонение – хороший показатель вариабельности признака, но эта величина именованная и зависит не только от степени варьирования, но и от единицы измерения средней арифметической, поэтому по ней можно сравнивать изменчивость лишь одних и тех же показателей, а сопоставлять вариативность разных признаков нельзя.

Нормированное отклонение (x или t)

Нормированное отклонение – статистический признак, позволяющий определить изменчивость и представляющий отклонение той или иной варианты от средней величины. С его помощью можно выразить в относительных единицах (долях среднего квадратичного отклонения) уровень отличия каждого конкретного члена совокупности от средней арифметической.

Формула нормированного отклонения:

$$x = \frac{v - M}{\delta},$$

где x – нормированное отклонение;

δ – среднее квадратичное отклонение;

M – средняя арифметическая;

v – результат измерения признака у каждого объекта.

Чем больше нормированное отклонение, тем дальше от средней арифметической отстоит величина показателя изучаемого признака. Знак у нормированного отклонения показывает, в какую сторону от средней арифметической отклоняется данная варианта, то есть будет ли она меньше (-x) или больше (+x), чем средняя арифметическая.

Вернемся к данным о средней длине туловища омуля байкальского возраста 5+ оз. Байкал (см): 25,9; 21,5; 23,9; 22,8; 21,6; 24,9; 26,0; 25,1; 23,6; 25,5; 22,9; 23,8; 26,5; 24,7; 26,3; 25,6; 25,9; 24,8; 24,2; 23,6; 24,7; 25,6; 25,2; 23,9; 25,4; 22,5; 21,7; 26,1; 23,0; 25,4; 23,6; 24,3; 24,0; 22,9; 22,1.

В результате проведенных расчетов средняя арифметическая составила: M = 24,27 см, среднее квадратичное отклонение - δ = 1,40 см. Определим нормированное отклонение первого показателя представленного вариационного ряда (25,9 см):

$$x= \frac{v\text{-}M}{\delta} = \frac{25{,}9 - 24{,}27}{1{,}40} = \frac{1{,}63}{1{,}4} = 1{,}16$$

С помощью нормированного отклонения можно оценить любое полученное значение по отношению к группе в целом, взвесить его и одновременно освободиться от именованных чисел. Нормированное отклонение можно использовать для сравнительной оценки индивидов по одному и тому же признаку. Оно помогает определять так называемые «выскакивающие» варианты и решать вопрос о возможности их отбрасывания как артефактов (исключать из дальнейшей обработки).

Решение о выбраковке резко выделяющихся значений (эти отклонения могли возникнуть в результате неточности измерений, ошибок внимания, методических погрешностей и т.д.) принимается на основании нормирования сомнительных вариант по отношению к их средней арифметической. Этой цели служит формула критерия выпада:

$$T = \frac{v - M}{\delta} \geq T_{st},$$

где T – критерий выпада;

δ – среднее квадратичное отклонение;

M – средняя арифметическая;

V – результат измерения признака у каждого объекта;

T_{st} – стандартные значения критериев выпадов.

Определим критерии выпада для минимального и максимального показателей длины туловища омуля байкальского возраста 5+ оз. Байкал: min. = 21,5 см, max. = 26,5 см, M = 24,27 см, δ = 1,40 см.

$$T_{min} = \frac{v - M}{\delta} = \frac{21{,}5 - 24{,}27}{1{,}40} = \frac{-2{,}77}{1{,}4} = -1{,}98$$

$$T_{max} = \frac{v - M}{\delta} = \frac{26{,}5 - 24{,}27}{1{,}40} = \frac{2{,}23}{1{,}4} = 1{,}59$$

Сравним полученные результаты со стандартными значениями критерия выпада. Величина табличного показателя зависит от числа показателей вариационном ряду.

Значение критерия выпада (T_{st}) для отбраковки «выскакивающих» вариант

n	T_{st}	n	T_{st}	n	T_{st}	n	T_{st}	n	T_{st}
5	3,04	11	2,33	17	2,18	35-39	2,06	80-89	2,00
6	2,78	12	2,29	18	2,17	40-44	2,05	90-99	2,00
7	2,62	13	2,26	19	2,16	45-49	2,04	100	1,99
8	2,51	14	2,24	20-24	2,15	50-59	2,03		
9	2,43	15	2,22	25-29	2,11	60-69	2,02		
10	2,37	16	2,20	30-34	2,08	70-79	2,01		

T_{st}, прип = 35 – 2,06

$$T_{min} = -1{,}98 < Tst$$

$$T_{max} = 1{,}59 < Tst$$

Следовательно, оба показателя можно считать невыпадающими, и они должны быть исключены из обработки данных.

Коэффициент изменчивости C_V или V

Методы определения степени варьирования с помощью лимитов и среднего квадратичного отклонения имеют один недостаток: они дают показатель изменчивости признака в именованных величинах, а не в относительных. Вследствие этого сопоставление разноименных признаков по величине изменчивости с их помощью невозможно. Коэффициент изменчивости или вариации C_V показывает вариативность признака в совокупности в относительных величинах (в процентах). В связи с этим использовать коэффициент вариации целесообразно в тех случаях, когда необходимо сравнить изменчивость разноименных признаков.

Формула коэффициента изменчивости:

$$C_v = \frac{\delta}{M} \cdot 100\%,$$

где C_v – коэффициент изменчивости;

δ – среднее квадратичное отклонение;

M – средняя арифметическая.

Из формулы видно, что C_v получается путем процентного выражения среднего квадратичного отклонения δ от своей средней арифметической M.

Так в примере по определению средней длины туловища омуля байкальского возраста 5+ оз. Байкал средняя арифметическая M = 24,27 см, а изменчивость, выраженная через среднее квадратичное отклонение δ = 1,40 см. Отсюда коэффициент вариации будет равен:

$$C_v = \frac{1{,}40}{24{,}27} \cdot 100 = 5{,}77\%$$

Чем больше значение коэффициента вариации, тем больше изменчивость признака у членов совокупности. Ориентировочно считают, что если $C_v < 5\%$ – изменчивость низкая, при C_v от 5 до 10% – средняя и при $C_v > 10\%$ – высокая.

Следовательно, можно констатировать, что изменчивость длины туловища омуля байкальского возраста 5+ оз. Байкал в вариационном ряду из 35 измерений оценивается, как средняя.

Коэффициент изменчивости для альтернативных признаков не вычисляют, так как его заменяет значение δ, выраженное в процентах.

Следует иметь в виду следующие особенности коэффициента изменчивости:

1. Величина коэффициента изменчивости не должна рассматриваться в отрыве от средней арифметической и стандартного отклонения.

Например, при близких значениях коэффициента изменчивости, полученных на двух выборках, абсолютные величины средней арифметической и среднего квадратичного отклонения для этих совокупностей могут быть совершенно на разных уровнях. Следовательно, одинаковая или близкая величина коэффициента изменчивости для двух выборок еще не означает, что они качественно близки.

2. Одинаковые величины коэффициента изменчивости двух выборок могут быть результатом разных причин, а именно C_v может увеличиваться за

счет повышенного числителя, то есть более высокой величины среднего квадратичного отклонения одной выборки, или за счет уменьшенного знаменателя, то есть средней арифметической. Эти особенности необходимо учитывать, чтобы не допустить ошибку и не сделать неправильные выводы, беря величину C_v вне связи с величинами М и δ.

3. Коэффициент изменчивости целесообразно использовать при изучении динамических рядов.

Например, в исследованиях, когда изучают показатели, характеризующие возрастные особенности животных, использование C_v совместно со средней арифметической и средним квадратичным отклонением дает ясное представление о динамике и закономерностях онтогенеза по тому или иному признаку.

4. Коэффициент изменчивости имеет большое значение при планировании объема опыта, так как это позволяет получать достоверные статистические параметры.

Типы вариационных рядов и их графическое изображение

Вариационный ряд позволяет проводить группировку данных по классам, показывающим закономерности варьирования признаков, и вычислять ряд статистических величин. Выше уже разбирался вопрос о способах составления вариационных рядов, служащих для выявления вариабельности признака у членов случайной выборки. Выборки имеют ограниченное, хотя и большое число наблюдений и служат способом изучения свойств генеральной совокупности.

По типу распределения частот вариационные ряды довольно сильно могут отличаться между собой. Их подразделяют на эмпирические и теоретические.

Эмпирическими являются ряды, составленные на основании конкретных данных, полученных из выборочной совокупности. Количество членов в них может быть малым и большим, но не достигающим бесконечно большого числа. Распределение членов выборки по классам выражается частотами (р) или частостями (р'), представляющими собой дробь, выраженную в долях единицы встречаемости членов выборки в том или ином классе вариационного ряда. Сумма частостей по всем классам ряда составляет единицу. Вычисляют ее путем деления частот класса на общий объем выборки.

$$p' = \frac{p}{n}$$

Теоретические вариационные ряды отражают закономерности распределения членов совокупности по классам варьирующего признака при бесконечно большом числе наблюдений (когда $n \to \infty$).То есть теоретический ряд служит пределом, к которому стремится эмпирический при увеличение объема выборки до бесконечности.

В теоретических рядах встречаемость членов совокупности, отклоняющихся на определенную величину от средней арифметической, выражается вероятностью (Р), а не частотой или частостью. При увеличении числа наблюдений частость стремиться к вероятности. Следовательно, вероятность служит мерой возможности появления объектов с данным отклонением от средней арифметической или мерой возможности появления какого-либо события. Например, возможность появления мутантов в популяции выражается вероятностью. Рождение особей мужского или женского пола также имеет определенную вероятность.

Вероятность представляет собой дробь, величина которой находится в границах от 0 до 1.Если вероятность события равна 0, то его осуществление не произойдет, если равна 1, то это событие с необходимостью будет осуществлено. При уровне вероятности больше 0,5, его реализация более вероятна, чем неосуществима. При вероятности случая меньше 0,5, его называют маловероятным. Сумма вероятностей двух противоположных фактов равна единице.

Пример: если из 1000 икринок вылупляется 550 мальков женского пола, то вероятность составляет:

$$P = \frac{550}{1000} = 0{,}55$$

Возможность противоположного действия, то есть рождения особей мужского пола составляет:

$$Q = 1 – P = 1\text{-}0{,}55 = 0{,}45.$$

Вариационные ряды бывают дискретными и интервальными.

Дискретным вариационным рядом распределения называют ранжированную (в порядке возрастания или убывания) совокупность вариантов с соответствующими им частотами (весами) или частостями. Их строят в том случае, если значения изучаемого признака отличается друг от друга не менее чем на некоторую конечную величину. В дискретных вариационных рядах задаются точечные значения признака. Графически его можно представить с помощью полигона распределения частот или частостей.

Интервальным вариационным рядом называют упорядоченную совокупность интервалов варьирования значений случайной величины с соответствующими частотами или частостями попаданий в каждый из них значений величины. Их создают в случае, если значения изучаемого признака отличаются друг от друга на сколь угодно малую величину. Значения признака в них задаются в виде интервалов, в которых выделяют верхнюю и нижнюю границы. Разность между ними называют интервальной разностью или размахом вариации. Интервальные вариационные ряды графически представляют с помощью гистограммы.

Распределение частот по классам варьирующего признака имеет несколько основных типов, для каждого из которых характерны свои особенности и закономерности. Основные типы распределения следующие:

- нормальное;
- биноминальное;
- асимметричное;
- эксцессивное;
- трансгрессивное;
- Пуассона.

Техника изображения вариационных рядов

Каждый вариационный ряд, составленный достаточно большим числом наблюдений может быть изображен в виде графика или диаграммы. При их построении необходимо соблюдать следующие требования:

- данные на графике должны размещаться слева направо или снизу вверх;
- шкалы на диаграммах должны быть снабжены указателями размеров;
- изображенные графически величины должны иметь цифровые обозначения на самом графике или в прилагаемой к нему таблице;
- геометрические знаки, фигуры, краски, штриховки должны быть пояснены;
- каждый график должен иметь четкое, ясное, по возможности краткое название, отражающее его содержание.

Графическое изображение вариационного ряда, разбитого на классы, называют гистограммой, это способ представления статистических данных в виде столбчатой диаграммы, позволяющий зрительно оценить закон распределения величины разброса данных. Ее строят в осях координат: на оси *x* откладывают классы варьирующего признака (размеры классов), по оси *y* – значения частот или их процентное выражение от общего числа наблюдений. Над каждым классом вычерчиваютстолбик, высота которого соответствует частотам или проценту частот класса отложенным по масштабу оси y. Гистограмму можно превратить в вариационную кривую. Для этого надо соединить прямыми линиями середины классов.

Все используемые в биометрии диаграммы по способу построения можно подразделить на:

- линейные или графические;
- плоскостные;
- объемные;
- поверхностные;
- фигурные.

Линейные диаграммы применяют при изучении различных связей между факторами и характеристики изменений этих факторов во времени. Они строятся в прямоугольной системе координат: горизонтальной (оси абсцисс – ось *x*) и вертикальной (оси ординат – ось *y)*. Точка пересечения осей служит началом отсчета.

На оси абсцисс в избранном масштабе откладывается время или другие факторные признаки, затем из точек, соответствующих определенным моментам или периодам времени, восстанавливают ординаты, отражающие размеры изучаемого результативного признака и их вершины соединяют прямыми линиями (рис.7).

На одном графике может быть одновременно построено несколько линейных диаграмм, что позволяет производить их наглядное сравнение (не рекомендуется строить более 4 диаграмм, так как это затрудняет восприятие).

Одним из видов плоскостной диаграммы является точечная, она показывает отношения между численными значениями в нескольких рядах данных или отображает две группы чисел как один ряд координат *x* и *y*, их используются для представления и сравнения числовых значений. Точечные диаграммы и графики очень похожи друг на друга, особенно если точечная диаграмма отображается с соединительными линиями. Однако между ними есть существенное различие. Точечная диаграмма имеет две оси значений, при этом один набор выводится вдоль горизонтальной оси (оси *x*), а другой - вдоль вертикальной (оси *y*). Эти значения объединяются в одну точку данных и выводятся с неравными интервалами, или кластерами (рис. 8).

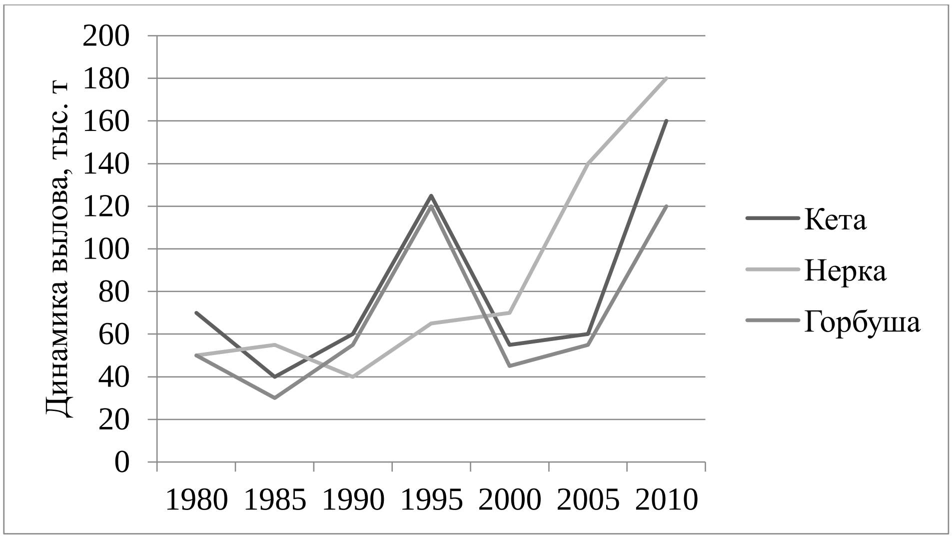


Рисунок 7 – Линейная диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

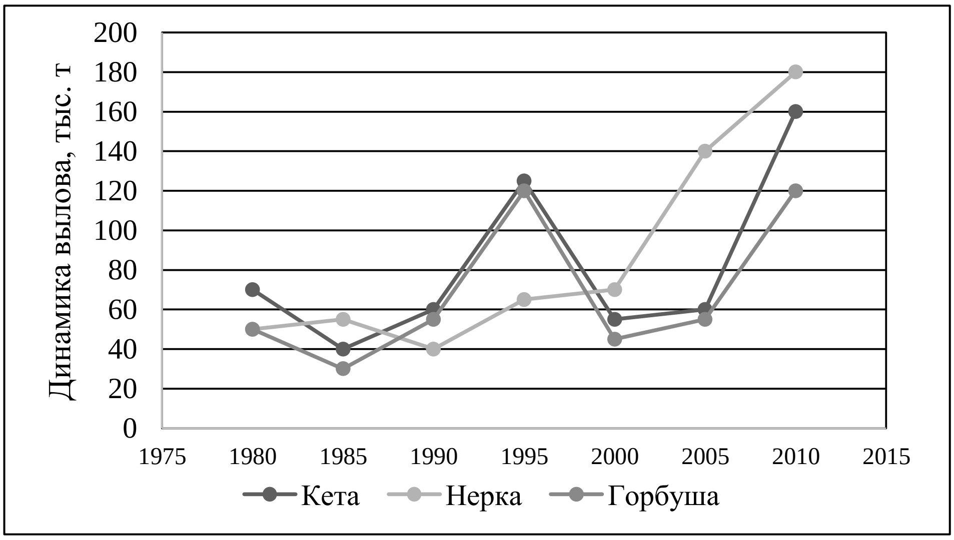


Рисунок 8 - Точечная диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

Так же разновидностью линейных диаграмм являются радиальные или лепестковые(диаграммы в системе полярных координат). Их используют для изображения значений или факторов, имеющих замкнутый циклический характер. Количество осей соответствует количеству частей, на которые разделен период времени. За длину радиуса окружности принимают среднюю величину, затем на каждой оси откладывают величину, соответствующую уровню значения. Полученные точки соединяют прямыми линиями (рис.9).

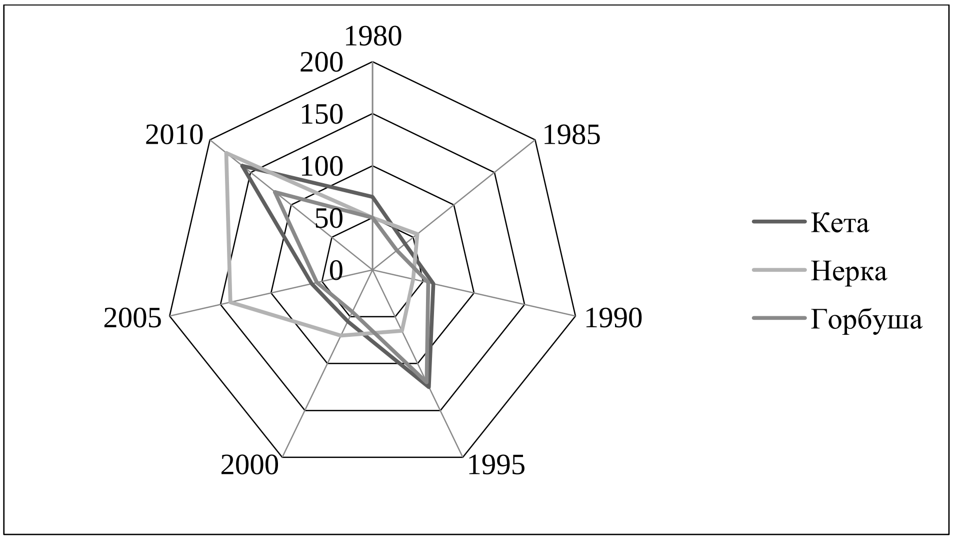


Рисунок 9 – Радиальная диаграмма динамики вылова рыб Камчатки за 1980-2010 гг., тыс. т.

Плоскостные диаграммы делятся на:

- столбиковые;
- секторные;
- внутристолбиковые.

Столбиковые диаграммы строят по такому же принципу, что и динамические кривые, но в них вертикально или горизонтально проводимым линиям соответствуют прямоугольники. Эти диаграммы особенно удобны для иллюстрации изменения величина в какой-либо определенный промежуток времени. Каждый столбик изображает величину отдельного уровня исследуемого статистического ряда. Размещение столбиков в поле графика может быть различным:

- на одинаковом расстоянии друг от друга;
- вплотную друг к другу;
- в частном наложении друг на друга.

При построении столбиковых диаграмм на горизонтальной оси располагают основания столбиков, величину основания выбирают произвольно, но она должна быть одинаковой для всех. Шкала, определяющая масштаб столбиков по высоте, расположена по вертикальной оси. Величина каждого столбика по вертикали соответствует размеру изображаемого на графике статистического показателя. У всех столбиков переменной величиной является только одно измерение (рис.10).

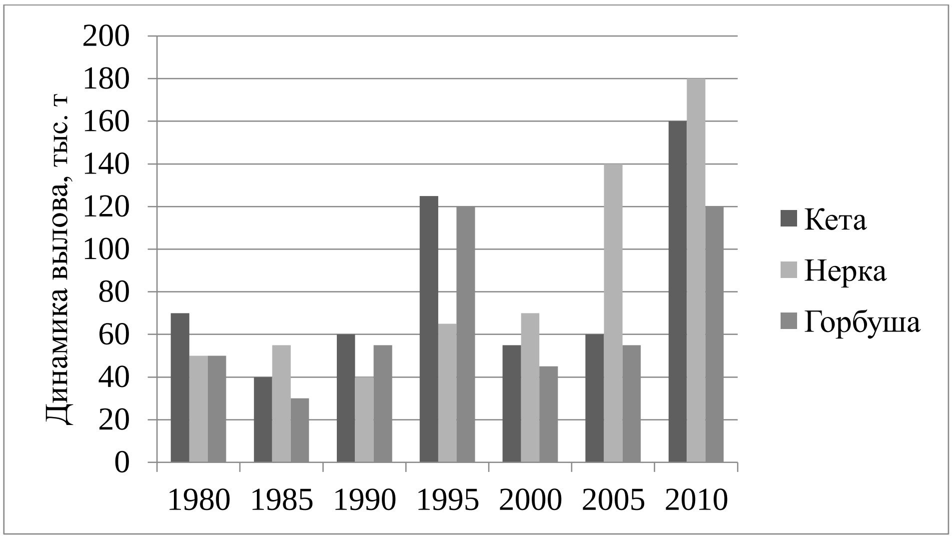


Рисунок 10 – Столбиковая диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

Секторные или круговые диаграммы представляют собой круг, который принимается за целое, а его отдельные секторы соответствуют частям изображаемого показателя. Такие диаграммы применяются для иллюстрации экстенсивных показателей и показывают вклад каждого значения в общую сумму (рис. 11).

Внутристолбиковые (полосовые, сложностолбиковые, ленточные) диаграммы представляют собой прямоугольник или квадрат, разделенный на части. При этом длина лент (столбиков) принимается за 100%, а их составные части соответствуют долям явления в процентах. Бывают двух видов:

1. Гистограмма с накоплением, которую применяют для выделения общей суммы по ряду в одной категории.

2. Нормированная гистограмма с накоплением – используют для показа относительной величины каждого ряда данных.

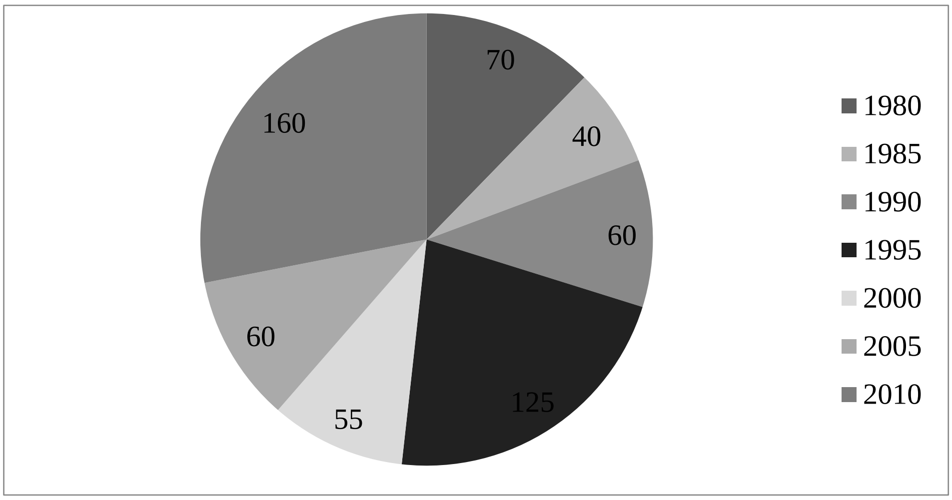


Рисунок 11 – Секторная диаграмма динамики вылова кеты на Камчатке за 1980-2010 гг., тыс.т.

Внутристолбиковые диаграммы используются сравнения структуры какого-либо фактора или значения (рис. 12).

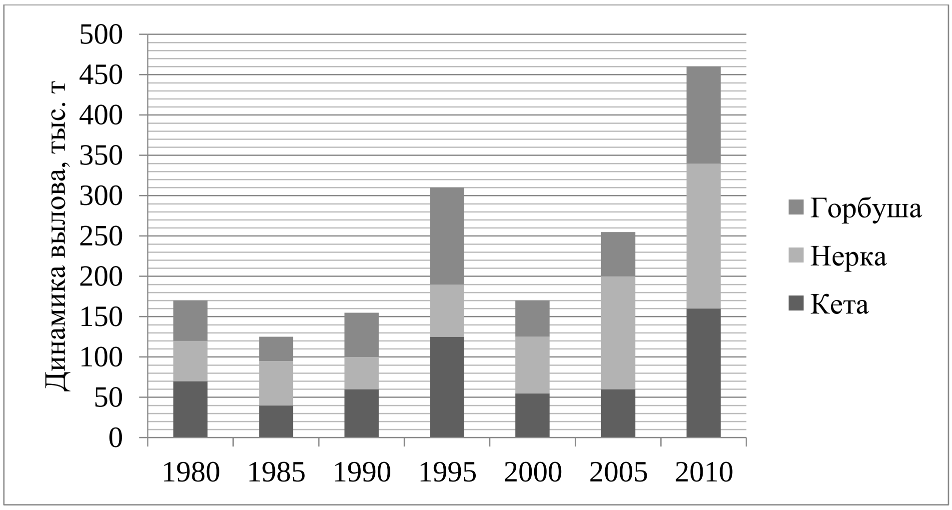


Рисунок 12 – Внутристолбиковая диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

При построении объемных диаграмм статистические данные изображают в виде геометрических фигур в трех измерениях (рис.13).

В том случае, когда требуется найти оптимальные комбинации в двух наборах данных, а показатели расположены в столбцах или строках, их можно изобразить в виде поверхностной диаграммы. Цвета выделяют зоны одинаковых диапазонов значений. Также можно их использовать для иллюстрации категорий и наборов данных, представляющих собой числовые значения (рис. 14).

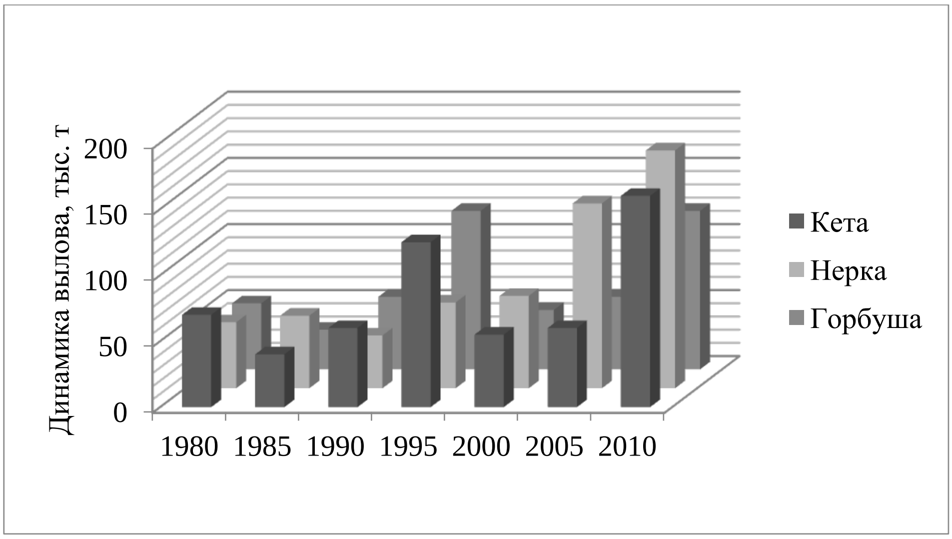


Рисунок 13 – Объемная диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

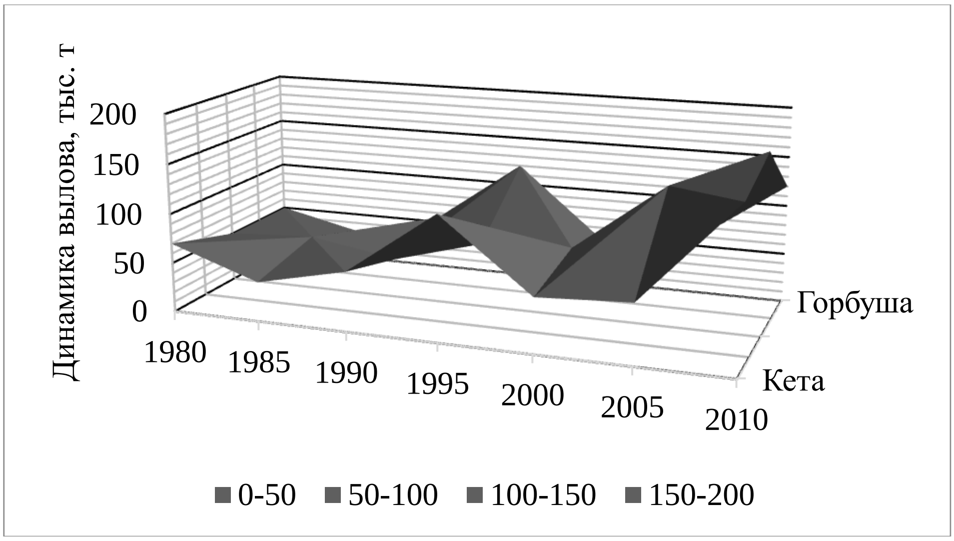
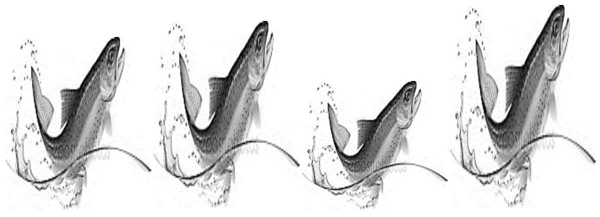


Рисунок 14 – Поверхностная диаграмма динамики вылова рыб Камчатки за 1980-2010 гг.

Фигурные диаграммы содержат соотношения определенных статистических величин, которые представлены в условном виде как определенные художественные фигуры. Их используют, обычно для наглядности и иллюстрации показателей. При этом обычно пользуются округленными цифровыми данными

Фигурные диаграммы строятся двумя методами:

1. Сравниваемые статистические величины изображаются фигурами разных размеров. Для построения диаграммы устанавливают определенный масштаб, а размер фигуры показывает величину показателя, например, изображение одной рыбы соответствует 10 тыс. тонн улова (рис. 15).



1980 г 1990 г 2000 г 2010 г

Рисунок 15 – Фигурная диаграмма динамики лососи на Камчатке за 1980-2010 гг по определенному масштабу

2. Сравниваемые статистические величины изображаются разной численностью фигур одинакового размера (рис.16).

Рисунок 16 – Фигурная диаграмма динамики лососи на Камчатке за 1980-2010 гг. разной численностью фигур одинакового размера

Для графического изображения дискретных вариационных рядов, то есть рядов с ранжированной совокупностью вариантов и соответствующим им частотам применяют полигон распределения. При его построении по оси абсцисс (оси x) откладывают количественные значения варьирующего признака - варианты, а по оси ординат (оси y) – частоты или частости.

Для демонстрации данного приема используем полученные нами результаты при определении моды для вариационного ряда массы байкальского омуля возраста 7+ (рис. 17).

Классы варьи-рующего признака	235-251	252-268	269-286	287-303	304-320	321-337	338-354	355-371	372-388	389-405
Частота, р	2	0	3	2	3	5	3	3	2	1

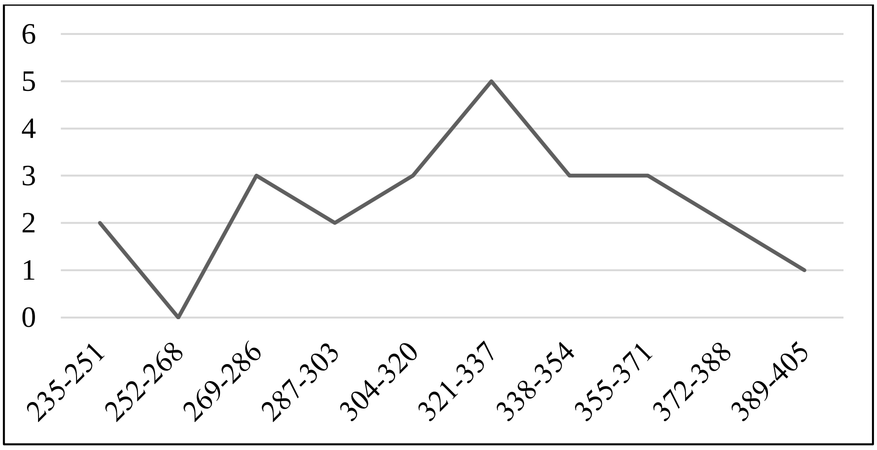


Рисунок 17 –Полигон распределения частот массы байкальского омуля возраста 7+

Распределение признака в вариационном ряду по накопленным частотам (частостям) изображают с помощью кумуляты (кумулятивной кривой). При этом на оси абсцисс помещают значения признака, а на оси ординат – накопленные частоты или частости. Также используем данные массы байкальского омуля возраста 7+, обработанные нами при определении медианы данного вариационного ряда и выстроим в виде кумуляты (рис. 18).

Классы варьирующего признака	235-251	252-268	269-286	287-303	304-320	321-337	338-354	355-371	372-388	389-405
Частота, p	2	0	3	2	3	5	3	3	2	1
Накопительные частоты	2	2	5	7	10	15	18	21	23	24

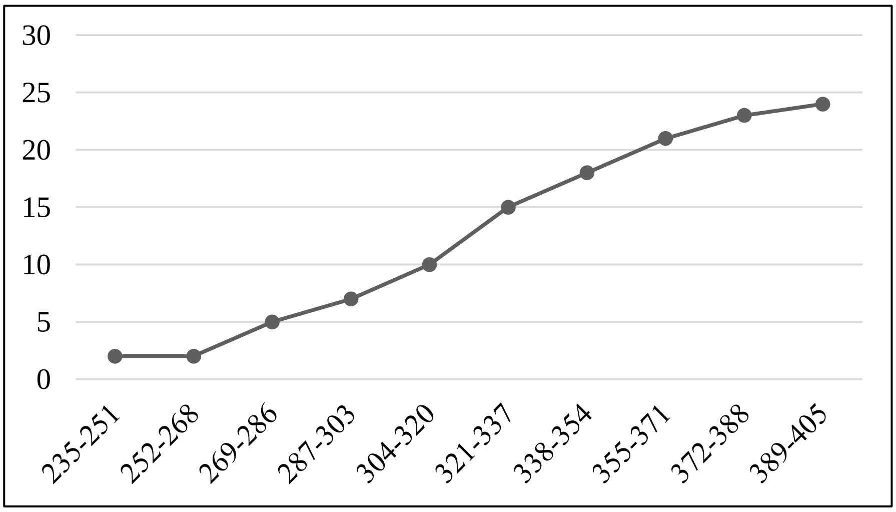


Рисунок 18 – Кумулята распределения накопительных частот массы байкальского омуля возраста 7+

Нормальное распределение и его свойства

Чаще всего распределение частот у биологических объектов характеризуется нормальным распределением. Его открытие связано с именами немецкого математика Иоганна Карла Фридриха Гаусса (1777-1855 гг.) и французского математика Пьера-Симона Лапласа (1949-1827 гг.), у которых оно впервые появилось связи с исследованиями по теории ошибок и методу наименьших квадратов. Поэтому нормальное распределение называют еще распределением Лапласа-Гаусса, или просто распределением Гаусса или Лапласа.

Если число наблюдений в выборке бесконечно большое, то нормальная вариационная кривая абсолютна симметрична, причем крайние значения (наибольшие и наименьшие) появляются редко, но чем ближе значения признака к центру (к средней арифметической), тем оно чаще встречается (рис. 19).

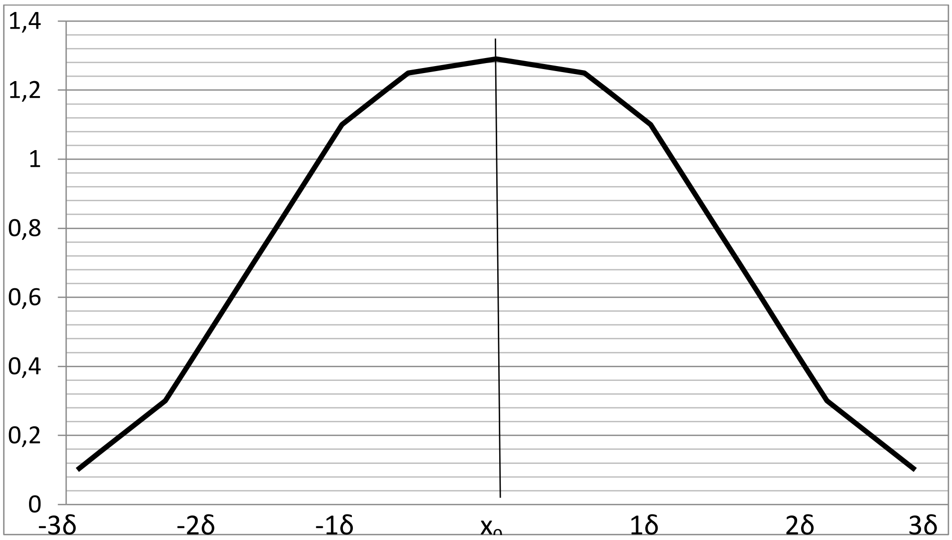


Рисунок 19 – Тип нормальной кривой распределения

Для нормальной кривой характерен ряд закономерностей, проявляющихся в распределении членов совокупности по классам.

Свойства нормальной кривой:

Ось абсцисс (ось *x*) служит основанием кривой, на котором откладываются значения варьирующего признака (v), выраженные в отклонениях его от средней арифметической в долях среднего квадратичного отклонения (δ).

Значение варьирующего признака, которое соответствует средней арифметической (M), берут в качестве начальной точки (x_0); вправо от нее откладывают значения, превышающие M и доходящие до $v_{макс}$, а влево –

значения меньше M, доходящие до $v_{мин}$.Точка x_0 соответствует варианту, величина которого равна средней арифметической, моде и медиане.Размах варьирования признака от $v_{мин}$ до $v_{макс}$ почти полностью ограничен отклонениями v от M на +3δ и -3δ.

Ось ординат (ось *y*) служит для отложения частот (p). Если число наблюдений в совокупности бесконечно большое (n →∞), то значения частот превращаются в значения вероятности (P).

Высота перпендикуляров, восстановленных из значений v, будет определять характер плавности боковых границ нормальной кривой.

Перпендикуляр, восстановленный из варианта $v=x_0=M$ образует вершину нормальной кривой и соответствует максимальной ординате, обозначаемой y_0, которая выражает максимальную частоту и максимальную вероятность для признака, соответствующего средней арифметической, моде и медиане данной совокупности.

Если соединить вершины перпендикуляров, восстановленных из каждого значения варьирующего признака в границах $v_{макс}$ до $v_{мин}$, то образуется линия, показывающя нормальную кривую распределения. Форма нормальной кривой колоколообразная, плавная. Площадь, ограниченная осью *x* и границей линии, образующей кривую, имеет симметричную форму. Если перегнуть рисунок кривой по ординате y_0, то правая и левая части совпадут.

Степень крутизны или округлости боковых ветвей, а так же высота вершины зависят от величины среднего квадратичного отклонения (δ). Поэтому форма различных вариационных кривых неодинакова. При одном и том же числе наблюдений, в завистмости от величины δ, форма нормальной кривой будет разная:

- чем больше значение δ, тем больше будет ее основание и ниже вершина;
- при уменьшении δ основание кривой уменьшается, а вершина (высота) кривой увеличивается.

Площадь, ограниченная нормальной кривой, принимается за 1 или за 100%. Она соответствует общему числу наблюдений выборки. Величина площади, отсеченной ординатами из точек $v_{мин}$ и $v_{мак}$, соответствующих – 3δ и +3δ, составляет 99,7% от всей площади, то есть от всех наблюдений, вошедших в совокупность. Площадь, заключенная между ординатой y_0 (ордината точки M) и какой-либо ординатой, является определенной величиной и может быть найдена по специальным таблицам (таблицы интеграла вероятности (табл. 9 приложения)). Следовательно, если мы знаем, что значение признака отклоняется например от средней арифметической на +2δ, то по таблицам интеграла вероятности можно определить, какое число объектов в нормальном вариационном ряду будет иметь величину признака в границах от М до +2δ.

Пример: средний горизонтальный диаметр глаза 100 выловленных щук М=10,5 мм, среднее квадратичное отклонение δ=0,97 мм. Определить, какое число щук будет иметь горизонтальный диаметр глаза на уровне от 10,5 до 11,0 мм.Для этого находим нормированное отклонение:

$$x= \frac{v\text{-}M}{\delta} = \frac{11{,}0 - 10{,}5}{0{,}97} = 0{,}52$$

То есть горизонтальный диаметр глаза щук, которые превышают среднюю на +0,5δ.

Ищем по таблице это значение нормированного отклонения.

Во втором столбце (показатель площади, ограниченной ординатами y_0 и y), в соответствующей строке (0,5) узнаем, что доля площади в этих границах равна 0,19146от всей площади, принятой за 1, а это соответствует 19,146% от всех наблюдений, вошедших в совокупность.

Площади и ординаты нормальной кривой распределения (из Миллса) (фрагмент)

Нормированное отклонение $x = \frac{v - M}{\delta}$	Вторая функция нормированного оклонения $\varphi(x)$		Первая функция нормированного отклонения f (x) ордината y при значениях $x=\frac{v-M}{\delta}$ То есть вероятность y_x при отклонениях v от М на x
	Площадь между ординатами y_0 и y $\frac{v - M}{\delta}$	% числа наблюдений, заключенных между ординатами y_0 и y	
0,0	0,00000	0	0,39894
0,1	0,03983	3,983	0,39695
0,2	0,07926	7,926	0,39104
0,3	0,11791	11,791	0,38139
0,4	0,15542	15,542	0,36827
0,5	0,19146	19,146	0,35207

Следовательно 19,146≈19 щук будут иметь горизонтальный диаметр глаза в границах от 10,5 до 11,0 мм.

Нормальная кривая характеризуется тем, что любая ее ордината (у), то есть значение частот соответствующих любому отклонению варианта от средней арифметичекой (v-M), может быть определена по уровнению нормальной кривой, которое выглядит следующим образом:

$$y_V = \frac{n}{\delta\sqrt{2\pi}} \cdot e^{-\frac{(v-M)}{2\delta^2}},$$

где y_v – ордината, соответствующая искомой теоретической частоте p конкретного значения варьирующего признака;

M – средняя арифметическая;

n – число наблюдений в выборке;

δ – среднее квадратичное отклонение;

e – основание натуральных логарифмов, равное 2,71828;

π – постоянное число, равное 3,1416

v – величина варьирующего признака, для которого вычисляют теоретическую частоту, или ордината y_v

Это уравнение может быть преобразовано и упрощено, что позволит быстрее найти теоретические частоты для конкретного эмпирического вариационного ряда. Преобразование формулы исходит из того, что выражение степени «е» может быть заменено нормированным отклонением:

$$x = \frac{v\text{-}M}{\delta}$$

Выраженное в долях δ, а выражение $\frac{n}{\delta\sqrt{2\pi}}$можно представить как:

$$\frac{n}{\delta} \cdot \frac{1}{\sqrt{2\pi}}$$

Здесь:

$$\frac{1}{\sqrt{2\pi}} = \frac{1}{\sqrt{2 \cdot 3{,}1416}} = 0{,}39894$$

То есть эта величина постоянна для любого вариационного ряда.

Имея в виду указанные преобразования и подставив в уравнение постоянные величины, мы получаем уравнение нормальной кривой в следующем виде:

$$y_V = \frac{n}{\delta} \cdot 0{,}39894 \cdot 2{,}71828^{-\frac{x^2}{2}}$$

Так как x – это нормированное отклонение $\left(\frac{v-M}{\delta}\right)$, выраженное в долях δ, то подставляя различные значения δ (от 0,1 до 4), можно заранее вычислить выражение, обозначаемое ∫(t), в которое войдут следующие величины:

$$\int (t) = 0{,}39894 \cdot 2{,}71828^{-\frac{x^2}{2}}$$

Величина $\int(t)$ – первая функция нормированного отклонения.

Зная выражения $\int(t)$, легко вычислить теоретичесчкие частоты вариантов с различной величиной нормированного отклонения, так как остается только определить отношение фактического числа наблюдений к значению δ и умножить эти выражения на величину $\int(t)$, которую находят в таблице.

Таким образом, конечное рабочее уравнение, по которым определяют теоретические частоты нормальной кривой, будет выглядеть следующим образом:

$$y_v = \frac{n}{\delta} \cdot \int (t)$$

В этом уравнении δ берут в относительных величинах, то есть без умножения на величину класса К. Если выражать δ в именованных величинах, то формула будет следующей:

$$y_V = \frac{n \cdot K}{\delta} \cdot \int (t)$$

Биноминальное распределение

Рассмотренное выше нормальное распределение характерезует распределение особей по количественным признакам. Но у биологических объектов очень часто встречаются качественные альтернативные признаки, такие как пол (самец или самка), тип наследования (доминантный или рецессивный), состояние здоровья (здоровый или зараженный) и т.п.

При изучении варьирования объектов с альтернативными признаками распределение приобретает иную форму, отличающуюся по внешнему виду кривой и по закономерностям от нормального распределения. Его называют биноминальным распределением, тоесть это частный случай нормального распределения. Оно отражает распределение членов совокупности, имеющих альтернативные признаки и характеризует поведение дискретных признаков (выраженных целыми числами).

Особенности биноминального распределения:

1. При биноминальном распределении имеются только два состояния альтернативного признака:признак присутствует v+ или признак отсутстует v-.

2. Вероятность появления признака в данной совокупности для всех ее членов постоянна и выражается через Р. Вероятность отсутствия данного признака также постоянна и выражается Q =1-Р.

3. Биноминальное распределение образуется распределением частных групп по классам альтернативного признака v+, в которые входят члены данной совокупности.

4. В биноминальном распределении частоты (p) могут быть эмпирическими (полученными из данных конкретного материала) и теоретическими. Теоретические частоты определяют с помощью разложения бинома Ньютона $(a+b)^n$. Коэффициенты разложения бинома составляют теоретические частоты биноминального вариационного ряда по классам альтернативного признака v+. Это означает, что коэффициенты покажут, как часто (то есть 1, 2…n раз) будут встречаться частные группы с отсутствием признака v+ или с присутствием этого признака у членов совокупности. Для получения этих теоретических частот бином Ньютона будет выглядеть так:

$$(\mathrm{P} + \mathrm{Q})^{\mathrm{k}},$$

где P – вероятность появления альтернативного признака v+;

Q - вероятность его отсутствия;

k - число наблюдений в частных группах выборки.

5. Биноминальные ряды характеризуются прирывистостью признака, поэтому кривая бюиноминального ряда имеет ломаную линию.Форма кривой зависит от величины вероятности P и величины k.Если вероятности альтернативного признака равны (P=Q), то есть P=0,5 и Q=0,5, биноминальный ряд симметричен.Если P и Q не равны, то ряд будет скошен (асимметричен). Но если даже P и Q не равны, а k (число наблюдений в частных группах) увеличивается, то асимметрия биноминального распределения уменьшается, и оно приближается к нормальному распределению.

6. Статистическими характеристиками биноминального распределения служат средняя арифметическая (M) и среднее квадратичное отклонение (δ).Если известна вероятность P появления признака v+, то тогда формулы M и δ будут следующие:

$$\mathrm{M} = \mathrm{k} \cdot \mathrm{P}; \;\; \delta = \sqrt{\mathrm{k} \cdot \mathrm{P} \cdot \mathrm{Q}}$$

Распределение Пуассона

Распределение Пуассона относится к тем случаям, когда имеют дело с появлением редких событий при большом числе опытов, то есть когда вероятность появления этого события очень мала, а варьирующий признак имеет распределение дискретного типа, то есть принимает только целые значения:

0,1,2,3 и т.д.

Распределение Пуассона моделирует случайную величину, представляющую собой число событий, произошедших за фиксированное время, при условии, что данные события происходят с некоторой фиксированной средней интенсивностью и независимо друг от друга, то есть когда вероятности альтернатив неравны: $P \neq Q$. В связи с этим ряды Пуассона имеют ясно выраженную асимметрию.

Распределение Пуассона имеет один параметр, обозначаемый символом λ – среднее количество успешных испытаний в заданной области возможных исходов. Количество успешных испытаний пуассоновской случайной величины изменяется от 0 до бесконечности. Распределение Пуассона описывается формулой:

$$P(X) = \frac{e^{-\lambda} \cdot \lambda^{X}}{X!},$$

где $P(X)$ – вероятность X успешных испытаний;

λ – ожидаемое количество успехов;

e – основание натурального логарифма=2,71828;

X – количество успехов в единицу времени;

X! – факториал количества успехов в единицу времени.

Характерной особенностью распределения Пуассона, четко отличающей его от нормального и биноминального, является то, что значение средней арифметической такого ряда (M) совпадает с величиной дисперсии (δ^2) или очень близко к ней и, следовательно, этот ряд имеет одну статистическую характеристику. Поэтому, если при вычислении получаются близкие значения M и δ^2, то это служит основанием считать данный ряд распределением Пуассона.

В биологии многие признаки характеризуются таким распределением: частота многоплодных рождений у одноплодных видов, появление у нормальных популяций экземпляров альбиносных животных, различных уродств, появление различных мутантных форм и т.п.

Асимметричные ряды

Кроме нормальных, биноминальных и Пуассоновых рядов, при обработке материала могут встречаться скошенные – ассиметричные. Для них характерно, что частоты уменьшаются в одну сторону быстрее, чем в другую, а это приводит к смещению вершины кривой в правую или левую сторону от средней арифметической.

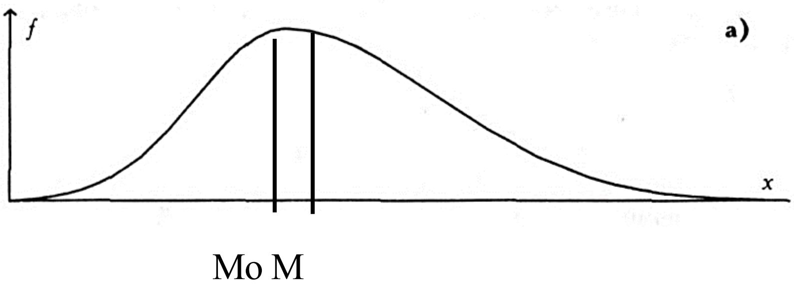
У асимметричных рядов и их кривых средняя арифметическая (M), мода (Mo) и медиана (Me) не совпадают, как у нормальных кривых. Если значение средней арифметической лежит правее моды, то такая асимметрия называется положительной или левосторонней, если M находится левее Mo - отрицательной или правосторонней.

Смещение распределения частот в правую или левую сторону от средней арифметической может быть вызвано следующими обстоятельствами:

1. Выборка сделана неправильно: в нее вошло непропорционально мало частот в левой или правой части ряда. Эта причина является методической ошибкой и в работе и не должна допускаться

2. Сдвиг частот и смещение моды в ту или иную сторону от средней арифметической обусловлен какими-то объективно существующими факторами, которые нарушают обычный для данного признака нормальный характер распределения. То есть, если причина асимметрии не в методической оплошности, то ее появление будет указывать на происходящие качественные сдвиги в изучаемой совокупности.

Асимметрию легко определить визуально по виду полигона или гистограммы распределения. При левосторонней асимметрии относительно центра распределения наблюдается длинная левая ветвь кривой, тогда как при правосторонней асимметрии – ее правая ветвь (рис. 20).



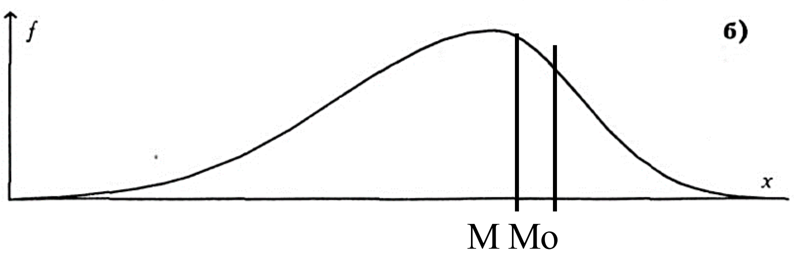


Рисунок 20 – Асимметричное распределение:

а) – левая, положительная асимметрия; б) – правая, отрицательнаяасимметрия

Степень асимметрии вариационного ряда определяется с помощью коэффициента асимметрии As:

$$\mathrm{As} = \frac{\sum(\mathrm{v} - \mathrm{M})^3}{\mathrm{n} \cdot \delta^3},$$

где As – коэффициент асимметрии;

v – значение варьирующего признака;

M – средняя арифметическая;

n – число наблюдений;

δ – среднее квадратичное отклонение.

Если As>0, то асимметрия положительная.

Если As<0, то асимметрия отрицательная.

Если As=0, то ряд нормальный.

Если As имеет значение от 0,25 до 0,5, то это свидетельствует об умеренной асимметрии (косости).

Эксцессивные ряды

Эксцессивными вариационными рядами называют ряды, у которых значительная доля частот накапливается около варианта, соответствующего средней арифметической. Модальный вариант при этом имеет значительно большее число наблюдений, чем в нормальных распределениях. Это приводит на графике к высоковершинности и островершинности. Общий вид такой кривой, называют положительным эксцессом (рис. 21).

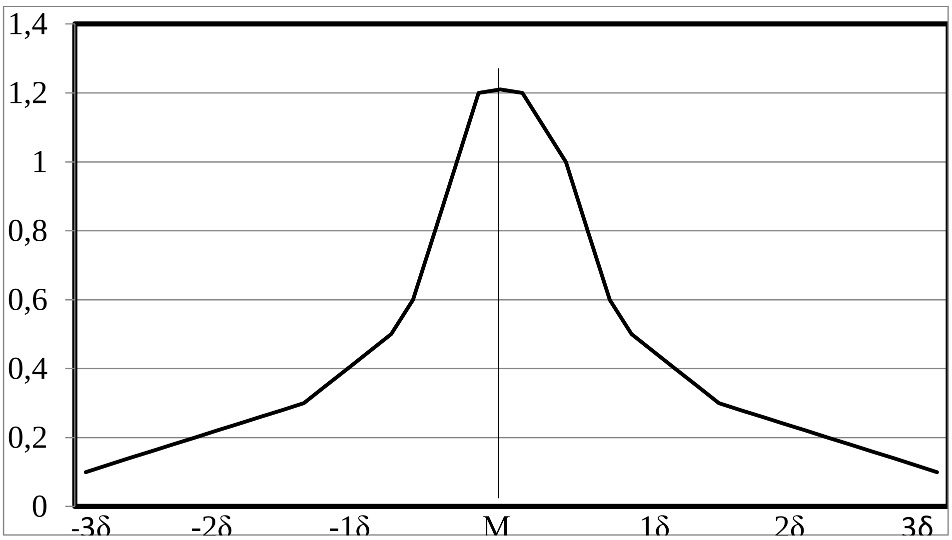
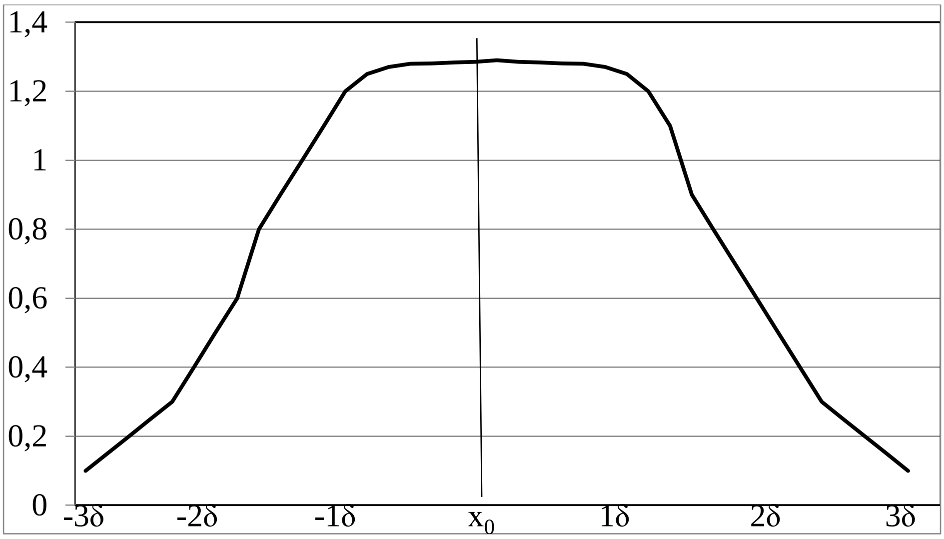


Рисунок 21 - Положительный эксцесс

Особенность эксцессивных кривых состоит в том, что ее крайние варианты $v_{макс}$. и $v_{мин.}$ отстоят от М не на ±3δ, как у нормальных кривых, а на большее значение δ и в приделах ±3δ находится не 99,7% всех наблюдений, а 97-98%.

Если эксцесс имеет отрицательный знак, то это свидетельствует о том, что кривая приобретает плосковершинность или двухвершинность (рис. 22).

У плосковершинных кривых ее крайние варианты $v_{макс.}$ и $v_{мин.}$ не доходят до границ +3δ и -3δ, то есть основание плосковершинной кривой меньше, что указывает на сниженную изменчивость признака.



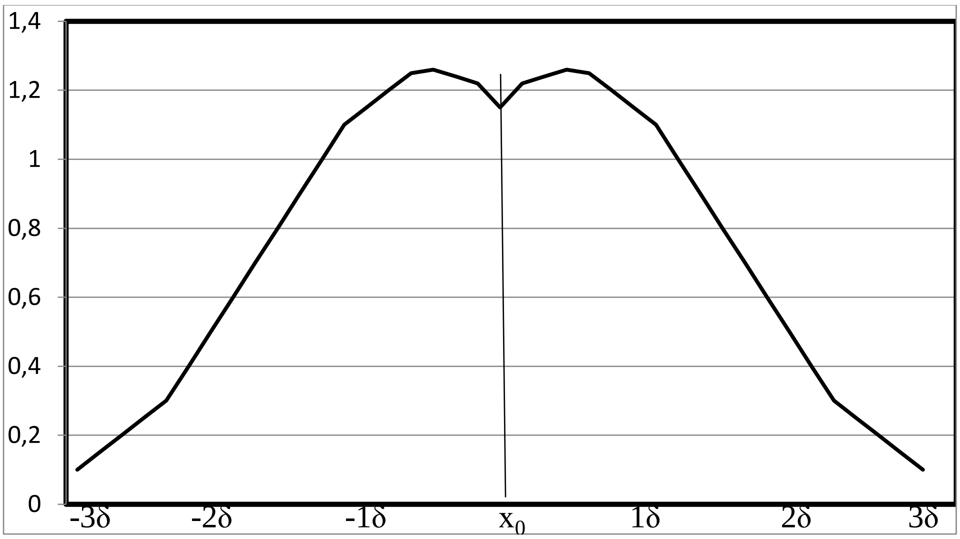


Рисунок 22 – Отрицательный эксцесс

Причины, вызывающие эксцесс, или в неправильно осуществленной выборке или в объективно существующих причинах, уменьшающих частоту появления особей в классах на концах ряда и увеличивающих их накопление в классах, близких к средней арифметической (М), моде (Мо) и медиане (Ме).

Мерой вытянутости вершины, или мерой эксцесса, служит коэффициент эксцесса Ех:

$$\mathrm{Ex} = \frac{\sum(\mathrm{v} - \mathrm{M})^4}{\mathrm{n} \cdot \delta^4} - 3,$$

где Ех – коэффициент эксцесса;

v – значение варьирующего признака;

М – средняя арифметическая;

n – число наблюдений;

δ – среднее квадратичное отклонение.

Если эксцесс близок 0,4, то это незначительное накопление частот. Если Ех=0, то тогда ряд имеет нормальное распределение.

Трансгрессивные ряды и трансгрессивные кривые

Явление трансгрессии прослеживается при обработке данных различных биологических особей. Трансгрессивными рядами и кривыми называются ряды, которые отличаются друг от друга величиной средней арифметической и их крайние классы, лежащие около максимального класса первой кривой, служат минимальными классами другой кривой, что создает в этих частях вариационных кривых их взаимное пересечение (рис. 23).

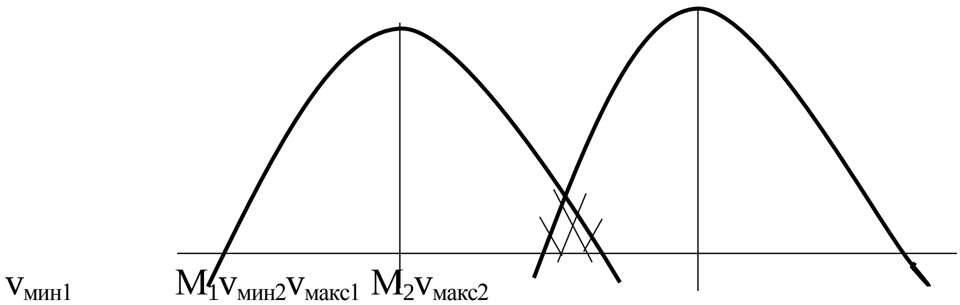


Рисунок 23 – Общий вид трансгрессирующих кривых

При изучении трансгрессивных рядов требуется решить следующие задачи:

1. Определить степень трансгрессии.

2. Определить, достоверна ли разность между средними арифметическими каждого ряда. Если разница между M_1 и M_2 достоверна, то это доказывает наличие двух трансгрессирующих рядов.

3. Определить, к какому из рядов следует отнести конкретную особь, которая имеет признак на уровне вариантов, отчленяющих пересекающиеся части обоих рядов.

Вычислим первый элемент трансгрессирующих рядов.

Определение степени трансгрессии проводится по следующей формуле:

$$T = \frac{n_1 \cdot p_1 + n_2 \cdot p_2}{n_1 + n_2},$$

где T – показатель трансгрессии;

n_1 и n_2 – общее число наблюдений в каждой из выборок;

p_1 и p_2 – доля трансгрессирующих частот в каждом из рядов,

ограниченных площадью кривой между $v_{мин2}$ и $v_{макс1}$.

Степень трансгрессии может быть большой и малой.

В формуле трансгрессии требуется найти доли трансгрессирующих частот p_1 и p_2. Для этого используют вторую функцию нормированного отклонения.

Долю трансгрессирующих частот первого ряда определяют с помощью следующих выражений:

$$p_1 = 0{,}5 \pm \varphi(x_1),$$

где

$$x_1 = \frac{v_{мин2} - M_1}{\delta_1},$$

где 0,5 – доля половинной площади от всей нормальной кривой первого ряда;

$\varphi(x_1)$ – буквенное выражение второй функции нормированного отклонения первого ряда (второй столбец таблицы «Площади и ординаты нормальной кривой распределения»);

$v_{мин2}$ – значение минимального варианта второго ряда, который может быть выражен как $Vмин_2 = M_2 - 3\delta_2$;

M_1 – средняя арифметическая первого ряда;

δ_1 – среднее квадратичное отклонение первого ряда.

Долю трансгрессирующих частот второго ряда определяют по аналогичной формуле, только соответственно меняют значки:

$$p_2 = 0{,}5 \pm \varphi(x_2),$$

где

$$x_1 = \frac{v_{мак1} - M_2}{\delta_2},$$

$$v_{мин2} = M_1 - 3\delta_1$$

Разберем на примере вычисление величины трансгрессии: имеются два вариационных ряда, характеризующих массу самок чавычи в р. Апуко и

Олюторском заливе Берингова моря возраста 4+(кг):

1 ряд: р. Апуко: 12,2; 10,1; 13,0;10,6; 10,4; 8,9; 12,5; 12,6; 12,2; 12,6; 10,9; 12,1; 10,8; 11,3; 11,8; 14,2; 13,2; 13,7; 12,1; 10,9; 9,8; 10,1; 10,6; 11,5; 12,7, 11,7; 11,5, 11,4; 11,8.

2 ряд: Олюторский залив:12,8; 13,1; 10,8; 12,6; 13,4; 12,8; 12,7; 11,1; 13,3; 13,1; 13,7; 10,9; 10,1; 12,0; 11,0; 10,8; 13,8; 13,4; 15,1; 12,8; 12,2; 11,6; 12,1; 12,6; 13,0; 13,2; 12,5; 12,9.

Вычисляем коэффициенты трансгрессии для первого и второго рядов.

1. Определяем крайние значения вариантов по каждому ряду:

$$v_{мин1}=8{,}9 \text{ кг}$$

$$v_{мак1}=14{,}2 \text{ кг}$$

$$v_{мин2}=9{,}8 \text{ кг}$$

$$v_{мак2}=15{,}1 \text{ кг}$$

2. Находим средние арифметические (M_1и M_2) и средние квадратичные отклонения (δ_1 и δ_2).

$$n_1=29$$

$$n_2=28$$

Следовательно, воспользуемся формулами простой средней арифметической и среднего квадратичного отклонения для малых выборок:

$$M_1 = \frac{\sum v_1}{n_1} = \frac{337{,}2}{29} = 11{,}6 \text{ (кг)}$$

$$M_2 = \frac{\sum v_2}{n_2} = \frac{349{,}4}{28} = 12{,}5 \text{ (кг)}$$

Для определения средних квадратичных отклонений найдем разницу между каждым значением вариационных рядов и средней арифметической, полученные результаты возведем в квадрат и суммируем. Найденную сумму разделим на количество членов ряда минус единица и извлечем квадратный корень:

1 ряд			2 ряд		
v	(v_1-M_1)	$(v_1-M_1)^2$	v	(v_2-M_2)	$(v_2-M_2)^2$
12,2	0,6	0,36	12,8	0,3	0,09
10,1	-1,5	2,25	13,1	0,6	0,36
13,0	1,4	1,96	10,8	-1,7	2,89
10,6	-1,0	1,00	12,6	0,1	0,01
10,4	-1,2	1,44	13,4	0,9	0,81
8,9	-2,7	7,29	12,8	0,3	0,09

12,5	0,9	0,81	12,7	0,2	0,04
12,6	1,0	1,00	11,1	-1,4	1,96
12,2	0,6	0,36	13,3	0,8	0,64
12,6	1,0	1,00	13,1	0,6	0,36
10,9	-0,7	0,49	13,7	1,2	1,44
12,1	0,5	0,25	10,9	-1,6	2,56
10,8	-0,8	0,64	10,1	-2,4	5,76
11,3	-0,3	0,09	12,0	-0,5	0,25
11,8	0,2	0,04	11,0	-1,5	2,25
14,2	2,6	6,76	10,8	-1,7	2,89
13,2	1,6	2,56	13,8	1,3	1,69
13,7	2,1	4,41	13,4	0,9	0,81
12,1	0,5	0,25	15,1	2,6	6,76
10,9	-0,7	0,49	12,8	0,3	0,09
9,8	-1,8	3,24	12,2	-0,3	0,09
10,1	-1,5	2,25	11,6	-0,9	0,81
10,6	-1,0	1,00	12,1	-0,4	0,16
11,5	-0,1	0,01	12,6	0,1	0,01
12,7	1,1	1,21	13,0	0,5	0,25
11,7	0,1	0,01	13,2	0,7	0,49
11,5	-0,1	0,01	12,5	0	0
11,4	-0,2	0,04	12,9	0,4	0,16
11,8	0,2	0,04			
n=29		∑41,26	n=28		∑33,72

$$\delta_1 = \sqrt{\frac{\sum(v_1 - M)^2}{n_1 - 1}} = \sqrt{\frac{41{,}26}{28}} = \sqrt{1{,}474} = 1{,}21 \text{ (кг)}$$

$$\delta_2 = \sqrt{\frac{\sum(v_2 - M)^2}{n_2 - 1}} = \sqrt{\frac{33{,}72}{27}} = \sqrt{1{,}249} = 1{,}12 \text{ (кг)}$$

3. Находим x_1 и x_2 в долях δ, то есть выражаем их через нормированное отклонение

$$x_1 = \frac{v_{\text{мин}2} - M_1}{\delta_1} = \frac{9{,}8 - 11{,}6}{1{,}21} = \frac{-1{,}8}{1{,}21} = -1{,}49 \approx -1{,}5$$

По таблице «Площади и ординаты нормальной кривой распределения»(табл. 5 приложения) находим, что при таком отклонении v от M значение $\varphi(x_1)$ будет равно 0,43319.

$$x_2 = \frac{v_{\text{макс}1} - M_2}{\delta_2} = \frac{14{,}2 - 12{,}5}{1{,}12} = \frac{1{,}7}{1{,}12} = 1{,}52 \approx 1{,}5$$

Соответственно φ (x_2) также равна 0,43319.

4. Находим площадь кривой, соответствующую доли или проценту частот p_1 и p_2трансгрессирующих рядов, то есть:

$$p = 0,5 \pm \varphi(x)$$

Для первого ряда это выражение берут в виде суммы $0,5+\varphi(x)$:

$$p_1 = 0,5+\varphi(x_1) = 0,5 + 0,43319 = 0,93319,$$

или 93,32% частот будут входить в трансгрессирующую часть ряда.

для второго ряда p_2 определяют, как разность 0,5- $\varphi(x_2)$:

$$p_2 = 0,5- \varphi(x_2) = 0,5-0,43319= 0,06681$$

или 6,68% частот второго ряда будут состоять в трансгрессии с частотами первого ряда.

5. Находим коэффициент трансгрессии:

$$T = \frac{n_1 \cdot p_1 + n_2 \cdot p_2}{n_1 + n_2} = \frac{29 \cdot 0,93319 + 28 \cdot 0,06681}{29 + 28} = \frac{27,06 \ + 1,87}{57} = \frac{28,93}{57}$$
$$= 0,50754,$$

или 50,75% частот имеют трансгрессию между обоими рядами.

Второй элемент анализа трансгрессии сводится к определению разности D между средними арифметическими каждого ряда:

$$D= M_1 – M_2$$

Если эта разность будет достоверна (когда статистическая ошибка этой разницы укладывается в ней не менее 3 раз), то такое различие между средними арифметическими обоих рядов будет свидетельствовать о трансгрессивном типе взаимоотношений между рядами. Если эта разность будет недостоверной, то один ряд как бы является частью другого и суммирование их частот по соответствующим классам даст единую кривую, проявляющую двухвершинность.

Для определения статистической ошибки разности между средними арифметическими обоих рядов используют формулу:

$$m_D = \sqrt{m_{M_1}^2 + \ m_{M_2}^2},$$

где m_{M1} и m_{M2} – ошибки средних арифметических каждого ряда.

Формулы этих ошибок:

$$m_M = \frac{\delta}{\sqrt{n}}$$

Используем данные нашего примера. Найдем статистическую ошибку для

M_1 и M_2.

$$m_{M_1} = \frac{\delta_1}{\sqrt{n_1}} = \frac{1,21}{\sqrt{29}} = \frac{1,21}{5,385} = 0,22\ (кг)$$

$$m_{M_2} = \frac{\delta_2}{\sqrt{n_2}} = \frac{1,12}{\sqrt{28}} = \frac{1,12}{5,292} = 0,21\ (кг)$$

Найдем разность средних арифметических:

$$D = M_2 - M_1 = 12,5\text{-}11,6 = 0,9 (кг)$$

Ошибка разности D равна:

$$m_D = \sqrt{m_1^2 + m_2^2} = \sqrt{0,22^2 + 0,21^2} = \sqrt{0,048 + 0,044} = \sqrt{0,09} = 0,3\ (кг)$$

Достоверность разностиD:

$$\frac{D}{M_D} = \frac{0,9}{0,3} = 3$$

Следовательно, ошибка M_D содержится 3 раз в разности D, что указывает на полную достоверность наличия трансгрессирующего наложения между вариационными рядами.

Третий элемент анализа трансгрессирующих рядов заключается в определении к какому из рядов следует отнести ту или иную особь, у которой величина признака находится в границах вариантов, являющихся общими для обоих рядов, то есть в границах трансгрессирующей части обоих вариационных рядов. Для определения принадлежности данной особи к тому или иному ряду, образующих трансгрессию, пользуются методом комбинированных признаков (предложенным ихтиологом Гейнеке), основанном на сопоставлении суммы квадратов отклонений $(v\text{-}M)^2$, вычисленных для трансгрессирующих рядов.

Поясним этот метод следующим примером: имеется особь чавычи массой 10,45 кг неизвестного происхождения. Требуется узнать, следует ли ее отнести на основании этого показателя к популяции р. Апуко или Олюторского залива Берингова моря.

Для решения этого вопроса возьмем несколько признаков, которые мало коррелируют друг с другом, но могут служить характеристикой для данных популяций. Положим, что в качестве показателей взяты длина тела и коэффициент упитанности по Фультану:

Район вылова	Вес, кг	Длина тела, см	Коэффициент упитанности по Фультану
р. Апуко (1)	11,6	100,2	1,01
Олюторский залив (2)	12,5	95,3	1,16
Показатели особи А	10,4	86,8	1,13

Сравним показатели интересующего нас особи А со средними показателями чавычи р. Апуко и Олюторского залива, выразив это через квадрат отклонения $(v_A\text{-}M_1)^2$ и $(v_A\text{-}M_2)^2$.

Использование метода комбинированных признаков

Признаки	Отклонения показателей особи от средних значений популяций чавычи	
	р. Апуко $(v_A\text{-}M_1)$	Олюторский залив
Вес, кг	10,4-11,6=-1,2	10,4-12,5=-2,1
Длина тела, см	86,8-100,2=-13,4	86,8-95,3=-8,5
Коэффициент упитанности по Фультану	1,13-1,01=0,12	1,13-1,16=-0,03
Квадраты отклонений по каждому признаку	$(v_A\text{-}M_1)$ $-1{,}2^2=1{,}44$ $-13{,}4^2=179{,}56$ $0{,}12^2=0{,}014$	$(v_A\text{-}M_2)$ $-2{,}1^2=4{,}41$ $-8{,}5^2=72{,}25$ $-0{,}03^2=0{,}001$
Сумма квадратов отклонений	$\sum(v_A\text{-}M_1)=181{,}014$	$\sum(v_A\text{-}M_2)=76{,}661$

Следовательно, особь А по сравниваемым признакам стоит ближе к популяции чавычи Олюторского залива, так как:

$$\sum(v_A\text{-}M_1)>\sum(v_A\text{-}M_2)$$

Этот случай удобен в тех случаях, когда исследователь располагает единичными экземплярами, в отношении которых необходимо выяснить, к какой группе особей они ближе относятся.

Статистические ошибки

Статистический метод изучения варьирующего признака основывается на использовании выборочной совокупности данных, взятых по принципу случайности из генеральной совокупности. Следовательно, выборочная совокупность представляет часть, а генеральная совокупность является целым, о котором мы судим по величинам, получаемым при обработке только части полученных значений. В результате этого, выборочный метод, лежащий в основе вариационно-статистической обработки материалов, может служит источником статистических ошибок. При этом необходимо добиваться того, чтобы ошибки были полностью устранены, а если их нельзя избежать, то следует свести их к минимуму.

Информация, полученная при проведении исследований или экспериментов с биологическими объектами, может не отвечать действительности, а расчетные показатели не соответствовать фактическим значениям. Расхождение между расчетным и фактическим данными называют ошибкой наблюдения. В зависимости от причин возникновения различают ошибки регистрации и ошибки репрезентативности, которые могут быть как случайными, так и систематическими.

Ошибки регистрации представляют собой отклонения между значением показателя, полученного в ходе статистического наблюдения, и его фактическим значением. Они бывают случайными, появляющимися в результате действий влияния каких-то факторов и систематическими, которые проявляются постоянно.

Ошибки репрезентативности возникают в случае, когда отобранная совокупность недостаточно точно воспроизводит исходную. Они характерны для несплошного наблюдения и заключаются в отклонении величины показателя исследуемой части совокупности от его величины в генеральной совокупности. То есть они обусловлены самим выборочным методом, при котором из генеральной совокупности отбирается по принципу случайности часть ее членов (случайная выборка).

Случайные ошибки являются результатом действия случайных факторов, вызванных неправильными действиями исследователя или неверными либо неучтенными параметрами наблюдения, или эксперимента. Они могут быть следствием просчетов, описок, неверных арифметических вычислений и т.п., происходящих из-за недостаточно внимательного отношения к работе. Устранение их можно осуществить перепроверкой всего исходного материала и более тщательной и внимательной работой.

Систематические ошибки всегда имеют одинаковую направленность к увеличению или уменьшению показателя по каждой единице наблюдения, вследствие чего значение показателя по совокупности в целом будет включать накопленную ошибку. Они появляются в результате неточного измерения какого-либо показателя. Это происходит из-за неточности или малой разрешающей силы используемых приборов. Так как величина неточности прибора всегда одинакова, то возникает систематическая ошибка измерения. Ее величину можно учесть, если определить путем проведения многократных измерений. Обычно ошибка прибора, то есть его точность, указывается на его паспорте и может быть внесена в виде поправки к проведенному измерению.

Чем меньше статистическая ошибка, вычисленная по отношению к какой-либо характеристике выборки, тем точнее итоговые результаты. По величине ошибки и соотношению ее с той выборочной характеристикой, для которой она вычислена, можно судить о том, достаточно ли точно выборочные данные отражают характеристики, присущие генеральной совокупности.

Прежде всего величина статистической ошибки зависит от числа наблюдений (n), вошедших в выборку, и от степени изменчивости изучаемого признака. Чем больше объем выборки, то есть чем больше в ней число наблюдений, тем меньше будет статистическая ошибка, и в то же время, чем больше изменчивость признака, тем она больше. Следовательно, чтобы уменьшить статистическую ошибку, нужно стремиться увеличить объем выборки, особенно в тех случаях, когда изучаемый признак обладает большой вариабельностью. В статистике разработаны приемы вычисления величины статистических ошибок. Методика их вычисления зависит от вида исследуемой величины или показателя.

Статистическую ошибку принято называть или средней или стандартной. Чаще всего ее обозначают буквой m, у которой подстрочно записывают значок, указывающий для какой величины она вычислена. Ошибки имеют то же именование, что и статистический показатель, для которого она определена.

Статистическая ошибка средней арифметической

Ошибка средней арифметической, является одним из важных статистических показателей. С ее помощью можно определить неоднородность выборки. Способ определения ошибки средней арифметической зависит от величины частот, вошедших в выборку, то есть от количества значений вариационного ряда.

Для средней арифметической большой выборки ошибка выражается следующей рабочей формулой:

$$m_M = \frac{\delta}{\sqrt{n}} \cdot \sqrt{1 - \frac{n}{N}},$$

где δ – выборочное среднее квадратичное отклонение;

n – число наблюдений в выборке;

N – число наблюдений в генеральной совокупности.

Когда N=∞, то $\sqrt{1-\frac{n}{\infty}} = \sqrt{1-0}$,формула ошибки может быть использована в виде:

$$m_М = \frac{δ}{\sqrt{n}}$$

Которую используют, если число наблюдений в выборке (n)составляет не менее 5-10% от величины генеральной совокупности (N). Из этой формулы видно, что ошибка средней арифметической прямо пропорциональна изменчивости признака и обратно пропорциональна числу наблюдений в выборке.

Как пользоваться статистической ошибкой, разберем на примере: из 320 выловленных в Камском водохранилище окуней возраста 3+ была взята десятипроцентная случайная выборка, то есть 32 экземпляров, у которых изучали их массу (г). Определить, правильно ли отражают данные выборки генеральную совокупность:60; 44; 65; 42; 49; 47; 50; 43; 42; 52; 61; 25; 33; 41; 28; 52; 33; 78; 64; 43; 71; 66; 39; 47; 49; 56; 50; 48; 39; 36; 47; 49.

Для оценки ошибки средней арифметической, прежде всего необходимо определить саму среднюю арифметическую и среднее квадратичное отклонение.

Находим минимальный и максимальный показатели:min– 25 г, max– 78 г. Размах колебаний признака – Lim78-25=53 г. Величина К=53:10=5,3≈5 г.

Классы	Средняя величина класса	Частоты p	Условное отклонение а	p·a	p·a²
25-29	27	2	-4	-8	32
30-34	32	2	-3	-6	18
35-39	37	3	-2	-6	12
40-44	42	6	-1	-6	6
45-49	А=47	7	0	0	0
50-54	52	4	1	4	4
55-59	57	1	2	2	4
60-64	62	3	3	9	27
65-69	67	2	4	8	32
70-74	72	1	5	5	25
75-79	77	1	6	6	36
		∑p=32		∑p·a=8	∑p·a²=196

Полученные данные таблицы подставим в формулы:

$$M = A + K \cdot \frac{\sum p \cdot a}{n} = 47 + 5 \cdot \frac{8}{32} = 48{,}25\ (г)$$

$$\delta = K \cdot \sqrt{\frac{\sum p \cdot a^2}{n} - \left(\frac{\sum p \cdot a}{n}\right)^2} = 5 \cdot \sqrt{\frac{196}{32} - \left(\frac{8}{32}\right)^2} = 5 \cdot \sqrt{6{,}12 - 0{,}06} = 5 \cdot \sqrt{6{,}06}$$

$$= 5 \cdot 2{,}46 = 12{,}3\ (см)$$

Ошибка средней арифметической:

$$m_M = \frac{\delta}{\sqrt{n}} = \frac{12{,}3}{\sqrt{32}} = \frac{12{,}3}{5{,}657} = 2{,}17\ (г)$$

Обычно статистическую величину записывают совместно с ее ошибкой, а именно:

$$M \pm m_M = 48{,}25 \pm 2{,}17\ (г)$$

Показателем того, насколько правильно выборочная средняя отражает генеральную среднюю, служит критерий достоверности (t), который получают путем деления любой средней статистической на ее ошибку.

$$t = \frac{средняя}{ееошибка}$$

В нашем примере:

$$t_M = \frac{M}{m_M} = \frac{48{,}25}{2{,}17} = 22{,}24$$

По величине критерия достоверности судят о том, насколько правильно выборочная средняя отражает генеральную среднюю. Чем больше его значение, тем достовернее выборочная средняя.

Для различных исследований установлены неодинаковые уровни критерия достоверности, то есть требования к уровню точности полученных выборочных величин.

Для производственных и научно-производственных опытов, и большинства биологических работ достаточен критерий достоверности t=2-2,5. При t=2,величина вероятности P=0,95, это означает, что в 95 опытах из 100 были получены значения статистических величин такие же, как в том, в котором критерий достоверности оказался равен 2 и только в 5 опытах могут быть получены иные характеристики выборочного материала. Следовательно, при t=2, что соответствует P=0,95, допускается 5%-ная ошибка в достоверности выборочных данных.

Если t= 2,5, то этому соответствует P=0,987, или один неверный опыт на приходится 76 правильных. В работах, где достоверность полученных данных требует более высокой вероятности, берут критерий достоверности на уровне t=3, когда P=0,997, что дает один неверный результат на 332 верных опыта.

В тех случаях, когда выводы от обработки материалов очень ответственны, критерий достоверности (t) берут на уровне 4, в этом случае P=999936, что дает один отрицательный опыт на 15625 положительных.

При малом числе наблюдений в выборке ошибка средней арифметической вычисляется по формуле:

$$m = \frac{\delta}{\sqrt{n-1}},$$

где δ – среднее квадратичное отклонение;

n – число наблюдений в выборке;

n-1 =v – число степеней свободы.

Обычное представление о достоверности средней арифметической по показателю критерия достоверности при малой выборке несколько изменяется. Как показали работы Стьюдента, величина критерия достоверности при малом числе наблюдений зависит от числа членов в выборке (n) и поэтому достоверная величина определяется с учетом числа степеней свободы (v), равным n-1 (числу наблюдений в выборке, уменьшенному на 1).

Достоверное значение средней арифметической определяют по специальной таблице Стьюдента (табл. 6 приложения), в которой приведены значения числа степеней свободы (v) и критериев достоверности (t) при разных уровнях вероятности (P).

Вычисление границ доверительных интервалов для генеральной средней арифметической при малом числе наблюдений осуществляется с помощью доверительных значений t.

Рассмотрим расчет ошибки средней арифметической для малой выборки на примере: определить границы доверительного интервала массы хариуса возраста 3+ р. Чусовой (г): 65; 130; 103; 51; 140; 88; 32; 75; 32; 77; 123; 46; 59; 124; 88; 95; 106; 112; 79; 135.

Найдем значения средней арифметической и среднего квадратичного отклонения, используя формулы для выборок с n<30.

$$M = \frac{\sum v}{n} = \frac{1760}{20} = 88 \text{ (г)}$$

Найдем разницу между каждым значением вариационных рядов и средней арифметической, полученные результаты возведем в квадрат и суммируем:

v	(v-M)	$(v\text{-}M)^2$
65	-23	529
130	42	1764
103	15	225
51	-37	1369
140	52	2704
88	0	0
32	-56	3136
75	-13	169
32	-56	3136
77	-11	121
123	35	1225
46	-42	1764
59	-29	841
124	36	1296
88	0	0
95	7	49
106	18	324
112	24	576
79	-9	81
135	47	2209
n=20		$\sum(v\text{-}M)^2$=21518

$$\delta = \sqrt{\frac{\sum(v - M)^2}{n - 1}} = \sqrt{\frac{21518}{19}} = \sqrt{1132{,}53} = 33{,}65 \text{ (г)}$$

На основании этих данных определяем ошибку m_M.

$$m_M = \frac{\delta}{\sqrt{n - 1}} = \frac{33{,}65}{\sqrt{20 - 1}} = \frac{33{,}65}{\sqrt{19}} = \frac{33{,}65}{4{,}36} = 7{,}72 \text{ (г)}$$

Отсюда

$$t = \frac{M}{m_M} = \frac{88}{7{,}72} = 11{,}4$$

По таблице значений критериев достоверности t (табл.6 приложения) находим, что при числе степеней свободыv=n-1=20-1=19: при вероятности P=0,95 - t=2,1; P=0,99 - t=2,9; P=0,999 - t=3,9.

Значение критериев достоверности t при различных уровнях вероятности P и числа степеней свободы v, дающие достоверную величину средней арифметической и достоверность разности (M_1– M_2) при малом и большом числе наблюдений n (фрагмент)

Число степеней свободы v	Уровень вероятности P		
	0,95	0,99	0,999
	Значение t		
1	12,71	63,7	637,0
2	4,3	9,9	31,6
3	3,2	5,8	12,9
4	2,88	4,6	8,6
5	2,6	4,0	6,9
6	2,4	3,7	6,0
7	2,4	3,5	5,3
8	2,3	3,4	5,0
9	2,3	3,3	4,8
10	2,2	3,2	4,6
11	2,2	3,1	4,4
12	2,2	3,1	4,3
13	2,2	3,0	4,1
14	2,15	3,0	4,1
15	2,1	3,0	4,1
16	2,1	2,9	4,0
17	2,1	2,9	4,0
18	2,1	2,9	3,9
19-20	2,1	2,9	3,9

Вычисленное же значение t больше табличного, следовательно, полученные данные массы 20 экземпляров хариуса возраста 3+ р. Чусовой правильно отражают среднюю арифметическую.

Статистическая ошибка при альтернативных признаках

При альтернативных признаках выборка образуется особями, подразделяющимися на два класса: имеющих данный альтернативный признак и группа у которой он отсутствует.

Для определения ошибки по каждой группе особей пользуются следующей формулой:

$$m_P = m_Q = \sqrt{\frac{P \cdot Q}{n}},$$

где P –число особей, имеющих альтернативные признаки;

Q – число особей, не имеющих альтернативные признаки;

n – число особей в выборке, равное P+Q.

Пример: определить ошибку для показателя численности речного окуня в двух группах. В первой группе, состоящей из 70 выловленных рыб, представлены экземпляры (P), зараженные гельминтами, во второй – 40 здоровых (Q).

Вычислим ошибку для обеих групп (для P и Q).

$$m_P = m_Q = \sqrt{\frac{P \cdot Q}{n}} = \sqrt{\frac{70 \cdot 40}{110}} = \sqrt{\frac{2800}{110}} = \sqrt{25{,}45} = 5{,}05 \text{ (рыб)}$$

Величина ошибки для обеих групп всегда одинаковая, то есть

$$m_P = m_Q$$

Но критерий достоверности (t) бывает разным, так как численность групп отличается друг от друга.

Определим t_P и t_Q.

$P \pm m_P = 70 \pm 5{,}05$ рыб

$$t_P = \frac{70}{5{,}05} = 13{,}86$$

Следовательно, численность группы зараженных окуней вполне достоверна, так как t>2.

$Q \pm m_Q = 40 \pm 5{,}05$рыб

$$t_Q = \frac{40}{5{,}05} = 7{,}92$$

Численность группы Q также достоверна.

Если численность особей в каждой группе альтернативных признаков выражается в процентах, то ошибка вычисляется по следующей формуле:

$$m_P = m_Q = \sqrt{\frac{P\% \cdot Q\%}{n}}$$

$$P = \frac{70 \cdot 100}{110} = 63{,}6\%$$

$$Q = \frac{40 \cdot 100}{110} = 36{,}4\%$$

Для нашего примера это дает следующее:

$$m_P = m_Q = \sqrt{\frac{63{,}6 \cdot 36{,}4}{110}} = \sqrt{\frac{2315{,}04}{110}} = \sqrt{21{,}045} = 4{,}59\%$$

Следовательно, достоверность показателей численности групп выраженных в процентах будет следующей:

По зараженной группе:

$$P' \pm m_{P'} = 63{,}4 \pm 4{,}59 \text{ рыб}$$

$$t_{P'} = \frac{P'}{m_{P'}} = \frac{63{,}4}{4{,}59} = 13{,}81$$

По здоровой группе:

$$Q' \pm m_{Q'} = 36.4 \pm 4{,}59 \text{ рыб}$$

$$t_{Q'} = \frac{Q'}{m_{Q'}} = \frac{36{,}4}{4.59} = 7{,}93$$

Что указывает на достоверность обеих долей.

Следовательно, указанные способы вычисления дают одинаковые значения достоверности, разница в сотых единицы возникла за счет округления.

Ошибку при альтернативных признаках можно получать и в относительных величинах, если численность особей в альтернативных классах выражается в долях единицы, то есть:

$$P' = \frac{P}{n} \text{ и } Q' = \frac{Q}{n}$$

Формула ошибки доли аналогична той, какая употреблялась и при абсолютных величинах численности альтернативных вариантов:

$$m_{P'} = m_{Q'} = \sqrt{\frac{P' \cdot Q'}{n - 1}},$$

где n- число наблюдений, или n=P+Q.

Для получения более точных значений ошибок доли в знаменателе берут n-1, что особенно имеет большое значение при малых выборках.

В нашем примере доля зараженных окуней составляет:

$$P' = \frac{P}{n} = \frac{70}{110} = 0{,}64 \text{ (рыб)},$$

а здоровых:

$$Q^{'} = \frac{Q}{n} = \frac{40}{110} = 0{,}36 \text{ (рыб)}$$

Ошибки долей будут равны:

$$m_{P^{'}} = m_{Q^{'}} = \sqrt{\frac{P^{'} \cdot Q^{'}}{n-1}} = \sqrt{\frac{0{,}64 \cdot 0{,}36}{109}} = \sqrt{\frac{0{,}23}{109}} = \sqrt{0{,}0022} = 0{,}047$$

Следовательно, достоверность долей будет следующей:

По зараженной группе:

$$P^{'} \pm m_{P^{'}} = 0{,}64 \pm 0{,}047 \text{ рыб}$$

$$t_{P^{'}} = \frac{P^{'}}{m_{P^{'}}} = \frac{0{,}64}{0{,}047} = 13{,}62$$

По здоровой группе:

$$Q^{'} \pm m_{Q^{'}} = 0{,}36 \pm 0{,}047 \text{ рыб}$$

$$t_{Q^{'}} = \frac{Q^{'}}{m_{Q^{'}}} = \frac{0{,}36}{0{,}047} = 7{,}66$$

Показатели критериев достоверности доказываю, что численность в обеих группах исследуемых окуней достаточна для получения точных результатов.

Следует иметь в виду, что указанный способ определения доверительных границ для долей альтернативных признаков дает верные значения только в тех случаях, когда P′ и Q′, близки к 0,5, или находятся в границах от 0,25 до 0,75. Если же доли приближаются к значениям 0 и 1, то эта формула может дать неверное представление.

Поэтому, для P′ и Q′, имеющих величину, меньшую 0,25 и большую 0,75, используется «Метод φ», или «Угловое преобразование Фишера». Критерий Фишера предназначен для сопоставления двух выборок по частоте встречаемости интересующего исследователя эффекта. Суть данного метода состоит в переводе процентных долей в величины центрального угла, который измеряется в радианах. Большей процентной доле будет соответствовать больший угол, а меньшей доле - меньший угол, но соотношения здесь не линейные. При увеличении расхождения между углами φ_1 и φ_2 и увеличения численности выборок значение критерия возрастает. Чем больше величина φ, тем более вероятно, что различия достоверны. Величина φ имеет ошибку, независящую от величины дисперсии и величины долей P′ и Q′.

Величина доли P′ связана с φ через величину синуса.

$$P' = \sin^2 \cdot \frac{\varphi}{2}$$

$$\varphi = 2\mathrm{Arcsin}\sqrt{P'}$$

Ошибка φ равна:

$$m_{\varphi} = \frac{1}{\sqrt{n}}$$

Значение величины φ находят по специальной таблице, в которой определенному значению доли P′ соответствует значение φ(табл. 8 приложения).

Для получения точных результатом при использовании метода φ, выборка должна отвечать некоторым условиям:

1. Ни одна из сопоставляемых долей не должна быть равной нулю. Формально нет препятствий для применения метода φ в случаях, когда доля наблюдений в одной из выборок равна 0, однако в этих случаях результат может оказаться неоправданно завышенным.

2. Верхний предел объема выборки в критерии φ отсутствует –она может быть сколь угодно большой.

3. Нижний предел объема выборки - 2 в одной из выборок.

4. Должны соблюдаться следующие соотношения в численности двух выборок:

- если в одной выборке 2 значения, то во второй их должно быть не менее 30:$n_1 = 2 \rightarrow n_2 \geq 30$;
- если в одной из выборок 3 значения, то во второй должно быть не менее 7: $n_1 = 3 \rightarrow n_2 \geq 7$;
- если в одной из выборок всего 4 значения, то во второй должно быть не менее 5:$n_1 = 4 \rightarrow n_2 \geq 5$;
- при n_1, $n_2 \geq 5$ - возможны любые сопоставления.

Определение ошибки для среднего квадратичного отклонения и коэффициента изменчивости

Статистические ошибки для квадратичного отклонения и коэффициента изменчивости выражаются следующими формулами:

Ошибка среднего квадратичного отклонения:

$$m_{\delta} = \frac{\delta}{\sqrt{2n}},$$

где δ – среднее квадратичное отклонение;

n – число наблюдений в выборке.

Ошибка коэффициента изменчивости:

$$m_{C_v} = \frac{C_v}{\sqrt{2n}},$$

где С – коэффициент изменчивости;

n – число наблюдений в выборке.

Эти формулы используют для выборок с большим числом наблюдений, а критерий достоверности берут при разных уровнях вероятности:

$t_{0,95}$= 2,0

$t_{0,99}$=2,6

$t_{0,999}$=3,3

При необходимости сопоставления степени изменчивости двух вариационных рядов можно использовать не только абсолютные значения коэффициентов вариации C_{v1} и C_{v2}, но и определять разность и ее достоверности между средними квадратичными отклонениями этих рядов:

$$D = \delta_1 - \delta_2$$

Для определения достоверности разности δ_1-δ_2 вычисляют ошибку разности и критерий достоверности t_D.

Ошибку разности между средними квадратичными отклонениями двух выборок вычисляют по следующей формуле:

$$m_D = m_{\delta 1 - \delta 2} = \sqrt{\frac{\delta_1^2}{2n_1} + \frac{\delta_2^2}{2n_2}},$$

где δ_1^2 иδ_2^2 - квадраты (или дисперсия) средних квадратичных отклонений

n_1 и n_2 – число наблюдений в каждой выборке

Эту формулу используют при большом числе наблюдений (n>30).

Критерий достоверности разности вычисляют обычным способом:

$$t_D = t_{\delta 1 - \delta 2} = \frac{\delta_1 - \delta_2}{m_D}$$

t_D должно быть ≥2.

Пример: определить наличие изменчивости показателей полной длину тела язя возраста 5+, выловленных в различных водоемах (мм):

1 ряд: Камское водохранилище: 233; 217; 222; 212; 271; 249; 232; 240; 257; 223; 241; 239; 241; 237; 251; 257; 253; 228.

2 ряд: Воткинское водохранилище: 264; 261; 273; 265; 279; 268; 271; 266; 262; 274; 278; 275; 259; 261; 274; 277; 263; 266.

Для нахождения достоверности разницы средних квадратичных отклонений двух вариационных рядов необходимо рассчитать M_1; M_2; δ_1; δ_2.

$$n_1=18$$

$$n_2=18$$

Следовательно, воспользуемся формулами простой средней арифметической и среднего квадратичного отклонения для малых выборок:

$$M_1 = \frac{\sum v_1}{n_1} = \frac{4303}{18} = 239{,}06 \approx 239 \text{ (мм)}$$

$$M_2 = \frac{\sum v_2}{n_2} = \frac{4836}{18} = 268{,}67 \approx 269 \text{ (мм)}$$

Для определения средних квадратичных отклонений найдем разницу между каждым значением вариационных рядов и средней арифметической, полученные результаты возведем в квадрат и суммируем.

1 ряд			2 ряд		
v	(v_1-M_1)	$(v_1-M_1)^2$	v	(v_2-M_2)	$(v_2-M_2)^2$
233	-6	36	264	-5	25
217	-22	484	261	-8	64
222	-17	289	273	4	16
212	-27	729	265	-4	16
271	32	1024	279	10	100
249	10	100	268	-1	1
232	-7	49	271	2	4
240	-1	1	266	-3	9
257	18	324	262	-7	49
223	-16	256	274	5	25
241	-2	4	278	9	81
239	0	0	275	6	36
241	-2	4	259	-10	100
237	-2	4	261	-8	64
251	12	144	274	5	25
257	18	324	277	8	64
253	14	196	263	-6	36
228	-11	121	266	-3	9
n=18		$\sum(v_1-M_1)^2=4089$	n=18		$\sum(v_2-M_2)^2=724$

$$\delta_1 = \sqrt{\frac{\sum(v_1 - M)^2}{n_1 - 1}} = \sqrt{\frac{4089}{17}} = \sqrt{240{,}53} = 15{,}51 \text{ (мм)}$$

$$\delta_2 = \sqrt{\frac{\sum(v_2 - M)^2}{n_2 - 1}} = \sqrt{\frac{724}{17}} = \sqrt{40{,}59} = 6{,}53 \text{ (мм)}$$

Достоверность разности: $D=\delta_1-\delta_2=15{,}51-6{,}53=8{,}98$ (мм).

Ошибка разности:

$$m_D = \sqrt{\frac{\delta_1^2}{2n_1} + \frac{\delta_2^2}{2n_2}} = \sqrt{\frac{15{,}51^2}{2 \cdot 18} + \frac{6{,}53^2}{2 \cdot 18}} = \sqrt{\frac{240{,}56}{36} + \frac{42{,}64}{36}} = \sqrt{6{,}682 + 1{,}184}$$
$$= \sqrt{7{,}866} = 2{,}80 \text{ (мм)}$$

Критерий достоверности разности:

$$t_D = \frac{\delta_1 - \delta_2}{m_D} = \frac{8{,}98}{2{,}80} = 3{,}21$$

По таблице (табл. 6 приложения) находим, что при числе степеней свободы $v=n-1=18-1=17$: при вероятности $P=0{,}95$ - $t=2{,}1$; $P=0{,}99$ - $t=2{,}9$; $P=0{,}999$ - $t=4{,}0$.

Полученный критерий указывает на наличие достоверности разности при $P=0{,}99$.

Определение ошибки для коэффициентов асимметрии и эксцесса

Для коэффициента асимметрии (A_s) ошибка определяется по следующей формуле:

$$m_{A_s} = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}},$$

где n – число наблюдений в выборке.

При больших выборках это может быть упрощено:

$$m_{A_s} = \sqrt{\frac{6}{n}}$$

Критерий достоверности асимметрии:

$$t_{A_s} = \frac{A_s}{m_{A_s}} \geq 3$$

Формула ошибки коэффициента эксцесса (Ex):

$$m_{E_x} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}},$$

где n – число наблюдений в выборке.

При большой выборке:

$$m_{E_x} = \sqrt{\frac{24}{n}} = 2\sqrt{\frac{6}{n}},$$

что дает те же результаты, если использовать удвоенную ошибку асимметрии, то есть:

$$m_{E_x} = 2m_{A_s}$$

Критерий достоверности эксцесса выражается следующей формулой:

$$t_{E_x} = \frac{E_x}{m_{E_x}} \geq 3$$

Статистические связи и методы вычисления их величин

Отличительной чертой водных биологических объектов является многообразие признаков, характеризующих каждый из них. Их можно устанавливать возрастом, размерами, весом, различными физиологическими показателями и т.д. Имея однородную совокупность предметов, можно изучать распределение их по любому из их признаков. Весьма часто можно усмотреть известную связь между вариациями по различным значениям. Например, чем больше размеры гидробионтов, тем обычно больше его вес.

В простейшем случае связь между двумя переменными величинами строго однозначна. Например, размер особей одного вида рыб, обитающих в одинаковых условиях, зависит от их возраста. Такого рода зависимость принято называть функциональной – связь между признаками или показателями, когда при изменении одного на определенную величину другой тоже меняется на определенную величину. Например, при повышении температуры воды на какое-то количество градусов, в результате удлинение водородных связей между молекулами, объем ее увеличится на определенную величину.

Для биологических объектов связь обычно бывает менее «жесткой». Экземпляры с одинаковым значением одного признака имеют, как правило, различные значения по другим показателям. Такую связь между вариациями называют корреляцией (дословный перевод: соотношение) между значениями.

При корреляционных связях изменение одного признака у ряда особей на определенную величину будет сопровождаться изменениями другого на различные, то есть варьирующие значения. Например, корреляционная связь между весом животных и их длиной выражается в том, что каждому значению длины соответствует определенное распределение веса (а не одно значение веса) и с увеличением длины увеличивается и средний вес животных. У животных и растений все процессы и все признаки взаимно связаны и вместе с тем каждый из них, в свою очередь, связан с внешней средой, следовательно, корреляционные связи являются широко распространенными и требуют углубленного изучения.

Наиболее правильный путь изучения корреляционных связей - определение их с использованием биологических методов, которые позволяют вскрыть природу взаимосвязи. Но, кроме того, дополнительно для выяснения величины, типа и направления связи вполне целесообразно пользоваться методом математического анализа на массе особей. Для этого можно применять несколько статистических коэффициентами, каждый из которых позволяет выяснить различные стороны корреляционной связи. По своим математическим особенностям корреляционные связи могут быть:

- прямыми (положительными) и обратными отрицательными;
- прямолинейными и криволинейными;
- простыми и множественными;
- между количественными признаками;
- между качественными признаками.

Если с увеличением (или уменьшением) одного признака другой также увеличивается (или уменьшается), то такая связь называется прямой (по мере роста увеличивается вес гидробионта).

Если с увеличением одного показателя другой будет уменьшаться, то есть изменяться в обратном направлении, то такая связь называется обратной (чем старше рыба, тем ниже тем ее роста, или снижение солености воды приводит к замедлению темпов роста морских гидробионтов).

Прямолинейный – это тип связи, при котором равным друг другу изменениям одного признака соответствуют равные же изменения другого (например, с увеличением питательности кормов, увеличивается размер гидробионтов).

Криволинейный тип связи характеризуется тем, что при увеличении одного признака (или его уменьшении) другой признак сначала увеличивается, а затем уменьшается. Например, по мере роста рыбы до определенного возраста плодовитость возрастает, а затем снижается. При этом сначала уменьшается относительная плодовитость, а затем абсолютная. При повышении температуры на каждые 10^0С уровень обмена веществ повышается в 2-4 раза до определенного лимитирующего значения, а затем снижается (правило Вант-Гоффа).

Простая корреляционная связь – это связь между двумя признаками, без учета имеющихся других связей (возраст рыб – вес).

При множественной корреляционной связи выясняется связь между несколькими показателями (темп роста рыб – запасы корма, температура, соленость, кислотность воды).

Корреляционные связи могут выясняться не только между количественными признаками, но и между качественными (состояние кормовой базы водоема – размер рыб).

Корреляционная связь между признаками бывает линейной и нелинейной, положительной и отрицательной. Задача корреляционного анализа сводится к установлению направления и формы связи между варьирующими признаками, измерению ее тесноты и, наконец, к проверке достоверности выборочных показателей корреляции.

Для характеристики связи, ее направления и степени сопряженности переменных применяют следующие параметрические показатели:

- коэффициент корреляции r;
- корреляционное отношение η.

Наряду с параметрическими показателями для измерения корреляционной зависимости между признаками применяют и непараметрические:

- бисериальный показатель связи r_b;
- полихорический показатель связи, или коэффициент взаимной сопряженности ρ;
- коэффициент ассоциации Юла ($r_{ac.}$);
- коэффициент контингенции Пирсона ($r_{кон.}$).

Наряду с анализом двумерных совокупностей в биологии широкое применение находит статистический анализ многомерных корреляционных связей. Если известна связь между несколькими различными признаками, можно определить частные или парциальные коэффициенты корреляции.

Коэффициент корреляции r

К настоящему времени создано множество различных коэффициентов корреляции. Однако самые важные меры связи разработаны британскими учеными – Пирсона (командой британских ученых во главе с Карлом (Чарльзом) Пирсоном в 90-х годах XIX века), Спирмена (в 1904 году психологом Чарльзом Эдвардом Спирменом) и Кендалла (английским статистиком Морисом Джорджем Кендаллом в середине XX века). Их общей особенностью является то, что они отражают взаимосвязь двух признаков, измеренных в количественной шкале - ранговой или метрической.

Критерий корреляции Пирсона позволяет определить, какова теснота (или сила) корреляционной связи между двумя показателями, измеренными в количественной шкале. При помощи дополнительных расчетов можно также определить, насколько статистически значима выявленная связь.

Коэффициент ранговой корреляции Спирмена используют для выявления и оценки тесноты связи между двумя рядами сопоставляемых количественных показателей. В том случае, если ранги показателей, упорядоченных по степени возрастания или убывания, в большинстве случаев совпадают (большему значению одного показателя соответствует большее значение другого), делают вывод о наличии прямой корреляционной связи. Если ранги показателей имеют противоположную направленность (большему значению одного показателя соответствует меньшее значение другого), то говорят об обратной связи между показателями.

Коэффициент корреляции Кендалла используют в случае, когда переменные представлены двумя порядковыми шкалами при условии, что связанные ранги отсутствуют. Вычисление коэффициента Кендалла связано с подсчетом числа совпадений и инверсий (перестановок).

Коэффициент корреляции ®позволяет определить величину и направление связи при прямолинейном ее типе или близком к прямолинейному. При криволинейной связи им пользоваться нельзя, так как он сильно ее преуменьшает, а в ряде случаев даже не может ее уловить. Выряжается десятичной дробью и может принимать значения от 0 до ±1. Чем ближе значение r к 1, тем больше связь между данными признаками. О тесной (сильной) связи корреляции говорят лишь тех случаях, когда коэффициент корреляции (r) не ниже 0,7:

- средняя связь – r = 0,5-0,69;
- умеренная связь – r=0,31-0,49;
- слабая связь – r=0,21-0,3;
- очень слабая связь r =0,2 (часто вообще не учитывается).

Формулы коэффициента корреляции могут быть выражены различно. В общем виде она может быть представлена как сумма произведений нормированного отклонения вариантов каждого признака от своей средней, деленной на число наблюдений:

$$r = \frac{\sum[\frac{(V_x - M_x)}{\delta_x} \cdot \frac{(V_y - M_y)}{\delta_y}]}{n},$$

или

$$r = \frac{\sum t_x \cdot t_y}{n},$$

где t_x и t_y – нормированные отклонения по признаку x и признаку y.

Использование в формуле коэффициента корреляции значений варьирующих признаков, выраженных через их нормированное отклонение, то есть в долях среднего квадратичного отклонения (δ), позволяет вычислить связь между признаками, измеренными мерами разных именований (литры с процентами, килограммы с сантиметрами и т.п.).

Рабочие формулы коэффициента корреляции применяются с учетом того, с какой выборкой (большой или малой) и с какими значениями вариантов (однозначными, многозначными или дробными) имеют дело.

Так, для малых выборок при многозначных показателях вариантов удобнее всего пользоваться следующими формулами:

$$r = \frac{\sum x \cdot y - \frac{\sum x \cdot \sum y}{n}}{\sqrt{a_x \cdot a_y}},$$

где

$$a_x = \sum x^2 - \frac{(\sum x)^2}{n}$$

и

$$a_y = \sum y^2 - \frac{(\sum y)^2}{n}$$

x – варианты первого признака;

y – варианты второго признака;

n – число наблюдений в выборке.

Эта формула может быть представлена в форме, где вводятся значения средних арифметических по каждому признаку:

$$r = \frac{\sum x \cdot y - n \cdot M_x \cdot M_y}{\sqrt{(\sum x^2 - n \cdot M_x^2)(\sum y^2 - n \cdot M_y^2)}}$$

Разберем использование этих формул для малых выборок: определить величину и направление связи между возрастом и плодовитостью леща Запорожского водохранилища (n=15):

Представим в виде вариационного ряда показатели плодовитости каждой самки определенного возраста, записывая их парно, и произведем вычисления таких показателей, как $x \cdot y$, x^2, y^2:

Возраст, лет x	Плодовитость, икринок, тыс. шт., y	$x \cdot y$	x^2	y^2
3	51	153	9	2601
4	132	528	12	17424
6	145	870	36	21025
5	128	640	25	16384
5	130	650	25	16900
4	124	496	16	15376
4	114	456	16	12996
3	54	162	9	2916
8	194	1552	64	37636
4	122	488	16	14884
7	173	1211	49	29929
3	53	159	9	2809
8	202	1616	64	40804
6	148	888	36	21904
6	137	822	36	18769
$\sum x$=76	$\sum y$=1907	$\sum x \cdot y$=10691	$\sum x^2$=422	$\sum y^2$=272357

Вычислим значения a_x и a_y.

$$a_x = \sum x^2 - \frac{(\sum x)^2}{n} = 422 - \frac{76^2}{15} = 422 - \frac{5776}{15} = 422 - 385{,}07 = 36{,}93$$

$$a_y = \sum y^2 - \frac{(\sum y)^2}{n} = 272357 - \frac{1907^2}{15} = 272357 - \frac{3636649}{15}$$
$$= 272357 - 242443{,}27 = 29913{,}73$$

Определим коэффициент корреляции:

$$r = \frac{\sum x \cdot y - \frac{\sum x \cdot \sum y}{n}}{\sqrt{a_x \cdot a_y}} = \frac{10691 - \frac{76 \cdot 1907}{15}}{\sqrt{36,93 \cdot 29913,73}} = \frac{10691 - \frac{144932}{15}}{\sqrt{1104714,05}} = \frac{10691 - 9662,13}{1051,05}$$
$$= \frac{1028,87}{1051,05} = +0,98$$

Таким образом, связь между возрастом и плодовитостью леща тесная и положительная, то есть чем выше возраст самки, тем выше ее плодовитость.

Если для этого примера использовать формулу вторую формулу, то расчеты упростятся. Для этой формулы нужны значения $\sum x^2$,$\sum y^2$ и $\sum x \cdot y$, которые у нас уже вычислены. Требуется найти средние арифметические возраста M_x и плодовитости M_y и их квадраты

$$M_x = \frac{\sum x}{n} = \frac{76}{15} = 5,07 \text{ (лет)}$$
$$M_x^2 = 5,07^2 = 25,70 \text{ (лет)}$$
$$M_y = \frac{\sum y}{n} = \frac{1907}{15} = 127,13 \text{ (тыс. шт.)}$$
$$M_y^2 = 127,13^2 = 16162,04 \text{ (тыс. шт.)}$$

Подставим все имеющиеся значения в формулу:

$$r = \frac{\sum x \cdot y - n \cdot M_x \cdot M_y}{\sqrt{(\sum x^2 - n \cdot M_x^2)(\sum y^2 - n \cdot M_y^2)}}$$
$$= \frac{10691 - 15 \cdot 5,07 \cdot 127,13}{\sqrt{(422 - 15 \cdot 25,70)(272357 - 15 \cdot 16162,04)}}$$
$$= \frac{10691 - 9668,24}{\sqrt{(422 - 385,50)(272357 - 242430,60)}} = \frac{1022,76}{\sqrt{36,5 \cdot 299264,4}}$$
$$= \frac{1022,76}{\sqrt{1092313,6}} = \frac{1022,76}{1045,14} = +0,98$$

Таким образом, вычисление по обеим формулам дает одинаковое значение коэффициента корреляции.

Для больших выборок вычисление коэффициента корреляции можно осуществлять по следующей формуле:

$$R = \frac{\sum p a_x A_y - n b_x b_y}{N \cdot \delta_x \cdot \delta_y}$$

Для обработки данных большой выборки строят корреляционную решетку, которая объединяет частоты (p) по обоим коррелирующим признакам. Основу решетки составляют классы, получаемые из данных о варьировании каждого признака. По горизонтали записывают классы одного признака, а по вертикали – классы другого. Пересечение столбцов и строчек классов образуют сетку решетки в виде клеток, в которые разносят данные с учетом величины обоих признаков у каждой особи, вошедшей в выборку. Обрабатывать данные корреляционной решетки можно известным уже методом произведений или методом сумм.

Разберем на примере вычисление методом произведений: требуется вычислить коэффициент корреляции между измерениями промысловой длины (мм) и массы тела (г) у байкальского омуля (n=35):

Длина, мм (x)	Масса, г (y)	Длина, мм (x)	Масса, г (y)	Длина, мм (x)	Масса, г (y)	Длина, мм (x)	Масса, г (y)
343	497	384	771	294	289	273	235
370	730	129	23	286	272	350	577
275	265	353	537	301	314	198	82
300	313	236	149	180	78	345	523
263	215	361	635	287	281	171	52
239	155	306	460	187	64	330	454
296	324	211	98	330	505	305	329
346	513	282	272	367	667	239	145
215	118	374	711	322	343		

Начинаем обработку с составления классов по каждому признаку. Для этого находим минимальные и максимальные значения длины (x) и массы (y)тела байкальского омуля.

$$x_{мин.}=129 \text{ мм} \quad y_{мин.}=23 \text{ г}$$

$$x_{мак.}=384 \text{ мм} \quad y_{мак.}=771 \text{ г}$$

Разность по длине:

$$D_x=384-129=255 \text{ (мм)}$$

Разность по массе:

$$D_y=771-23=748 \text{ (г)}$$

Определяем размер класса K для каждого признака:

$$K_x = \frac{D_x}{8} = \frac{255}{10} = 25{,}5 \approx 26$$

$$K_y = \frac{D_y}{8} = \frac{748}{10} = 75{,}8 \approx 76$$

Строим корреляционную решетку.

x	y										p_y	a_y	$p_y a_y$	$p_y a_y^2$
	23-98	99-174	175-250	251-326	327-402	403-478	479-554	555-630	631-706	707-782				
	1 квадрат				2 квадрат									
129-154	.1										1	-6	-6	36
155-180	..2										2	-5	-10	50
181-206	..2										2	-4	-8	32
207-232	.1	.1									2	-3	-6	18
233-258		…3									3	-2	-6	12
259-284			..2	..2							4	-1	-4	4
285-310				……6	.1	.1					8	0	0	0
	3 квадрат				4 квадрат									
311-336					.1	.1	.1				3	1	3	3
337-362							…..5		.1		6	2	12	24
363-388									.1	…3	4	3	12	36
p_x	6	4	2	8	2	2	6	0	2	3	n=35	-	∑=-13	∑=215
a_x	-3	-2	-1	0	1	2	3	4	5	6				
$p_x a_x$	-18	-8	-2	0	2	4	18	0	10	18	∑=24			
$p_x a_x^2$	54	16	2	0	2	8	54	0	50	108	∑=294			

Графы p_x и p_y заполняют суммированием частот по каждому из признаков. Для этого по каждому признаку выделяют условно средние классы, которые удобнее брать исходя из их центрального расположения и наибольшего числа частот. В данном примере условно средние классы: по длине - с градациями 285-310 мм, по массе - 251-326 г. Эти классы выделяют, отчего в решетке образуется фигура «креста»; они служат нулевыми классами, от которых остальные перечисляют по порядку в условных отклонениях a_x и a_y. В сторону уменьшения признака от нулевого класса отклонения идут со знаком минус, в сторону увеличения признака – со знаком плюс.

В формуле коэффициента корреляции требуется рассчитать b_x и b_y,δ_x и δ_y. Эти значения определяют обычным приемом обработки вариационного ряда. Для этого заполняют графы $p \cdot a$ и $p \cdot a^2$ и получают их суммы:

$$b_x = \frac{\sum p_x \cdot a_x}{n} = \frac{24}{35} = 0{,}69 \text{ (мм)}$$

$$b_x^2 = 0{,}48 \text{ (мм)}$$

$$b_y = \frac{\sum p_y \cdot a_y}{n} = \frac{-13}{35} = -0{,}37 \text{ (г)}$$

$$b_y^2 = 0{,}14 \text{ (г)}$$

Вычисляем δ_x и δ_y. При этом следует иметь в виду, что для формулы коэффициента корреляции средние квадратичные отклонения выражаются в относительных величинах, то есть без умножения корня на классовый промежуток К.

$$\delta_x = \sqrt{\frac{\sum p_x \cdot a_x^2}{n} - b_x^2} = \sqrt{\frac{294}{35} - 0{,}48} = \sqrt{7{,}92} = 2{,}81 \text{ (мм)}$$

$$\delta_y = \sqrt{\frac{\sum p_y \cdot a_y^2}{n} - b_y^2} = \sqrt{\frac{215}{35} - 0{,}14} = \sqrt{6{,}00} = 2{,}45 \text{ (г)}$$

Для формулы коэффициента корреляции осталось неизвестным выражение $\sum p \cdot a_x \cdot a_y$. Эту сумму получают путем умножения каждого значения частот p по клеткам решетки на условные отклонения a_x и a_y. При этом действие умножения осуществляют только для тех частот, которые расположены в клетках за пределами нулевых классов (за пределами «креста»), то есть в клетках 1,2,3,4 квадратов решетки.

Произведем построчное умножение pa_xa_y по каждому квадрату:

1 квадрат

1-я строка: $1 \cdot -3 \cdot -6 = +18$

2-я строка: $2 \cdot -3 \cdot -5 = +30$

3-я строка: $2 \cdot -3 \cdot -4 = +24$

4-я строка: $1 \cdot -3 \cdot -3 = +9$

$1 \cdot -2 \cdot -3 = +6$

5-я строка: $3 \cdot -2 \cdot -2 = +12$

6-я строка: $2 \cdot -1 \cdot -1 = +2$

$\sum p \cdot a_x \cdot a_y = 101$

2 квадрат

$\sum p \cdot a_x \cdot a_y = 0$

3 квадрат

$\sum p \cdot a_x \cdot a_y = 0$

4 квадрат:

8-я строка: 1·1·1=+1

1·2·1=+2

1·3·1=+3

9-я строка: 5·3·2=+30

1·5·2=+10

10-я строка: 1·5·3=+15

3·6·3=+54

$\sum p \cdot a_x \cdot a_y = 115$

Суммарное значение по четырем квадратам: $\sum p \cdot a_x \cdot a_y = 101 + 115 = +216$

Следует иметь в виду, что 1 и 4 квадраты имеют всегда положительное значение $\sum p \cdot a_x \cdot a_y$, 2 и 3 – всегда отрицательное значение $\sum p \cdot a_x \cdot a_y$.

Подставим все найденные величины в формулу коэффициента корреляции:

$$r = \frac{\sum p \cdot a_x \cdot a_y - n \cdot b_x \cdot b_y}{n \cdot \delta_x \cdot \delta_y} = \frac{216 - 35 \cdot 0{,}69 \cdot -0{,}37}{35 \cdot 2{,}81 \cdot 2{,}45} = \frac{216 + 8{,}94}{240{,}96} = \frac{224{,}94}{240{,}96} = +0{,}93$$

Таким образом, связь между промысловой длиной и массой тела у байкальского омуля значительная и положительная (чем больше длина, тем выше масса).

По расположению частот в клетках решетки по диагонали можно судить о направлении связи, ее типе и уровне. Так, если числа частот в основном группируются близко к диагонали, идущей из верхнего левого угла в правый нижний угол, то коэффициент корреляции будет иметь знак плюс. Если частоты группируются вдоль другой диагонали решетки, то связь отрицательная и коэффициент корреляции будет иметь знак минус. Чем ближе группируются частоты к диагонали, тем больше значение r. При распределении частот в клетках решетки беспорядочно связь будет незначительная. Если частоты располагаются в решетке дугообразно или как бы образуют фигуру полумесяца, то связь имеет криволинейный тип и вычислять коэффициент корреляции нецелесообразно.

Ошибка коэффициента корреляции

При большом числе наблюдений (n≥100), и при высоком значении коэффициента корреляции ошибку вычисляют по следующей формуле:

$$m_r = \frac{1 - r^2}{\sqrt{n - 1}},$$

где r – коэффициент корреляции, вычисленный при n≥100;

n – число наблюдений в выборке.

Если объем выборки имеет меньше 100 наблюдений (n<100), то такое вариационное распределение коэффициента корреляции начинает отклоняться от нормального и использование вышеприведенной формулы может дать искаженное значение ошибки.

Поэтому при малых выборках формула ошибки коэффициента корреляции видоизменяется следующим образом:

$$m_r = \sqrt{\frac{1 - r^2}{n - 2}}$$

Критерий достоверности определяют по формуле:

$$t_r = \frac{r}{m_r}$$

Определим ошибку коэффициента корреляции для нашего примера связи между промысловой длиной и массой тела у байкальского омуля. Поскольку число наблюдений меньше 100, ошибка коэффициента корреляции определяется по формуле:

$$m_r = \frac{1 - r^2}{\sqrt{n - 2}} = \frac{1 - 0{,}93^2}{\sqrt{36 - 2}} = \frac{1 - 0{,}89}{\sqrt{34}} = \frac{0{,}11}{5{,}83} = 0{,}02$$

$$t_r = \frac{r}{m_r} = \frac{0{,}93}{0{,}02} = 46{,}5$$

Полученная величина критерия достоверности настолько велика, что вероятность достоверности результатов составляет более 0,999. Таким образом достоверность связи не вызывает сомнений.

Корреляционное отношение η

Описанный выше коэффициент корреляции является полноценным показателем тесноты связи лишь в случае линейной зависимости между переменными. Однако часто возникает необходимость в достоверном показателе интенсивности связи при любой форме корреляции. В случае наличия линейной или нелинейной зависимости между двумя признаками для измерения тесноты связи применяют корреляционное отношение.

Различают эмпирическое и теоретическое корреляционное отношение. Теоретическое или индекс корреляции характеризует тесноту связи при любой форме зависимости, эмпирическое рассчитывают по данным группировки.

Величина корреляционного отношения изменяется от 0 до 1. Близость ее к нулю говорит об отсутствии связи, близость к единице – о тесноте связи.

Для оценка связи на основе теоретического корреляционного отношения используют шкалу Чеддока:

Значение	Характер связи	Значение	Характер связи
$\eta = 0$	Отсутствует	$0,5 \leq \eta < 0,7$	Заметная
$0 < \eta < 0,2$	Очень слабая	$0,7 \leq \eta < 0,9$	Сильная
$0,2 \leq \eta < 0,3$	Слабая	$0,9 \leq \eta < 1$	Весьма сильная
$0,3 \leq \eta < 0,5$	Умеренная	$\eta = 1$	Функциональная

Для линейной зависимости теоретическое корреляционное отношение тождественно линейному коэффициенту корреляции: $\eta = r$

Отметим основные свойства корреляционных отношений (при достаточно большом объеме выборки):

- корреляционное отношение есть неотрицательная величина, не превосходящая 1: $0 < \eta < 1$;
- если $\eta = 0$, то корреляционная связь отсутствует;
- если $\eta = 1$, то между переменными существует функциональная зависимость;
- если $\eta_{yx} \neq \eta_{xy}$, то есть, в отличие от коэффициента корреляции (r), для которого $r_{xy} = r_{yx}$, при вычислении корреляционного отношения существенно, какую переменную считать независимой, а какую - зависимой.

Корреляционное отношение определяют через отношение межгрупповой дисперсии к общей дисперсии. В связи с этим подробный ее расчет будет рассмотрен в главе «Дисперсионный анализ».

Бисериальный показатель связи r_b

Бисериальный показатель связи r_b применяют в тех случаях, когда один признак выражен количественно, а другой имеет качественное и при том альтернативное выражение. Бисериальные коэффициенты корреляции изменяются в диапазоне от -1 до +1, однако следует помнить, что в данном случае знак для интерпретации не имеет значения (это исключение из общего правила).

Формула бисериального показателя:

$$r_b = \frac{\frac{\sum p_+ \cdot a}{n_+} - \frac{\sum p \cdot a}{n}}{\sqrt{\frac{\alpha}{n_+} - \frac{\alpha}{n}}},$$

где

$$\alpha = \sum p \cdot a^2 - \frac{(\sum p \cdot a)^2}{n},$$

где p – частоты ряда всей выборки, распределенные по количественному признаку;

p_+ – частоты ряда по одному из альтернативных признаков (+);

n – число наблюдений в выборке;

n_+ – число наблюдений в ряду одного из альтернативных признаков (+);

α – величина, входящая в значение δ для вариационного ряда;

a – условное отклонение для ряда всей выборки по количественному признаку.

Для вычисления r_b строят корреляционную решетку обычным методом, в которой будут два класса для альтернативного признака (+ и –) и классы для количественного признака. Материал по классам решетки разносят обычным способом.

Пример: определить связь между полной длины тела байкальского омуля селенгинской и северобайкальской рас возраста 7+ (мм):

Селенгинская раса: 331; 338; 327; 329; 302; 313; 307; 267; 304; 322; 312; 305; 277; 269; 335; 320; 278; 284; 328; 265; 314; 333; 321; 318; 313; 299; 287; 296; 278; 330; 281; 280.

Северобайкальская раса: 268; 261; 299, 304; 307; 277; 270; 283; 304; 255; 312; 306; 278; 265; 263; 273; 288; 281; 301; 270; 269; 260; 256; 280; 294; 290; 303; 305; 285; 282; 262; 284.

Находим минимальный и максимальный показатели для обоих вариационных рядов:min – 255мм, max – 338 мм. Размах колебаний признака – Lim338-255=83 мм. Величина К=83:10=8,3≈9 мм.

Группы	Полная длина тела (мм)										p_2
	255-263	264-272	273-281	282-290	291-299	300-308	309-317	318-326	327-335	336-344	
Селенгинская раса (p_+)	0	3	5	2	2	4	4	4	7	1	n_+=32
Северобайкальская раса (p_-)	5	6	5	6	2	7	1	0	0	0	n_-=32
p_1	5	9	10	8	4	11	5	4	7	1	n=64
a	-5	-4	-3	-2	-1	0	1	2	3	4	-
$p_+ \cdot a$	0	-12	-15	-4	-2	0	4	8	21	4	$\sum p_+ \cdot a$ =4
$p_1 \cdot a$	-25	-36	-30	-16	-4	0	5	8	21	4	$\sum p_1 \cdot a$ =-73
$p_1 \cdot a^2$	125	144	90	32	4	0	5	16	63	16	$\sum p_1 \cdot a^2$ =495

Берем условное отклонение (а) в классе, имеющим полная длину телав интервале 300-308 мм, и выражаем в строчку остальные классы по плодовитости a_0 вправо со знаком плюс и влево со знаком минус. Следующая строчка образуется от умножения частот Селенгинской расы (p_+) на отклонение (а), то есть составляем ряд $p_+ \cdot a$. Следующие две строчки составляют обычным способом для получения ряда $p_1 \cdot a$ и $p_1 \cdot a^2$. Находим α:

$$\alpha = \sum p_1 \cdot a^2 - \frac{(\sum p \cdot a)^2}{n} = 495 - \frac{(-73)^2}{64} = 495 - \frac{5329}{64} = 495 - 83{,}27 = 411{,}73$$

Подставим полученные значения в формулу бисериального показателя связи:

$$r_b = \frac{\frac{\sum p_+ \cdot a}{n_+} - \frac{\sum p_1 \cdot a}{n}}{\sqrt{\frac{\alpha}{n_+} - \frac{\alpha}{n}}} = \frac{\frac{4}{32} - (\frac{-73}{64})}{\sqrt{\frac{411{,}73}{32} - \frac{411{,}73}{64}}} = \frac{0{,}12 - (-1{,}14)}{\sqrt{12{,}87 - 6{,}43}} = \frac{0{,}12 + 1{,}14}{\sqrt{6{,}44}} = \frac{1{,}26}{2{,}54}$$
$$= 0{,}50$$

Рассчитанный бисериальный коэффициент показывает, что связь полной длины тела байкальского омуля селенгинской и северобайкальской рас возраста 7+ существует и близка к умеренной.

Полихорический показатель связи ρ

В практике биологических исследований встречаются случаи, когда полученные результаты не могут быть выражены численно (окраска, интенсивность освещения, состояние развития и т.д.). Для определения степени сопряженности между качественными признаками с числами вариант, большими двух, служит полихорический показатель связи или коэффициент взаимной сопряженности, предложенный К. Пирсоном и усовершенствованный А.А. Чупровым.

$$К_ч = \sqrt{\frac{\varphi^2}{(n_x - 1) \cdot (n_y - 1)}},$$

где $К_ч$ – коэффициент взаимной сопряженности Чупрова;

n_x и n_y – численность групп по строка и столбцам многопольной таблицы;

φ^2 – показатель взаимной сопряженности, определяемый, как сумма отношений квадратов частот каждой клетки таблицы к произведению итоговых частот соответствующего столбца и строки минус 1:

$$\varphi^2 = \sum \frac{n_{xy}^2}{n_x \cdot n_y} - 1$$

Разберем на примере вычисление коэффициента взаимной сопряженности между размером и жирностью у леща:

Размер	Жирность			Всего (n_x)
	Жирные	Средней жирности	Не жирная	
Крупные	45	10	2	57
Средние	12	28	4	44
Мелкие	7	8	15	30
Всего (n_y)	64	46	21	131

Определим коэффициент взаимной сопряженности между этими признаками, предварительно рассчитав величину φ^2:

$$\varphi^2 = \sum \frac{n_{xy}^2}{n_x \cdot n_y} - 1$$

$$= \frac{45^2}{57 \cdot 64} + \frac{10^2}{57 \cdot 46} + \frac{2}{57 \cdot 21} + \frac{12^2}{44 \cdot 64} + \frac{28^2}{44 \cdot 46} + \frac{4^2}{44 \cdot 21} + \frac{7^2}{30 \cdot 64} + \frac{8^2}{30 \cdot 46} + \frac{15^2}{30 \cdot 21} - 1$$

$$= \frac{2025}{3648} + \frac{100}{2622} + \frac{4}{1197} + \frac{144}{2816} + \frac{784}{2024} + \frac{16}{924} + \frac{49}{1920} + \frac{64}{1380} + \frac{225}{630} - 1$$

$$= 0{,}555 + 0{,}038 + 0{,}003 + 0{,}051 + 0{,}387 + 0{,}017 + 0{,}026 + 0{,}046 + 0{,}357 - 1 = 1{,}48 - 1 = 0{,}48$$

Подставим найденные значения в формулу коэффициент взаимной сопряженности Чупрова:

$$К_ч = \sqrt{\frac{\varphi^2}{(n_x - 1) \cdot (n_y - 1)}} = \sqrt{\frac{0{,}48^2}{(3 - 1) \cdot (3 - 1)}} = \sqrt{\frac{0{,}23}{4}} = \sqrt{0{,}058} = 0{,}24$$

Рассчитанный коэффициент указывает на наличие слабой связи между размером и жирностью у леща.

Коэффициенты корреляции для альтернативных признаков r_a

Для измерения тесноты зависимости альтернативных признаков или показателей, имеющих лишь два возможных различных значения используют коэффициенты ассоциации Юла ($r_{ас.}$) (разработан британскими статистиками Джорджем Эдни Юлом и Морисом Джорджем Кендалом) и контингенции Пирсона ($r_{кон.}$) (предложен К. Пирсоном). Данные коэффициенты принимают значения от -1 до +1. Отрицательные значения свидетельствуют об обратном направлении связи. Нужно иметь в виду, что для одних и тех же данных коэффициент контингенции всегда меньше коэффициента ассоциации ($r_{кон.} < r_{ас.}$). Если $r_{кон} \geq 0{,}3$, а $r_{ас.} \geq 0{,}5$, то это указывает о наличии связи.

При вычислении коэффициентов корреляции для альтернативных признаков строят четырехпольную корреляционную решетку, в которой два класса будут по одному признаку и два – по другому признаку.

Формула коэффициента ассоциации Юла:

$$r_{ас} = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c}$$

Формула коэффициента контингенции Пирсона:

$$r_{кон} = \frac{a \cdot d - b \cdot c}{\sqrt{(a + b) \cdot (c + d) \cdot (a + c) \cdot (b + d)}}$$

Где a, b, c, d– частоты в каждой клетке корреляционной решетки.

Рассмотрим это на конкретном примере:

Для борьбы с триходиниозом у мальков карпа в аквакультуре использовали солевые ванны с 1,5% раствором поваренной соли при температуре воды 21-30^{0}С. Необходимо определить имеется ли связь между данным мероприятием и числом заболевших особей по выборке из 500 рыб.

	Заболевших триходиниозом	Не заболевших триходиниозом	$\sum$
Обработанные мальки	40 (a)	210(b)	250
Не обработанные мальки	220 (c)	30 (d)	250
$\sum$	260	240	500

$$r_{ac} = \frac{a \cdot d - b \cdot c}{a \cdot d + b \cdot c} = \frac{40 \cdot 30 - 210 \cdot 220}{40 \cdot 30 + 210 \cdot 220} = \frac{1200 - 46200}{1200 + 46200} = \frac{-45000}{47400} = -0{,}949$$

$$r_{кон} = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (c+d) \cdot (a+c) \cdot (b+d)}}$$

$$= \frac{40 \cdot 30 - 210 \cdot 220}{\sqrt{(40+210) \cdot (220+30) \cdot (40+220) \cdot (210+30)}}$$

$$= \frac{1200 - 46200}{\sqrt{250 \cdot 250 \cdot 260 \cdot 240}} = \frac{-45000}{\sqrt{3900000000}} = \frac{-45000}{62449{,}98} = -0{,}721$$

В нашем примере оба коэффициента характеризуют достаточно большую обратную зависимость между исследуемыми признаками. Однако коэффициент контингенции показывает менее тесную связь, чем коэффициент ассоциации, но он считается более точным. Поэтому можно сделать вывод о том, что между мероприятием для борьбы с заболеваемостью мальков карпа с помощью раствора поваренной соли и числом пораженных особей триходиниозом существует значительная обратная зависимость.

Множественный и частный коэффициент корреляции

Выше изложены способы вычисления связи между двумя признаками. Но часто представляет большой интерес вопрос выявления связь между тремя и большим числом факторов. Например, изучая популяцию какого-то вида рыб необходимо знать зависимость ее продуктивности не только от возрастного состава, но и от соотношения полов, кормовой базы, условий обитания и т.д. Следовательно, показатель множественной корреляции характеризует тесноту связи рассматриваемого набора факторов с исследуемым признаком, или, иначе, оценивает тесноту совместного влияния факторов на результат.

Множественная корреляция решает три задачи. Она определяет:

- форму связи;
- тесноту связи;
- влияние отдельных факторов на общий результат.

При множественной корреляции можно вычислить свободный коэффициент корреляции, который служит мерой силы связи между признаком *z* и признаками *x* и *y*, определяющими изменения *z*.

Формула свободного коэффициента связи:

$$R_{св\,(yxz)} = +\sqrt{\frac{r_{xy}^2 + r_{yz}^2 - 2r_{yx} \cdot r_{yz} \cdot r_{xz}}{1 - r_{xz}^2}},$$

где r_{xy}, r_{yz}, r_{xz} – парные коэффициенты линейной корреляции между признаками.

Свободный коэффициент $R_{св}$ – число всегда положительное и изменяется от 0 до 1.Чем он ближе $R_{св}$ к 1, тем в большей мере учтены факторы, определяющие конечный результат. Если $R_{св}$=0, то признак *z* не имеет линейной связи с признаками *x* и *y*. Если же $R_{св}$=1, то связь между *z* и *x* и *y* линейная.

Свободный коэффициент используется редко, чаще требуется определить частные коэффициенты корреляции, которые позволяют выделять влияние каждого фактора из числа нескольких действующих.

Для определения частного коэффициента корреляции между признаками *x* и *y* при постоянной величине признака *z* применяют формулу:

$$R_{xy(z)} = \frac{r_{xy} - r_{xz} \cdot z_{yz}}{+\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

Заключение знака *z* в скобки обозначает, что влияние признака *z* на корреляцию между *x* и *y* исключено.

Также вычленяют действие *x* на фактор *z* при изоляции фактора *y*:

$$R_{xz(y)} = \frac{r_{xz} - r_{xy} \cdot z_{zy}}{+\sqrt{\left(1 - r_{xy}^2\right)(1 - r_{zy}^2)}}$$

Выявляется влияние *y* на *z* при постоянстве *x* выявляют по аналогичной формуле:

$$R_{yz(x)} = \frac{r_{yz} - r_{yx} \cdot z_{zx}}{+\sqrt{\left(1 - r_{yx}^2\right)(1 - r_{zx}^2)}}$$

Частные коэффициенты корреляции также измеряют от -1 до +1.

Сравнение частных коэффициентов корреляции друг с другом позволяет ранжировать факторы по тесноте их связи с результатом, то есть они показывают меру тесноты связи каждого фактора с результатом в чистом виде.

Регрессия

Коэффициент корреляции указывает лишь на степень (тесноту) связи в изменчивости двух переменных величин, но не позволяет судить о том, как меняется одна величина по мере изменения другой. Ответ на этот вопрос дает регрессивный анализ – статистический метод исследования взаимосвязи переменных.

Коэффициент регрессии (R) – величина именованная и показывает, насколько изменяется в среднем признак x, если коррелирующий с ним признак y изменяется на определенную величину, точнее, на какую величину один признак отклоняется от средней при одновременном отклонении другого от своей средней.

Вычисление коэффициента регрессии позволяет:

- определить, насколько изменяется одна величина относительно другой;
- прогнозировать результаты.

Формула коэффициента регрессии включает в себя коэффициент корреляции и средние квадратичные отклонения по обоим признакам и выражается следующим образом:

$$R_{xy} = r\frac{\delta_x}{\delta_y} \text{ и } R_{yx} = r\frac{\delta_y}{\delta_x}$$

Из формул видно, что коэффициенты регрессии имеют тот же знак, что и коэффициент корреляции, размерность, равную отношению размерностей изучаемых показателей x и y, и связаны соотношением:

$$R_{xy} \cdot R_{yx} = r^2$$

Рассмотрим определение коэффициентов регрессии на примере, где мы вычисляли корреляционную связь между измерениями промысловой длины (x) и массы тела (y) у байкальского омуля (n=35).

Связь, выраженная через коэффициент корреляции равна:

$$r = +0{,}93$$

$$\delta_x = 2{,}81 \text{ (мм)}; \ \delta_y = 2{,}45 \text{ (г)}$$

Подставим эти значения в формулу коэффициента регрессии:

Регрессия длины тела омуля от его массы составляет:

$$R_{xy} = r\frac{\delta_x}{\delta_y} = 0{,}93 \cdot \frac{2{,}81}{2{,}45} = 0{,}93 \cdot 1{,}15 = 1{,}07 \text{ (мм)}$$

То есть увеличение массы омуля на 1 г ведет к увеличению его промысловой длины тела на 1,07 мм.

Обратная регрессия влияния длины на массу:

$$R_{yx} = r\frac{\delta_y}{\delta_x} = 0{,}93 \cdot \frac{2{,}45}{2{,}81} = 0{,}93 \cdot 0{,}87 = 0{,}81 \text{ (г)}$$

Увеличение промысловой длины тела омуля на 1 мм, повышает массу на 0,81 г.

Выборочные показатели регрессии являются оценками соответствующих генеральных параметров и, как величины случайные, сопровождаются статистическими ошибками. При этом, показатели связи имеют реальный смысл, если они оказываются статистическими достоверными. Практическое же значение они приобретают лишь тогда, когда имеют достаточную величину.

Выбор формулы для расчета ошибки коэффициента регрессии зависит от величины исследованной выборки:

$$m_R = \sqrt{\frac{1 - R^2}{n - 2}}, \text{если } n < 30$$

$$m_R = \sqrt{\frac{1 - R^2}{n}}, \text{если } n \geq 30$$

В нашем случае n=35, следовательно, воспользуемся формулой для большой выборки:

$$m_{R_{xy}} = \sqrt{\frac{1 - R_{xy}^2}{n}} = \sqrt{\frac{1 - 1{,}07^2}{35}} = \sqrt{\frac{1 - 1{,}14}{35}}\sqrt{\frac{-0{,}14}{35}} = \sqrt{0{,}004} = 0{,}063(\text{мм})$$

$$m_{R_{yx}} = \sqrt{\frac{1 - R_{yx}^2}{n}} = \sqrt{\frac{1 - 0{,}81^2}{35}} = \sqrt{\frac{1 - 0{,}66}{35}}\sqrt{\frac{0{,}34}{35}} = \sqrt{0{,}0098} = 0{,}099(\text{г})$$

Критерий достоверности определяют по стандартной формуле, как частное от деления оцениваемого показателя к его ошибке:

$$t_{R_{xy}} = \frac{R_{xy}}{m_{xy}} = \frac{1{,}07}{0{,}063} = 16{,}98$$

$$t_{R_{yx}} = \frac{R_{yx}}{m_{xy}} = \frac{0{,}81}{0{,}099} = 8{,}18$$

Полученные критерии достоверности сравнивают со стандартами значений по таблице Стьюдента (табл. 6 приложения) для установленного числа степеней свободы и порога вероятности безошибочных прогнозов равным n-2.

Вычисленные значения t больше табличных при уроне вероятности P=0,999 ($t_{тб.}$=3,7), следовательно, полученные коэффициенты регрессии верно отражают зависимость исследованных факторов.

Дисперсионный анализ

Дисперсионный анализ составляет своеобразный раздел биометрии, разработанный Р. Фишером.

При изучении и анализе сложных и многообразных причинноследственных отношений между объектами и явлениями ихтиологу приходится учитывать целый комплекс внешних и внутренних факторов, от которых, в конечном итоге зависят уровень и ход наблюдаемых процессов, те или иные свойства живых организмов, их динамика и разнообразие. При этом важно оценивать не только значение одного из факторов, но и их взаимодействие при сопряженном влиянии на популяцию и организм. Решение подобных задач с помощью корреляционного или регрессивного анализа, как правило, не дает удовлетворительного результата. Гораздо более удобным и совершенным статистическим приемом, позволяющим охватить весь комплекс наблюдений и процесс в целом и обладающим рядом других существенных достоинств, является метод дисперсионного анализа, который строится на обработке выборки, полученной по принципу случайного отбора объектов, но при этом допускает малочисленность материала и его качественную разнородность.

При дисперсионном анализе обработке подвергаются выборочные данные, сведенные в статистический комплекс, который оформляется в виде таблицы, состоящей из граф и строчек, по клеткам которой размещают сведения со значениями варьирующего признака. В этой части он схож с корреляционной таблицей.

Основное назначение дисперсионного анализа состоит в том, что он позволяет выявить статистически влияние различных факторов на изменчивость изучаемого признака. При этом можно определять как влияние каждого фактора в отдельности, так и суммарное их воздействие, приводящее к определенной изменчивости в величине данного показателя.

Общую изменчивость или дисперсию выражают путем суммирования квадратов отклонений каждого варианта от средней арифметической, то есть в виде:

$$\sum(\mathrm{v} - \mathrm{M})^2 = \mathrm{C_y}$$

Это выражение называют общей дисперсией признака.

Общая дисперсия C_y может быть разложена на составные части:

- C_x – дисперсия, возникающая под влиянием различных учтенных факторов – факторная дисперсия;
- C_z – дисперсия, возникающая под влиянием различных случайных (неучтенных) факторов – остаточная дисперсия.

Следовательно, в общей форме дисперсия, то есть разнообразие и изменчивость любого признака, может быть записана как:

$$C_y = C_x + C_z$$

В задачу дисперсионного анализа входят вычисления и определение величины факторной (C_x) и остаточной (Cz) дисперсий.

В самом общем виде факторная дисперсия может быть представлена как сумма квадратов разностей между частными средними значениями признака (Мх), полученных в графиках статистического комплекса по классам действующих факторов и общей арифметической (Мобщ.), вычисленной для всего статистического комплекса из показателей варьирующего признака.

Это можно выразить в следующем виде:

$$C_x = \sum(\text{Мчастн.} - \text{Мобщ.})^2, \quad \text{или}$$

$$C_x = \sum n_x\,(\text{Мчастн} - \text{Мобщ.})^2,$$

где n_x – число наблюдений в каждом классе фактора

Случайную, или остаточную дисперсию (Cz) можно вычислить через сумму квадратов разностей варьирующего признака (v) по отношению к частной средней арифметической (Мчастн.).

$$C_z = \sum(\text{v} - \text{Мчастн})^2$$

Если ведут изучение изменчивости признака под влиянием нескольких факторов (размер – А, возраст – В), то факторная дисперсия (Сх) может быть представлена суммой из дисперсий каждого фактора раздельно (А и В) и дисперсии совместного влияния обоих факторов (АВ).

Это выражается в следующей форме:

$$C_x = C_A + C_B + C_{AB}$$

Общая дисперсия будет выражена такой суммой:

$$C_y = C_A + C_B + C_{AB} + C_Z$$

Куда входят частные факторные и остаточная дисперсии.

Чем меньше величина факторных дисперсий и чем больше величина остаточной дисперсии, тем меньше познана изменчивость изучаемого признака при помощи дисперсионного анализа.

С помощью дисперсионного анализа можно вычислить долю или степень влияния Сх и Cz на варьирующий признак. Для этого берут отношение между дисперсиями и обозначают эти отношения через (η^2).

Доля влияния всех учтенных факторов на изменчивость признака выразится формулой:

$$\eta_x^2 = \frac{Cx}{Cy},$$

а доля влияния неучтенных факторов выразится формулой

$$\eta_z^2 = \frac{Cz}{Cy}$$

Если общую изменчивость принять за 1 или за 100%, то доли ее будут составлять влияние каждого фактора на изменчивость признака:

$$\eta_y^2 = \eta_x^2 + \eta_z^2 = 1$$

Извлечение квадратного корня из этих выражений дает величину корреляционного отношения:

$$\eta_x = \sqrt{\frac{Cx}{Cy}} \text{ и } \eta_z = \sqrt{\frac{Cz}{Cy}}$$

Таким образом, в ходе дисперсионного анализа можно получить коэффициенты связи, не проводя специальной обработки выборочного материала.

Дисперсионный анализ осуществляют в несколько этапов путем обработки таблицы статистического комплекса и составлением сводной таблицы дисперсионного анализа.

Первый этап дисперсионного анализа заключается в обработке статистического комплекса для получения общей, факторных и остаточной дисперсий (Cy, Cx, Cz). Второй этап состоит в вычислении долей каждой частной дисперсии в общей дисперсии, для чего рассчитывают величины η_x^2 и η_z^2.Третий этап сводится к корректированию полученных дисперсий на число степеней свободы, вычисленных для каждой дисперсии по определенным формулам. Корректированные дисперсии (или девианты) обозначают через дисперсию (δ^2)и вычисляют по формулам:

Корректированную общую дисперсию с учетом числа степеней свободы v_y определяют как:

$$\delta_y^2 = \frac{Cy}{v_y}$$

Корректированную факторную дисперсию с учетом числа степеней свободы v_x:

$$\delta_x^2 = \frac{Cx}{v_x}$$

Корректированную остаточную дисперсию с учетом числа степеней свободы vz:

$$\delta_z^2 = \frac{Cz}{v_z}$$

Если из корректированных дисперсий извлечь квадратный корень, то получим величины средних квадратичных отклонений (δ):

$$\delta_y = \sqrt{\frac{Cy}{v_y}} = \sqrt{\frac{\sum(v - \text{Мобщ.})^2}{v_y}}$$

$$\delta_x = \sqrt{\frac{Cx}{v_x}} = \sqrt{\frac{\sum(\text{Мчастн.} - \text{Мобщ.})^2}{v_x}}$$

$$\delta_z = \sqrt{\frac{Cz}{v_z}} = \sqrt{\frac{\sum(v - \text{Мобщ.})^2}{v_z}}$$

Четвертый этап дисперсионного анализа дает суждение о том, достоверно ли значение факторной дисперсии, то есть достоверно ли влияние данного фактора (x или A, B, AB и т.п.) на варьирующий признак.

Для этой цели используют коэффициент Фишера (F), который получают в результате деления факторных дисперсий на остаточную:

$$F = \frac{\delta_x^2}{\delta_z^2}\ ;\ F = \frac{\delta_A^2}{\delta_z^2};\ F = \frac{\delta_B^2}{\delta_z^2};\ F = \frac{\delta_{AB}^2}{\delta_z^2}$$

Для суждения о достоверности факторных дисперсий сравнивают вычисленное значение Fвычисл. с величиной Fтабл., которую определяют по специальным таблицам Фишера (табл. приложения 8,9).

Если вычисленное значение F окажется больше или равным табличному значению, то дисперсия и влияние данного фактора считают достоверными.

Типы статистических комплексов

Рабочие формулы и техника обработки выборки при дисперсионном анализе зависят от того, большая или малая выборка подвергается обработке, а также от структуры статистического комплекса.

Статистические комплексы различают по тому, сколько факторов включено в каждом из них для изучения дисперсии.

Статистические комплексы бывают:

- однофакторными;
- двухфакторными;
- трехфакторными;

- с большим числом факторов.

Статистические комплексы различают между собой еще и по соотношению частот в классах факторов, входящих в них.

Статистические комплексы, имеющие больше одного фактора, бывают:

- равномерными;
- пропорциональными;
- неравномерными.

В равномерных комплексах число наблюдений по классам факторов одинаковое и их отношения равны 1:1:1 и т.д.

В таблице приведен равномерный статистический комплекс, в котором рассматривается влияние двух факторов: вес тела речного окуня (А) и его возраст (В) на показатель веса мозга (v)

Выборка имела малое число наблюдений:

Структура статистического комплекса включает три класса по фактору А-вес тела (г): A_1 –0-100 г; A_2 –101-200 г; A_3 –201-300 г.

И два класса по фактору В (возраст): B_1 –3+; B_2 –4+.

В каждый класс фактора А входят факторы B_1 и B_2.

Таким образом, в комплексе имеется 6 классов, или градаций.

Исследованные 18 особи окуня распределены равномерно, в каждом классе по 3 шт.; следовательно, комплекс имеет равномерный, двухфакторный тип.

Равномерный статистический комплекс при малом числе наблюдений (n=18)
(фактор А – вес тела, фактор В – возраст)

Фактор А (вес тела)	A_1 =0-50 г		A_2= 51-100 г		A_3=101-150 г	
Фактор В (возраст)	B_1-3+	B_2-4+	B_1-3+	B_2-4+	B_1-3+	B_2-4+
Варьирующий признак v (вес мозга, мг)	102	109	138	155	189	213
	104	126	115	167	169	202
	109	145	137	143	177	216
Частоты р или n_x	3	3	3	3	3	3

В каждом классе таблицы строчки (v) проставлены сведения о весе мозга раздельно по двум показателя фактора В: возраст 3+ и возраст 4+.

Равномерный комплекс является частным случаем пропорционального. Соотношение частот в классах В для каждого класса А одно и то же и составляет 1:2, хотя число наблюдений по классам различное.

Пропорциональный двухфакторный комплекс при малом числе наблюдений(n=21)

Фактор А (вес тела)	A_1 =0-50 г		A_2= 51-100 г		A_3=101-150 г	
Фактор В (возраст)	B_1-3+	B_2-4+	B_1-3+	B_2-4+	B_1-3+	B_2-4+
Варьирующий признак v (вес мозга, мг)	102	109	138	155	189	213
	104	126	115	167	169	202
		145	137	143		216
		129		148		206
				136		
				140		
Частоты р или n_x	2	4	3	6	2	4
Отношение частот В по классам А	1:2		1:2		1:2	

В равномерных и пропорциональных комплексах сумма частных дисперсий равна общефакторной дисперсии, то есть

$$C_A + C_B + C_{AB} = C_x$$

Остановимся на схеме неравномерного комплекса. Распределение частот фактора В по классам фактора А неравномерное: в классе A_1 оно равно 1:2, в классе A_2 – 1:1, а классе A_3 – 1:1.

Неравномерный двухфакторный комплекс при малом числе наблюдений (n=22)

Фактор А (вес тела)	A_1 =0-50 г		A_2= 51-100 г		A_3=101-150 г	
Фактор В (возраст)	B_1-3+	B_2-4+	B_1-3+	B_2-4+	B_1-3+	B_2-4+
Варьирующий признак v (вес мозга, мг)	102	109	138	155	189	189
	104	126	115	167	169	169
		145	137	143		189
		129	129	148		169
			116	136		
Частоты р или n_x	2	4	5	5	2	4
Отношение частот В по классам А	1:2		1:1		1:2	

Рабочие формулы и техника обработки статистического комплекса меняется в зависимости от его типа.

Обработка однофакторного комплекса при малом числе наблюдений

Однофакторные комплексы не бывают неравномерными, так как в их структуре представлены классы только по одному фактору.

Для вычисления общей дисперсии пользуются следующими рабочими формулами:

$$C_y = \sum v^2 - \frac{(\sum v)^2}{n}, \qquad \text{или } C_y = \sum v^2 - H,$$

где v – величина варьирующего признака.

Для удобства выражение $\frac{(\sum v)^2}{n}$ обозначают через Н.

Остаточную дисперсию вычисляют по следующей формуле:

$$C_z = \sum v^2 - \sum h_x,$$

где

$$\sum h_x = \frac{(\sum v_x)^2}{n_x},$$

$\sum v_x$ - получают от суммирования варьирующего признака по каждому классу изучаемого фактора;

n_x – число наблюдений по каждому классу изучаемого фактора.

Факторную дисперсию вычисляют по формуле:

$$C_x = \sum h_x - H$$

Разберем пример обработки однофакторного комплекса. Исходные данные и техника вычисления приведены в таблице:

Фактор А – вес тела речного окуня (г): A_1 – 0-50; A_2 – 51-100; A_3 –101-150. Варьирующий признак вес мозга речного окуня (мг).

Фактор А (вес тела)	Классы по фактору А			Сводные показатели
	A_1 =0-50 г	A_2= 51-100 г	A_3=101-150 г	
Варьирующий признак v (вес мозга, мг)	102, 104, 109, 126,145, 129, 108, 113	138, 115, 137, 155, 167, 143, 148, 136	189, 169, 213, 202, 216, 206, 184, 208	$\sum v$=3662
v^2	10404, 10816, 11881, 15876, 21025, 16641, 11664, 12769	19044, 13225, 18769, 24025, 27889, 20449, 21904, 18496	35721, 28561, 45369, 40804, 46656, 42436, 33856, 43264	$\sum v^2$= 591544
n_x	8	8	8	$\sum n_x$=24
$\sum v_x$	936	1139	1587	$\sum v_x$=3662
$(\sum v_x)^2$	876096	1297321	2518569	-
$h_x = \frac{(\sum v_x)^2}{n_x}$	$\frac{876096}{8} = 109512{,}0$	$\frac{1297321}{8} = 162165{,}1$	$\frac{2518569}{8} = 314821{,}1$	$\sum h_x$=586498,2
$M_x = \frac{\sum v_x}{n_x}$	$\frac{936}{8} = 117{,}0$	$\frac{1139}{8} = 142{,}4$	$\frac{1587}{8} = 198{,}4$	$M_{общ} = \frac{\sum v}{n} = \frac{3662}{24} = 152{,}6$

$$H = \frac{(\sum v)^2}{n} = \frac{3662^2}{24} = \frac{13410244}{24} = 558760{,}2$$

Суммарное значение по этой строчке дает $\sum v^2$, входящее в формулу общей и остаточной дисперсии.

Строчку n_x составляют из числа наблюдений по каждому классу. $\sum v_x$ образуют путем суммирования величин веса мозга в каждом классе.$(\sum v)^2$ получают возведением в квадрат данных по каждому классу из предыдущей строчки. Строчку h_x составляют для каждого класса из соотношений данных строчек, а именно $(\sum v_x)^2$ и n_x. Суммарное значение по этой строчке дает величину $\sum h_x$, входящую в формулу факторной дисперсии. Последнюю строчку комплекса составляют для получения частных средних арифметических для каждого класса и общей средней арифметической.

После обработки статистического комплекса можно приступать к дисперсионному анализу.

Методом дисперсионного анализа определим достоверность и долю влияния повеса тела по классам A_1, A_2, A_3на вес мозга.

Для этого вычислим дисперсии Cy, Cx, Cz.

$$C_y = \sum v^2 - H = \sum v^2 - \frac{(\sum v)^2}{n} = 591544 - 558760{,}2 = 32783{,}8$$

$$C_x = \sum h_x - H = 586498{,}2 - 558760{,}2 = 27738$$

$$C_z = \sum v^2 - \sum h_x = 591544 - 586498{,}2 = 5045{,}8$$

Для проверки правильности расчетов произведем суммирование:

$$C_y = C_x + C_z, \quad \text{то есть} \quad 32783{,}8 = 27738 + 5045{,}8$$

Исходя из полученных данных о дисперсиях, составим сводную таблицу дисперсионного анализа. В сводной таблице выделяют графы для факторной, остаточной и общей дисперсии, что обозначается заголовками x, z и y.

Построчно проведем дальнейшую вычислительную работу. В первой строчке записывают уже вычисленные значения дисперсий. Вторая строка отводится для вычисления значений η^2, то есть выявлению доли (или проценту) влияния изучаемого фактора x и неучтенных факторов z на изменчивость признака v. Для этого вычисляют отношение каждой дисперсии к общей дисперсии. Для проверки правильности вычисления поводят суммирование.

Далее в третьей строчке определяют число степеней свободы v. Для Cx число степеней свободы равно числу классов *l* по фактору А минус единица. Для остаточной дисперсии Cz число степеней свободы определяют по разнице между числом наблюдений n и числом классов *l*. Число степеней свободы для общей дисперсии равно числу наблюдений n без единицы.

После этих вычислений можно рассчитать корректированную дисперсию (или девиату) δ^2. Для этого каждую дисперсию (факторную и остаточную) делят на соответствующее число классов степеней свободы v.

Дисперсии С	x	z	y
	27738	5045,8	32783,8
Степень влияния фактора x и z на С η^2	$\eta_x^2 = \frac{C_x}{C_y} = \frac{27738}{32783,8} = 0,846 = 84,6\%$	$\eta_z^2 = \frac{C_z}{C_y} = \frac{5045,8}{32783,8}$ =0,154 = 15,4%	$\eta_x^2 + \eta_z^2 = \eta_y^2 = 1$
Число степеней свободы v	$v_x = l_x - 1 = 3 - 1 = 2$	$v_z = n - l_x = 24 - 3 = 21$	$v_y = n - 1 = 24 - 1 = 23$
Корректированная дисперсия δ^2	$\delta_x^2 = \frac{C_x}{v_x} = \frac{27738}{2} = 13869$	$\delta_z^2 = \frac{C_z}{v_z} = \frac{5045,8}{21} = 240,28$	-
Коэффициент достоверности F	$\frac{\delta_x^2}{\delta_z^2} = \frac{13869}{240,28} = 57,7$	-	-
Табличное значение F при v_z=21, v_x=2	При 0,95=3,5 При 0,99=5,8 При 0,999=9,8	-	-

Последний этап дисперсионного анализа заключается в определении достоверности факторной дисперсии, то есть достоверно ли влияние и доля влияния фактора на изменчивость признака. Для этого вычисляют коэффициент достоверности Фишера (F) путем деления факторной дисперсии на остаточную. Далее сравнивают вычисленное значение F со значением F табличным.

Так как наше вычисленное F равно 57,7, то можно сделать вывод, что влияние веса тела речного окуня на вес мозга достоверно при уровне вероятности 0,999. Причем факторная дисперсия, то есть вес тела речного окуня имеет подавляющее влияние на вес его мозга (84,6%), а на долю неучтенных факторов приходится только 15,4%.

ГЛОССАРИЙ

Абсолютная (индивидуальная) плодовитость – количество икры, отложенное самкой за один нерестовый период.

Антедорсальное расстояние, **антедорсальное пространство** – расстояние от вершины рыла до основания первого луча спинного плавника.

Асимметричные ряды – ряды, для которых характерно, что частоты уменьшаются в одну сторону быстрее, чем в другую, что приводит к смещению вершины кривой в правую или левую сторону от средней арифметической.

Биноминальное распределение – частный случай нормального распределения. Отражает распределение членов совокупности, имеющих альтернативные признаки.

Бисериальный показатель связи – применяют в случаях, когда один признак выражен количественно, а другой имеет качественное и при том альтернативное выражение.

Боковая или латеральная линия – число прободенных чешуи (число трубочек или канальчиков) в боковой части тела.

Вариации – ступени, на которые разбивается весь вариационный ряд.

Вариационная статистика – наука, разрабатывающая методы изучения варьирующего признака на массовых материалах в различных областях знаний.

Вариационный ряд – ряд цифр по величине изучаемого признака, расположенных по возрастающей или убывающей степени с соответствующими им частотами появления признака.

Варьирующие признаки- признаки, проявляющие определенную закономерность в изменчивости (колеблемости) своих значений.

Высота головы у затылка – верхняя точка берется по окончанию черепа, нижняя, противоположная ей – по вертикали.

Горизонтальный (продольный) диаметр глаза – диаметр роговицы, веки, если они есть, в расчет не принимают.

Дисперсия – указывает на степень разнообразия показателя у членов совокупности средней арифметической, вычисленной для данной совокупности.

Дисперсионный анализ – метод, направленный на поиск зависимости данных путем исследования значимости различий средних значений, или анализ изменчивости результативного признака под влиянием каких-либо контролируемых переменных факторов. Строится на обработке выборки, полученной по принципу случайного отбора объектов.

Длина без хвостового плавника – расстояние от начала рыла до конца чешуйного покрова.

Длина верхней лопасти хвостового плавника С – длина наибольшего луча верхней лопасти хвостового плавника.

Длина всей рыбы – общая или абсолютная длина - от вершины рыла до вертикали конца наиболее длинной лопасти хвостового плавника при горизонтальном положении рыбы.

Длина головы – расстояние сбоку от вершины рыла (при закрытом рте) до заднего, наиболее удаленного края жаберной крышки (без жаберной перепонки).

Длина основанияА – длина анального плавника - расстояние от основания переднего, луча до основания последнего луча анального плавника.

Длина основания D – длина спинного (дорсального) плавника – расстояние от основания переднего, луча до основания последнего луча, или до конца перепонки спинного плавника. Если спинных плавников не один – расстояние до конца последнего луча каждого плавника.

Длина Р – длина грудных плавников – расстояние от передней линии их прикрепления до вершины наиболее длинного луча.

Длина V – длина брюшных плавников – расстояние от передней линии их прикрепления до вершины наиболее длинного луча.

Длина нижней лопасти С – длина наибольшего луча нижней лопасти хвостового плавника.

Длина по Смитту – расстояние от переднего края рыла до конца средних лучей.

Длина рыла – предглазничный отдел, предглазничное пространство головы, предглазье – расстояние от вершины рыла до переднего края глаза, до переднего наружного края глазного яблока.

Длина туловища – расстояние от жаберной щели до конца чешуйного покрова или до корней средних лучей у рыб без чешуи.

Длина хвостового стебля – расстояние от вертикали заднего края основания анального плавника до основания хвостового плавника или до конца чешуйного покрова по середине тела рыбы.

Дорсальный плавник – спинной плавник.

Заглазничный отдел головы – заглазничное пространство – расстояние от заднего края глаза до наиболее удаленной от конца рыла точки жаберной крышки. Жаберная перепонка, окаймляющая сзади жаберную крышку, в расчет не принимается.

Коэффициент изменчивости или вариации – изменчивость признака в совокупности в относительных величинах (в процентах).

Коэффициент корреляции-определяет величину и направление связи при прямолинейном ее типе или близком к прямолинейному.

Коэффициент регрессии – величина показывающая, насколько изменяется в среднем признак, если коррелирующий с ним признак изменяется на определенную величину.

Критерий достоверности – показатель того, насколько правильно выборочная средняя отражает генеральную среднюю.

Лимиты – показывают размах значений и тем самым характеризуют разнообразие признака в группе.

Медиана – вариант, значение которого делит всю совокупность наблюдений на две равные части. Одна половина объектов совокупности будет иметь значения варьирующего признака меньше, а другая половина объектов больше чем медиана.

Мода, или модальный вариант – наиболее часто встречающиеся значения варианта.

Наибольшая высота А – высота анального плавника - высота наибольшего луча плавника.

Наибольшая высота D – высота спинного плавника - высота наибольшего луча плавника.

Наибольшая высота тела – расстояние от самой высокой точки спины до брюшка по вертикали. Плавники (у осетровых и костяные щитки) в расчет не входят.

Наибольший обхват тела – место наибольшей толщины и высоты тела без плавников.

Наибольшая толщина тела – наибольшее расстояние между боками.

Наименьшая высота тела – высота хвостового стебля – расстояние около основания хвостового плавника.

Нормированное отклонение – статистический признак, позволяющий определить изменчивость признаков. С его помощью можно выразить в относительных единицах отклонение каждого конкретного члена совокупности.

Отолит – слуховой камешек – твердое в виде зернышек или более крупных частиц.

Ошибка средней арифметической – прямо пропорциональна изменчивости признака и обратно пропорциональна числу наблюдений в выборке.

Плодовитость –количество икринок, отложенных самкой.

Постдорсальное расстояние, **постдорсальное пространство –** расстояние от вертикали заднего конца основания спинного плавника до основания хвостового плавника, считая по середине тела.

Промысловая длина тела рыб – расстояние от конца рыла до заднего края чешу иного покрова или до основания, средних лучей хвостового плавника.

Простая корреляционная связь – связь между двумя признаками, без учета имеющихся других связей.

Распределение Пуассона – случай, когда имеют дело с появлением редких событий при большом числе опытов, то есть когда вероятность появления этого события очень мала.

Расстояние между P и V – расстояние между основаниями грудного и брюшного плавников.

Склериты – известковые образования на чешуйной пластинке.

Средняя арифметическая – величина, сумма отрицательных и положительных отклонений от которой равна нулю.

Средняя арифметическая для альтернативных признаков – показатель доли, которую составляют члены совокупности, имеющие данный альтернативный признак.

Средняя гармоническая – необходима для вычисления средних значений, получаемых во времени.

Средняя геометрическая – средняя величина, которая выявляет средний прирост (или среднее уменьшение) какого-либо показателя за определенный период времени.

Средняя квадратическая – используется для признаков, которые характеризуются площадью круга и для ее получения измеряют величину диаметра.

Среднее квадратичное отклонение – величина для измерения изменчивости как количественных, так и качественных признаков членов совокупности.

Трансгрессивные ряды и кривые – ряды, которые отличаются друг от друга величиной средней арифметической и у которых крайние классы, лежащие около максимального класса первой кривой, служат минимальными классами другой кривой, что создает в этих частях вариационных кривых их взаимное пересечение.

Функциональная связь – связь между какими-либо показателями, когда при изменении одного признака или показателя на определенную величину другой признак или показатель изменяется тоже на определенную величину.

Частные коэффициенты корреляции – позволяют выделять влияние каждого фактора из числа нескольких действующих.

Ширина лба или межглазничное пространство, **межглазничный промежуток –** расстояние между глазами сверху, то есть ширина черепа между глазами.

Эксцессивные вариационные ряды – ряды, у которых значительная доля частот накапливается около варианта, соответствующего средней арифметической.

СПИСОК ЛИТЕРАТУРЫ

1. Абросов Н.С., Дегерменджи А.Г., Косолапова Л.Г. Биофизические принципы организации микроэкосистем – основа эколого-экономического прогнозного моделирования. Красноярск: ИФ, 1985. 41 с.

2. Базыкин А.Д. Математическая биофизика взаимодействующих популяций. М.: Наука, 1985. 181 с.

3. Башаринов А.Е., Флейшман Б.С. Методы статистического последовательного анализа и их приложения. М.: Советское радио, 1962. 352 с.

4. Беляев В.И. Обработка и теоретический анализ океанографических наблюдений. Киев. Наук. Думка, 1973. 295 с.

5. Беляев В.И. Моделирование морских систем. Киев: Наукова думка, 1987. 202 с.

6. Боголюбов А.Г. Столетие биометрии в России// Известия Самарского научного центра Российской академии наук. т.4. №2. 2002. С.189-196.

7. БондаренкоА.С., Жигунов А.В. Статистическая обработка материалов лесоводственных исследований: учебное пособие. СПб.: Изд-во Политехнического университета, 2016. 123 с.

8. Василевич В.И. Статистические методы в геоботанике. Л.: Наука, 1969. 232 с.

9. Вольф В.Г. Статистическая обработка опытных данных. М.: Колос, 1966. 253 с.

10. Глотов Н.В., Животовский Л.А., Хованов Н.В., Хромов-Борисов Н.Н. Биометрия. Л.: Издательство Ленингр. университета, 1982. 264 с.

11. Горстко А.Б. Познакомьтесь с математическими моделями. М.: Знание, 1991. 160 с.

12. Горстко А.Б. Модели управления эколого-экономическими. М.: Наука, 1984. 119 с.

13. Горшенина М.В., Горшенина О.В. Вклад русских ученых в развитие статистики как науки // Молодой ученый. 2012. №12. С. 190-192.

14. Гринин А.С. Математическое моделирование в экологии. М.: Юнити, 2003. 269 с.

15. Дмитриев Е.А. Математическая статистика в почвоведении. М.: Издательство МГУ, 1995. 320 с.

16. Доспехов Б.А. Планирование полевого опыта и статистическая обработка его данных: учебное пособие для высших сельскохозяйственных учебных заведений. М.: Колос, 1972. 207 с.

17. Доспехов Б.А. Методика полевого опыта (с основами статистической

обработки результатов исследований). М.: Агропромиздат, 1985. 351 с.

18. Ефимов В.М., Ковалева В.Ю. Многомерный анализ биологических данных изд.2, испр. и допол. СПб, 2008. 86 с.

19. Жаков Л.А., Меншуткин В.В. Практические занятия по ихтиологии. Ярославль, 1982. 112 с.

20. Животовский Л.А. Популяционная биометрия. М.: Наука, 1991. 270 с.

21. ЖигуновА.В., Маркова И.А., Бондаренко А.С. Статистическая обработка материалов лесокультурных исследований. СПб.: Санкт-Петербургская государственная лесотехническая академия, 2002. 86 р.

22. Зайцев Г.Н. Математическая статистика в экспериментальной ботанике. М. Наука, 1984. 424 с.

23. Зайцев Г.Н. Математика в экспериментальной ботанике. М.: Наука, 1990. 296 с.

24. Зиновьев Е.А., Мандрица С.А. Методы исследования пресноводных рыб. Пермь, 2003. 113 с.

25. Ивантер Э.В. Основы практической биометрии. Петрозаводск: изд-во Карелия, 1969. 96 с.

26. Ивантер Э.В., Коросов А.В. Введение в количественную биологию. Петрозаводск: изд-во ПетрГУ, 2011. 302 с.

27. Ивантер Э.В., Коросов А.В. Элементарная биометрия. Петрозаводск: изд-во ПетрГУ, 2013. 110 с.

28. Колесов Ю.С. Математические модели экологии // Исследования по устойчивости и теории колебаний. Ярославль: ЯрГУ, 1979. С. 3-40.

29. Крюков В.И. Статистические методы изучения изменчивости. Орел: Изд-во ГАУ, 2006. 208 с.

30. Лакин Г.Ф. Биометрия: учебное пособие. М.: Высшая школа, 1990. 351 с.

31. Лапин Н.И. Эмпирическая социология в Западной Европе: Учебное пособие. М.: Издательский дом ГУ ВШЭ, 2004.С.33-36.

32. Левинский В.П. Краткий курс вариационной статистики. Государственное учебно-педагогическое издательство. М., 1935. 153с.

33. Левич А.П. Искусство и метод в моделировании систем: вариационные методы в экологии сообществ, структурные и экстремальные принципы, категории и функторы. М.-Ижевск: Институт компьютерных исследований, 2012. 728 с.

34. Леонтович А.В. Биологическая статистика в применении к сельскому хозяйству. М., 1922. 78 с.

35. Леонтович А.В., Григорьев Г.А., Мандзюк А.И. Вариационная статистика. Государственное издательство колхозной и совхозной литературы М.: Сельхозгиз,1935. 204с.

36. Любищев А.А. Дисперсионный анализ в биологии. М.: Издательство Московского университета, 1986. 200 с.

37. Мальчевский А.С., Полянский Ю.И., Хозяцкий Л.И. Памяти Павла Викторовича Терентьева // Вестник Ленингр. университета. 1971. №.9. С.156-158.

38. Меньшуткин В.В. Метод моделирования в динамике численности рыб. М.: Всесоюзный научно-исследовательский институт морского рыбного хозяйства и океанографии «ВНИРО», 1964. 60 с.

39. Меншиткин В.В. Имитационное моделирование водных экологических систем. СПб.: Наука, 1993. 160 с.

40. Меншиткин В.В. Путь к моделированию в экологи. СПб.: Нестор-История, 2007. 394 с.

41. Меркурьева Е.К. Биометрия в животноводстве. М: Колос, 1964. 311 с.

42. Минкевич И.И., Захарова Т.И. Математические методы в фитопатологии. Л.: Колос, 1977. 48 с.

43. Митропольский А.К. Методы статистических вычислений. М.: Наука. 1971. 576 с.

44. Моисеев Н.Н., Александров В.В., Тарко А.М. Человек и биосфера: Опыт системного анализа и эксперименты с моделями. М.: Наука, 1985. 271 с.

45. Молчанов А.М. Математические модели в экологии. Роль критических режимов. Пущино-на-Оке: НЦБИ АН СССР, 1973. 11 с.

46. Никольский Г.В. О биологических основах математического моделирования динамики популяций рыб // Вопросы ихтиологии. Т.3. Вып.2. 1963. С 591-610.

47. Пасеков В.П. О теоретических проблемах биометрического и причинного подходов в популяционных исследованиях. М.: ВЦ им. А.А. Дородницына РАН, 2005. 63 с.

48. Плохинский Н.А. Дисперсионный анализ Новосибирск: Изд-во СО АН СССР, 1960. 124 с.

49. Плохинский Н.А. Биометрия. М.: Изд-во МГУ, 1970. 367 с.

50. Плохинский Н.А. Алгоритмы биометрии. М.: МГУ, 1980. 150 с.

51. Полуэктов Р.А., Пых Ю.А., Швытов И.А. Динамические модели экологических систем. Ленинград: Наука, 1980. 288 с.

52. Поморский Ю.Л. Методы биометрических исследований. Л.:

Ленинградское областное издательство, 1935. 400 с.

53. Поморский Ю.Л. Статистический анализ комплексов признаков (конспект лекций). М: Изд. Всесоюзного института защиты растений ВАСХНИЛ, 1938. 80 с.

54. Поморский Ю.Л. Новейшие методы вариационной статистики. Ленинградский областной научно-исследовательский институт охраны здоровья детей и подростков. Л, 1939. 308 с.

55. Поморский Ю.Л. Методы статистического анализа экспериментальных данных. Методическое руководство для научных работников и аспирантов. Л: ГУЗ Наркомпрос СССР, 1940. 174 с.

56. Правдин И.Ф. Руководство по изучению рыб. М.: Пищепромиздат, 1966. 376 с.

57. Ратнер В.А. Математические модели в популяционной генетике: частотные детерминированные модели // Итоги науки и техники. Сер. Математические методы в биологии. М.: ВИНИТИ, 1969. С. 88–115.

58. Рождественский А.В Чеботарев А.И. Статистические методы в гидрологии. Л.: Гидрометиздат, 1974. 424 с.

59. Рождественский А.В., Ежов А. В. Оценка точности гидрологических расчетов. Л.: Гидрометеоиздат. 1986. 277 с.

60. Розенберг Г.С. Модели в фитоценологии. М.: Наука, 1984. 265 с.

61. Рокицкий П.Ф. Основы вариационной статистики для биологов: Изд-во Белгосуниверситета, 1961. 220 с.

62. Рокицкий П.Ф. Биологическая статистика. Изд. 3-е, испр. Минск: Вышэйш. школа, 1973. 320 с.

63. Рокицкий П.Ф. Введение в статистическую генетику. Минск: Вышэйшая школа, 1974. 448 с.

64. Рыжков Л.П., Дзюбук И.М., Кучко Т.Ю. Ихтиологические исследования на водоемах: учебное пособие для студентов эколого-биологического и агротехнического факультетов. Петрозаводск: Изд-во ПетрГУ, 2013. 72 с.

65. Сапегин А.А Вариационная статистика. М-Л: Госиздат, 1929. 135 с.

66. Сапегин А.А. Вариационная статистика. Практическое элементарное пособие для опытников. М.: Государственное издательство колхозной и совхозной литературы Сельхозгиз, 1935. 93 с.

67. Сапегин А.А. Вариационная статистика. Практическое элементарное пособие для агрономов, опытников и биологов. М.: Государственное издательство колхозной и совхозной литературы Сельхозгиз, 1937. 88 с.

68. Свирежев Ю.М., Логофет Д.О. Устойчивость биологических сообществ. М.: Мир, 1983. 319 с.

69. Сиделев С.И. Математические методы в биологии и экологии: введение в элементарную биометрию: учебное пособие. Ярославль: ЯрГУ, 2012. 140с.

70. Терентьев П.В., Ростова Н.С. Практикум по биометрии. Л.: изд-во ЛГУ, 1977. 152 с.

71. УлановаЕ.С., Сиротенко О.Д. Методы статистического анализа в агрометеорологии. Л.: Гидрометеоиздат, 1968. 197 с.

72. Уланова Е.С., Забелин В.Н. Методы корреляционного и регрессионного анализа в агрометеорологии Л.: Гидрометеоиздат, 1990. 207с.

73. Урбах В.Ю. Биометрические методы (Статистическая обработка опытных данных в биологии, сельском хозяйстве и медицине). М.: Наука, 1964. 415 с.

74. Филипченко Ю.А. Изменчивость и методы ее изучения. М.-Л.: Госиздат, 1929. 275 с.

75. Фишер Р. Статистические методы для исследователей. М.: Госстатиздат, 1958. 267 с.

76. Химич Г.З., Хлущевская О.А. Введение в биометрию. Учебное пособие для биологических специальностей вузов. Павлодар, 2009. 97 с.

77. Чугунова Н.И. Методика изучения возраста и роста рыб. М.: Советская наука, 1952. 115 с.

78. Чудновская Г.В. Математические методы в биологии: учебное пособие. Иркутск: Изд-во ИрГСХА, 2013. 112 с.

79. Шапиро А.П. Исследования по математической популяционной экологии. Владивосток: ДВНЦ АН СССР, 1983. 147 с.

80. Яковенко А. М. Биометрические методы анализа качественных и количественных признаков в зоотехнии: учебное пособие для студентов вузов, магистров, аспирантов. М.: СтГАУ, 2013. 91 с.

ПРИЛОЖЕНИЕ (ТАБЛИЦЫ)

Таблица 1 – Десятичные логарифмы целых чисел от 0 до 99

Десятки	Единицы									
	0	1	2	3	4	5	6	7	8	9
0	–	0	0,30103	0,47712	0,60206	0,69897	0,77815	0,84510	0,90309	0,95424
1	1	1,04139	1,07918	1,11394	1,14613	1,17609	1,20412	1,23045	1,25527	1,27875
2	1,30103	1,32222	1,34242	1,36173	1,38021	1,39794	1,41497	1,43136	1,44716	1,46240
3	1,47712	1,49136	1,50515	1,51851	1,53148	1,54407	1,55630	1,56820	1,57978	1,59106
4	1,60206	1,61278	1,62325	1,63347	1,64345	1,65321	1,66276	1,67210	1,68124	1,69020
5	1,69897	1,70757	1,71600	1,72428	1,73239	1,74036	1,74819	1,75587	1,76343	1,77085
6	1,77815	1,78533	1,79239	1,79934	1,80618	1,81291	1,81954	1,82607	1,83251	1,83885
7	1,84510	1,85126	1,85733	1,86332	1,86923	1,87506	1,88081	1,88649	1,89209	1,89763
8	1,90309	1,90849	1,91381	1,91908	1,92428	1,92942	1,93450	1,93952	1,94448	1,94939
9	1,95424	1,95904	1,96379	1,96848	1,97313	1,97772	1,98227	1,98677	1,99123	1,99564

Таблица 2 – Десятичные логарифмы
В столбцах единицы, в строках - десятки.

00 -∞	01 **0**	02 0,30103	03 0,47712	04 0,60206	05 0,69897	06 0,77815	07 0,8451	08 0,90309	09 0,95424
10 **1**	11 1,04139	12 1,07918	13 1,11394	14 1,14613	15 1,17609	16 1,20412	17 1,23045	18 1,25527	19 1,27875
20 1,30103	21 1,32222	22 1,34242	23 1,36173	24 1,38021	25 1,39794	26 1,41497	27 1,43136	28 1,44716	29 1,4624
30 1,47712	31 1,49136	32 1,50515	33 1,51851	34 1,53148	35 1,54407	36 1,5563	37 1,5682	38 1,57978	39 1,59106
40 1,60206	41 1,61278	42 1,62325	43 1,63347	44 1,64345	45 1,65321	46 1,66276	47 1,6721	48 1,68124	49 1,6902
50 1,69897	51 1,70757	52 1,716	53 1,72428	54 1,73239	55 1,74036	56 1,74819	57 1,75587	58 1,76343	59 1,77085
60 1,77815	61 1,78533	62 1,79239	63 1,79934	64 1,80618	65 1,81291	66 1,81954	67 1,82607	68 1,83251	69 1,83885
70 1,8451	71 1,85126	72 1,85733	73 1,86332	74 1,86923	75 1,87506	76 1,88081	77 1,88649	78 1,89209	79 1,89763
80 1,90309	81 1,90849	82 1,91381	83 1,91908	84 1,92428	85 1,92942	86 1,9345	87 1,93952	88 1,94448	89 1,94939
90	91	92	93	94	95	96	97	98	99

1,95424	1,95904	1,96379	1,96848	1,97313	1,97772	1,98227	1,98677	1,99123	1,99564
100 **2**	101 2,00432	102 2,0086	103 2,01284	104 2,01703	105 2,02119	106 2,02531	107 2,02938	108 2,03342	109 2,03743
110 2,04139	111 2,04532	112 2,04922	113 2,05308	114 2,0569	115 2,0607	116 2,06446	117 2,06819	118 2,07188	119 2,07555
120 2,07918	121 2,08279	122 2,08636	123 2,08991	124 2,09342	125 2,09691	126 2,10037	127 2,1038	128 2,10721	129 2,11059
130 2,11394	131 2,11727	132 2,12057	133 2,12385	134 2,1271	135 2,13033	136 2,13354	137 2,13672	138 2,13988	139 2,14301
140 2,14613	141 2,14922	142 2,15229	143 2,15534	144 2,15836	145 2,16137	146 2,16435	147 2,16732	148 2,17026	149 2,17319
150 2,17609	151 2,17898	152 2,18184	153 2,18469	154 2,18752	155 2,19033	156 2,19312	157 2,1959	158 2,19866	159 2,2014
160 2,20412	161 2,20683	162 2,20952	163 2,21219	164 2,21484	165 2,21748	166 2,22011	167 2,22272	168 2,22531	169 2,22789
170 2,23045	171 2,233	172 2,23553	173 2,23805	174 2,24055	175 2,24304	176 2,24551	177 2,24797	178 2,25042	179 2,25285
180 2,25527	181 2,25768	182 2,26007	183 2,26245	184 2,26482	185 2,26717	186 2,26951	187 2,27184	188 2,27416	189 2,27646
190 2,27875	191 2,28103	192 2,2833	193 2,28556	194 2,2878	195 2,29003	196 2,29226	197 2,29447	198 2,29667	199 2,29885
200 2,30103	201 2,3032	202 2,30535	203 2,3075	204 2,30963	205 2,31175	206 2,31387	207 2,31597	208 2,31806	209 2,32015
210 2,32222	211 2,32428	212 2,32634	213 2,32838	214 2,33041	215 2,33244	216 2,33445	217 2,33646	218 2,33846	219 2,34044
220 2,34242	221 2,34439	222 2,34635	223 2,3483	224 2,35025	225 2,35218	226 2,35411	227 2,35603	228 2,35793	229 2,35984
230 2,36173	231 2,36361	232 2,36549	233 2,36736	234 2,36922	235 2,37107	236 2,37291	237 2,37475	238 2,37658	239 2,3784
240 2,38021	241 2,38202	242 2,38382	243 2,38561	244 2,38739	245 2,38917	246 2,39094	247 2,3927	248 2,39445	249 2,3962
250 2,39794	251 2,39967	252 2,4014	253 2,40312	254 2,40483	255 2,40654	256 2,40824	257 2,40993	258 2,41162	259 2,4133
260 2,41497	261 2,41664	262 2,4183	263 2,41996	264 2,4216	265 2,42325	266 2,42488	267 2,42651	268 2,42813	269 2,42975
270 2,43136	271 2,43297	272 2,43457	273 2,43616	274 2,43775	275 2,43933	276 2,44091	277 2,44248	278 2,44404	279 2,4456
280 2,44716	281 2,44871	282 2,45025	283 2,45179	284 2,45332	285 2,45484	286 2,45637	287 2,45788	288 2,45939	289 2,4609

290 2,4624	291 2,46389	292 2,46538	293 2,46687	294 2,46835	295 2,46982	296 2,47129	297 2,47276	298 2,47422	299 2,47567
300 2,47712	301 2,47857	302 2,48001	303 2,48144	304 2,48287	305 2,4843	306 2,48572	307 2,48714	308 2,48855	309 2,48996
310 2,49136	311 2,49276	312 2,49415	313 2,49554	314 2,49693	315 2,49831	316 2,49969	317 2,50106	318 2,50243	319 2,50379
320 2,50515	321 2,50651	322 2,50786	323 2,5092	324 2,51055	325 2,51188	326 2,51322	327 2,51455	328 2,51587	329 2,5172
330 2,51851	331 2,51983	332 2,52114	333 2,52244	334 2,52375	335 2,52504	336 2,52634	337 2,52763	338 2,52892	339 2,5302
340 2,53148	341 2,53275	342 2,53403	343 2,53529	344 2,53656	345 2,53782	346 2,53908	347 2,54033	348 2,54158	349 2,54283
350 2,54407	351 2,54531	352 2,54654	353 2,54777	354 2,549	355 2,55023	356 2,55145	357 2,55267	358 2,55388	359 2,55509
360 2,5563	361 2,55751	362 2,55871	363 2,55991	364 2,5611	365 2,56229	366 2,56348	367 2,56467	368 2,56585	369 2,56703
370 2,5682	371 2,56937	372 2,57054	373 2,57171	374 2,57287	375 2,57403	376 2,57519	377 2,57634	378 2,57749	379 2,57864
380 2,57978	381 2,58092	382 2,58206	383 2,5832	384 2,58433	385 2,58546	386 2,58659	387 2,58771	388 2,58883	389 2,58995
390 2,59106	391 2,59218	392 2,59329	393 2,59439	394 2,5955	395 2,5966	396 2,5977	397 2,59879	398 2,59988	399 2,60097
400 2,60206	401 2,60314	402 2,60423	403 2,60531	404 2,60638	405 2,60746	406 2,60853	407 2,60959	408 2,61066	409 2,61172
410 2,61278	411 2,61384	412 2,6149	413 2,61595	414 2,617	415 2,61805	416 2,61909	417 2,62014	418 2,62118	419 2,62221
420 2,62325	421 2,62428	422 2,62531	423 2,62634	424 2,62737	425 2,62839	426 2,62941	427 2,63043	428 2,63144	429 2,63246
430 2,63347	431 2,63448	432 2,63548	433 2,63649	434 2,63749	435 2,63849	436 2,63949	437 2,64048	438 2,64147	439 2,64246
440 2,64345	441 2,64444	442 2,64542	443 2,6464	444 2,64738	445 2,64836	446 2,64933	447 2,65031	448 2,65128	449 2,65225
450 2,65321	451 2,65418	452 2,65514	453 2,6561	454 2,65706	455 2,65801	456 2,65896	457 2,65992	458 2,66087	459 2,66181
460 2,66276	461 2,6637	462 2,66464	463 2,66558	464 2,66652	465 2,66745	466 2,66839	467 2,66932	468 2,67025	469 2,67117
470 2,6721	471 2,67302	472 2,67394	473 2,67486	474 2,67578	475 2,67669	476 2,67761	477 2,67852	478 2,67943	479 2,68034
480 2,68124	481 2,68215	482 2,68305	483 2,68395	484 2,68485	485 2,68574	486 2,68664	487 2,68753	488 2,68842	489 2,68931

490	491	492	493	494	495	496	497	498	499
2,6902	2,69108	2,69197	2,69285	2,69373	2,69461	2,69548	2,69636	2,69723	2,6981
500	501	502	503	504	505	506	507	508	509
2,69897	2,69984	2,7007	2,70157	2,70243	2,70329	2,70415	2,70501	2,70586	2,70672
510	511	512	513	514	515	516	517	518	519
2,70757	2,70842	2,70927	2,71012	2,71096	2,71181	2,71265	2,71349	2,71433	2,71517
520	521	522	523	524	525	526	527	528	529
2,716	2,71684	2,71767	2,7185	2,71933	2,72016	2,72099	2,72181	2,72263	2,72346
530	531	532	533	534	535	536	537	538	539
2,72428	2,72509	2,72591	2,72673	2,72754	2,72835	2,72916	2,72997	2,73078	2,73159
540	541	542	543	544	545	546	547	548	549
2,73239	2,7332	2,734	2,7348	2,7356	2,7364	2,73719	2,73799	2,73878	2,73957
550	551	552	553	554	555	556	557	558	559
2,74036	2,74115	2,74194	2,74273	2,74351	2,74429	2,74507	2,74586	2,74663	2,74741
560	561	562	563	564	565	566	567	568	569
2,74819	2,74896	2,74974	2,75051	2,75128	2,75205	2,75282	2,75358	2,75435	2,75511
570	571	572	573	574	575	576	577	578	579
2,75587	2,75664	2,7574	2,75815	2,75891	2,75967	2,76042	2,76118	2,76193	2,76268
580	581	582	583	584	585	586	587	588	589
2,76343	2,76418	2,76492	2,76567	2,76641	2,76716	2,7679	2,76864	2,76938	2,77012
590	591	592	593	594	595	596	597	598	599
2,77085	2,77159	2,77232	2,77305	2,77379	2,77452	2,77525	2,77597	2,7767	2,77743
600	601	602	603	604	605	606	607	608	609
2,77815	2,77887	2,7796	2,78032	2,78104	2,78176	2,78247	2,78319	2,7839	2,78462
610	611	612	613	614	615	616	617	618	619
2,78533	2,78604	2,78675	2,78746	2,78817	2,78888	2,78958	2,79029	2,79099	2,79169
620	621	622	623	624	625	626	627	628	629
2,79239	2,79309	2,79379	2,79449	2,79518	2,79588	2,79657	2,79727	2,79796	2,79865
630	631	632	633	634	635	636	637	638	639
2,79934	2,80003	2,80072	2,8014	2,80209	2,80277	2,80346	2,80414	2,80482	2,8055
640	641	642	643	644	645	646	647	648	649
2,80618	2,80686	2,80754	2,80821	2,80889	2,80956	2,81023	2,8109	2,81158	2,81224
650	651	652	653	654	655	656	657	658	659
2,81291	2,81358	2,81425	2,81491	2,81558	2,81624	2,8169	2,81757	2,81823	2,81889
660	661	662	663	664	665	666	667	668	669
2,81954	2,8202	2,82086	2,82151	2,82217	2,82282	2,82347	2,82413	2,82478	2,82543
670	671	672	673	674	675	676	677	678	679
2,82607	2,82672	2,82737	2,82802	2,82866	2,8293	2,82995	2,83059	2,83123	2,83187
680	681	682	683	684	685	686	687	688	689
2,83251	2,83315	2,83378	2,83442	2,83506	2,83569	2,83632	2,83696	2,83759	2,83822

690 2,83885	691 2,83948	692 2,84011	693 2,84073	694 2,84136	695 2,84198	696 2,84261	697 2,84323	698 2,84386	699 2,84448
700 2,8451	701 2,84572	702 2,84634	703 2,84696	704 2,84757	705 2,84819	706 2,8488	707 2,84942	708 2,85003	709 2,85065
710 2,85126	711 2,85187	712 2,85248	713 2,85309	714 2,8537	715 2,85431	716 2,85491	717 2,85552	718 2,85612	719 2,85673
720 2,85733	721 2,85794	722 2,85854	723 2,85914	724 2,85974	725 2,86034	726 2,86094	727 2,86153	728 2,86213	729 2,86273
730 2,86332	731 2,86392	732 2,86451	733 2,8651	734 2,8657	735 2,86629	736 2,86688	737 2,86747	738 2,86806	739 2,86864
740 2,86923	741 2,86982	742 2,8704	743 2,87099	744 2,87157	745 2,87216	746 2,87274	747 2,87332	748 2,8739	749 2,87448
750 2,87506	751 2,87564	752 2,87622	753 2,87679	754 2,87737	755 2,87795	756 2,87852	757 2,8791	758 2,87967	759 2,88024
760 2,88081	761 2,88138	762 2,88195	763 2,88252	764 2,88309	765 2,88366	766 2,88423	767 2,8848	768 2,88536	769 2,88593
770 2,88649	771 2,88705	772 2,88762	773 2,88818	774 2,88874	775 2,8893	776 2,88986	777 2,89042	778 2,89098	779 2,89154
780 2,89209	781 2,89265	782 2,89321	783 2,89376	784 2,89432	785 2,89487	786 2,89542	787 2,89597	788 2,89653	789 2,89708
790 2,89763	791 2,89818	792 2,89873	793 2,89927	794 2,89982	795 2,90037	796 2,90091	797 2,90146	798 2,902	799 2,90255
800 2,90309	801 2,90363	802 2,90417	803 2,90472	804 2,90526	805 2,9058	806 2,90634	807 2,90687	808 2,90741	809 2,90795
810 2,90849	811 2,90902	812 2,90956	813 2,91009	814 2,91062	815 2,91116	816 2,91169	817 2,91222	818 2,91275	819 2,91328
820 2,91381	821 2,91434	822 2,91487	823 2,9154	824 2,91593	825 2,91645	826 2,91698	827 2,91751	828 2,91803	829 2,91855
830 2,91908	831 2,9196	832 2,92012	833 2,92065	834 2,92117	835 2,92169	836 2,92221	837 2,92273	838 2,92324	839 2,92376
840 2,92428	841 2,9248	842 2,92531	843 2,92583	844 2,92634	845 2,92686	846 2,92737	847 2,92788	848 2,9284	849 2,92891
850 2,92942	851 2,92993	852 2,93044	853 2,93095	854 2,93146	855 2,93197	856 2,93247	857 2,93298	858 2,93349	859 2,93399
860 2,9345	861 2,935	862 2,93551	863 2,93601	864 2,93651	865 2,93702	866 2,93752	867 2,93802	868 2,93852	869 2,93902
870 2,93952	871 2,94002	872 2,94052	873 2,94101	874 2,94151	875 2,94201	876 2,9425	877 2,943	878 2,94349	879 2,94399
880 2,94448	881 2,94498	882 2,94547	883 2,94596	884 2,94645	885 2,94694	886 2,94743	887 2,94792	888 2,94841	889 2,9489

890 2,94939	891 2,94988	892 2,95036	893 2,95085	894 2,95134	895 2,95182	896 2,95231	897 2,95279	898 2,95328	899 2,95376
900 2,95424	901 2,95472	902 2,95521	903 2,95569	904 2,95617	905 2,95665	906 2,95713	907 2,95761	908 2,95809	909 2,95856
910 2,95904	911 2,95952	912 2,95999	913 2,96047	914 2,96095	915 2,96142	916 2,9619	917 2,96237	918 2,96284	919 2,96332
920 2,96379	921 2,96426	922 2,96473	923 2,9652	924 2,96567	925 2,96614	926 2,96661	927 2,96708	928 2,96755	929 2,96802
930 2,96848	931 2,96895	932 2,96942	933 2,96988	934 2,97035	935 2,97081	936 2,97128	937 2,97174	938 2,9722	939 2,97267
940 2,97313	941 2,97359	942 2,97405	943 2,97451	944 2,97497	945 2,97543	946 2,97589	947 2,97635	948 2,97681	949 2,97727
950 2,97772	951 2,97818	952 2,97864	953 2,97909	954 2,97955	955 2,98	956 2,98046	957 2,98091	958 2,98137	959 2,98182
960 2,98227	961 2,98272	962 2,98318	963 2,98363	964 2,98408	965 2,98453	966 2,98498	967 2,98543	968 2,98588	969 2,98632
970 2,98677	971 2,98722	972 2,98767	973 2,98811	974 2,98856	975 2,989	976 2,98945	977 2,98989	978 2,99034	979 2,99078
980 2,99123	981 2,99167	982 2,99211	983 2,99255	984 2,993	985 2,99344	986 2,99388	987 2,99432	988 2,99476	989 2,9952
990 2,99564	991 2,99607	992 2,99651	993 2,99695	994 2,99739	995 2,99782	996 2,99826	997 2,9987	998 2,99913	999 2,99957
1000 **3**	1001 3,00043	1002 3,00087	1003 3,0013	1004 3,00173	1005 3,00217	1006 3,0026	1007 3,00303	1008 3,00346	1009 3,00389

Таблица 3 – Мантиссы десятичных логарифмов

N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
10	0	43									4	9	13	17	22	26	30	35	39
			86	128	170						4	9	13	17	21	25	30	34	38
						212	253				4	8	12	16	21	25	29	33	37
								294	334	374	4	8	12	16	20	24	28	32	36
11	414	453	492								4	8	12	16	20	24	27	31	35
				531	569	607					4	8	11	15	19	23	27	30	34
							645	682	719	755	4	7	11	15	18	22	26	29	33
12	792	898	864	899	934						3	7	11	14	18	21	25	28	32
						969					4	7	11	14	17	21	24	28	31

							1004	1038	1072	1106	3	7	10	14	17	20	24	27	30
13	1139	1173									3	7	10	13	17	20	23	27	30
			1206	1239	1271	1303	1335				3	6	10	13	16	19	23	26	29
								1367	1399	1430	3	6	9	13	16	19	22	25	28
14	1461	1492									3	6	9	13	16	19	22	25	28
			1523	1553	1584	1614	1644	1673			3	6	9	12	15	18	21	24	27
									1703	1732	3	6	9	11	14	17	20	23	26
15	1761	1790	1818	1847	1875	1903	1931				3	6	9	11	14	17	20	23	26
								1959	1987	2014	3	5	8	11	14	16	19	22	25
16	2041	2068	2095	2122	2148	2175	2201	2227			3	5	8	11	13	16	19	21	24
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
									2253	2279	3	5	8	10	13	15	18	20	23
17	2304	2330	2355	2380	2405	2430					3	5	8	10	13	15	18	20	23
							2455	2480	2504	2529	2	5	7	10	12	15	17	19	22
18	2553	2577	2601	2625	2648	2672	2695	2718			2	5	7	9	12	14	16	19	21
									2742	2765	2	5	7	9	11	13	16	18	20
19	2788	2810	2833	2856	2878	2900					2	4	7	9	11	13	16	18	20
							2923	2945	2967	2989	2	4	6	8	11	13	15	17	19
20	3010	3032	3054	3075	3096						2	4	6	8	11	13	15	17	19
						3118	3139	3160	3181	3201	2	4	6	8	10	12	14	17	19
21	3222	3243	3263	3284	3304	3324	3345	3365	3385	3404	2	4	6	8	10	12	14	16	18
22	3424	3444	3464	3483	3502	3522	3541	3560	3579	3598	2	4	6	8	10	12	14	15	17
23	3617	3636	3655	3674	3692	3711	3729	3747	3766	3784	2	4	6	7	9	11	13	15	17
24	3802	3820	3838	3856	3874	3892	3909	3927	3945	3962	2	4	5	7	9	11	12	14	16
25	3979	3997	4014	4031	4048	4065	4082	4099	4116	4133	2	3	5	7	9	10	12	14	15
26	4150	4166	4183	4200	4216	4232	4249	4265	4281	4298	2	3	5	7	8	10	11	13	15
27	4314	4330	4346	4362	4378	4393	4409	4425	4440	4456	2	3	5	6	8	9	11	13	14
28	4472	4487	4502	4518	4533	4548	4564	4579	4594	4609	2	3	5	6	8	9	11	12	14
29	4624	4639	4654	4669	4683	4698	4713	4728	4742	4757	1	3	4	6	7	9	10	12	13
30	4771	4786	4800	4814	4829	4843	4857	4871	4886	4900	1	3	4	6	7	9	10	11	13
31	4914	4928	4942	4955	4969	4983	4997	5011	5024	5038	1	3	4	6	7	8	10	11	13
32	5051	5065	5079	5092	5105	5119	5132	5145	5159	5172	1	3	4	5	7	8	9	11	12
33	5185	5198	5211	5224	5237	5250	5263	5276	5289	5302	1	3	4	5	6	8	9	10	12
34	5315	5328	5340	5353	5366	5378	5391	5403	5416	5428	1	3	4	5	6	8	9	10	11

35	5441	5453	5465	5478	5490	5502	5514	5527	5539	5551	1	2	4	5	6	7	9	10	11
36	563	5575	5587	5599	5611	5623	5635	5647	5658	5670	1	2	4	5	6	7	8	10	11
37	5682	5694	5705	5717	5729	5740	5752	5763	5775	5786	1	2	3	5	6	7	8	9	10
38	5798	5809	5821	5832	5843	5855	5866	5877	5888	5899	1	2	3	5	6	7	8	9	10
39	5911	5922	5933	5944	5955	5966	5977	5988	5999	6010	1	2	3	4	5	7	8	9	10
40	6021	6031	6042	6053	6064	6075	6085	6096	6107	6117	1	2	3	4	5	6	8	9	10
41	6128	6138	6149	6160	6170	6180	6191	6201	6212	6222	1	2	3	4	5	6	7	8	9
42	6232	6243	6253	6263	6274	6284	6294	6304	6314	6325	1	2	3	4	5	6	7	8	9
43	6335	6345	6355	6365	6375	6385	6395	6405	6415	6425	1	2	3	4	5	6	7	8	9
44	6435	6444	6454	6464	6474	6484	6493	6503	6513	6522	1	2	3	4	5	6	7	8	9
45	6532	6542	6551	6561	6571	6580	6590	6599	6609	6618	1	2	3	4	5	6	7	8	9
46	6628	6637	6646	6656	6665	6675	6684	6693	6702	6712	1	2	3	4	5	6	7	7	8
47	6721	6730	6739	6749	6758	6767	6776	6785	6794	6803	1	2	3	4	5	5	6	7	8
48	6812	6821	6830	6839	6848	6857	6866	6875	6884	6893	1	2	3	4	4	5	6	7	8
49	6902	6911	6920	6928	6937	6946	6955	6964	6972	6981	1	2	3	4	4	5	6	7	8
50	6990	6998	7007	7016	7024	7033	7042	7050	7059	7067	1	2	3	3	4	5	6	7	8
51	7076	7084	7093	7101	7110	7118	7126	7135	7143	7152	1	2	3	3	4	5	6	7	8
52	7160	7168	7177	7185	7193	7202	7210	7218	7226	7235	1	2	2	3	4	5	6	7	7
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
53	7243	7251	7259	7267	7275	7284	7292	7300	7308	7316	1	2	2	3	4	4	6	6	7
54	7324	7332	7340	7348	7356	7364	7372	7380	7388	7396	1	2	2	3	4	4	6	6	7
55	7404	7412	7419	7427	7435	7443	7451	7459	7466	7474	1	2	2	3	4	5	5	6	7
56	7482	7490	7497	7505	7513	7520	7528	7536	7543	7551	1	2	2	3	4	5	5	6	7
57	7559	7566	7574	7582	7589	7597	7604	7612	7619	7627	1	2	2	3	4	5	5	6	7
58	7634	7642	7649	7657	7664	7672	7679	7686	7694	7701	1	1	2	3	4	4	5	6	7
59	7709	7716	7723	7731	7738	7745	7752	7760	7767	7774	1	1	2	3	4	4	5	6	7
60	7782	7789	7796	7803	7810	7818	7825	7832	7839	7846	1	1	2	3	4	4	5	6	6
61	7853	7860	7868	7875	7882	7889	7896	7903	7910	7917	1	1	2	3	4	4	5	6	6
62	7924	7931	7938	7945	7952	7959	7966	7973	7980	7987	1	1	2	3	4	4	5	6	6
63	7993	8000	8007	8014	8021	8028	8035	8041	8048	8055	1	1	2	3	3	4	5	5	6
64	8062	8069	8075	8082	8089	8096	8102	8109	8116	8122	1	1	2	3	3	4	5	5	6
65	8129	8136	8142	8149	8156	8162	8169	8176	8182	8189	1	1	2	3	3	4	5	5	6
66	8195	8202	8209	8215	8222	8228	8235	8241	8248	8254	1	1	2	3	3	4	5	5	6
67	8261	8267	8274	8280	8287	8293	8299	8306	8312	8319	1	1	2	3	3	4	5	5	6

68	8325	8331	8338	8344	8351	8357	8363	8370	8376	8382	1	1	2	3	3	4	4	5	6
69	8388	8395	8401	8407	8414	8420	8426	8432	8439	8445	1	1	2	2	3	4	4	5	6
70	8451	8457	8463	8470	8476	8482	8488	8494	8500	8506	1	1	2	2	3	4	4	5	6
71	8513	8519	8525	8531	8537	8543	8549	8555	8561	8567	1	1	2	2	3	4	4	5	5
72	8573	8579	8585	8591	8597	8603	8609	8615	8621	8627	1	1	2	2	3	4	4	5	5
73	8633	8639	8645	8651	8657	8663	8669	8675	8681	8686	1	1	2	2	3	4	4	5	5
74	8692	8698	8704	8710	8716	8722	8727	8733	8739	8745	1	1	2	2	3	4	4	5	5
75	8751	8756	8762	8768	8774	8779	8785	8791	8797	8802	1	1	2	2	3	3	4	5	5
76	8808	8814	8820	8825	8831	8837	8843	8848	8854	8859	1	1	2	2	3	3	4	5	5
77	8865	8871	8876	8882	8887	8893	8899	8904	8910	8915	1	1	2	2	3	3	4	4	5
78	8921	8927	8932	8938	8943	8949	8954	8960	8965	8971	1	1	2	2	3	3	4	4	5
79	8976	8982	8987	8993	8998	9004	9009	9015	9020	9025	1	1	2	2	3	3	4	4	5
80	9031	9036	9042	9047	9053	9058	9063	9069	9074	9079	1	1	2	2	3	3	4	4	5
81	9085	9090	9096	9101	9106	9112	9117	9122	9128	9133	1	1	2	2	3	3	4	4	5
82	9138	9143	9149	9154	9159	9165	9170	9175	9180	9186	1	1	2	2	3	3	4	4	5
83	9191	9196	9201	9206	9212	9217	9222	9227	9232	9238	1	1	2	2	3	3	4	4	5
84	9243	9248	9253	9258	9263	9269	9274	9279	9284	9289	1	1	2	2	3	3	4	4	5
85	9294	9299	9304	9309	9315	9320	9325	9330	9335	9340	1	1	2	2	3	3	4	4	5
86	9345	9350	9355	9360	9365	9370	9375	9380	9385	9390	1	1	2	2	3	3	4	4	5
87	9395	9400	9405	9410	9415	9420	9425	9430	9435	9440	0	1	1	2	2	3	3	4	4
88	9445	9450	9455	9460	9465	9469	9474	9479	9484	9489	0	1	1	2	2	3	3	4	4
89	9494	9499	9504	9509	9513	9518	9523	9528	9533	9538	0	1	1	2	2	3	3	4	4
90	9542	9547	9552	9557	9562	9566	9571	9576	9581	9586	0	1	1	2	2	3	3	4	4
91	9590	9595	9600	9605	9609	9614	9619	9624	9628	9633	0	1	1	2	2	3	3	4	4
92	9638	9643	9647	9652	9657	9661	9666	9671	9675	9680	0	1	1	2	2	3	3	4	4
93	9685	9689	9694	9699	9703	9708	9713	9717	9722	9727	0	1	1	2	2	3	3	4	4
N	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
94	9731	9736	9741	9745	9750	9754	9759	9763	9768	9773	0	1	1	2	2	3	3	4	4
95	9777	9782	9786	9791	9795	9800	9805	9809	9814	9818	0	1	1	2	2	3	3	4	4
96	9823	9827	9832	9836	9841	9845	9850	9854	9859	9863	0	1	1	2	2	3	3	4	4
97	9868	9872	9877	9881	9886	9890	9894	9899	9903	9908	0	1	1	2	2	3	3	4	4
98	9912	9917	9921	9926	9930	9934	9939	9943	9948	9952	0	1	1	2	2	3	3	4	4
99	9956	9961	9965	9969	9974	9978	9983	9987	9991	9996	0	1	1	2	2	3	3	3	4

Таблица 4 – Десятичные антилогарифмы

m	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
.00	1000	1002	1005	1007	1009	1012	1014	1016	1019	1021	0	0	1	1	1	1	2	2	2
.01	1023	1026	1028	1030	1033	1035	1038	1040	1042	1045	0	0	1	1	1	1	2	2	2
.02	1047	1050	1052	1054	1057	1059	1062	1064	1067	1069	0	0	1	1	1	1	2	2	2
.03	1072	1074	1076	1079	1081	1084	1086	1089	1091	1094	0	0	1	1	1	1	2	2	2
.04	1096	1099	1102	1104	1107	1109	1112	1114	1117	1119	0	1	1	1	1	2	2	2	2
.05	1122	1125	1127	1130	1132	1135	1138	1140	1143	1146	0	1	1	1	1	2	2	2	2
.06	1148	1151	1153	1156	1159	1161	1164	1167	1169	1172	0	1	1	1	1	2	2	2	2
.07	1175	1178	1180	1183	1186	1189	1191	1194	1197	1199	0	1	1	1	1	2	2	2	2
.08	1202	1205	1208	1211	1213	1216	1219	1222	1225	1227	0	1	1	1	1	2	2	2	3
.09	1230	1233	1236	1239	1242	1245	1247	1250	1253	1256	0	1	1	1	1	2	2	2	3
.10	1259	1262	1265	1268	1271	1274	1276	1279	1282	1285	0	1	1	1	1	2	2	2	3
.11	1288	1291	1294	1297	1300	1303	1306	1309	1312	1315	0	1	1	1	2	2	2	2	3
.12	1318	1321	1324	1327	1330	1334	1337	1340	1343	1346	0	1	1	1	2	2	2	2	3
.13	1349	1352	1355	1358	1361	1365	1368	1371	1374	1377	0	1	1	1	2	2	2	3	3
.14	1380	1384	1387	1390	1393	1396	1400	1403	1406	1409	0	1	1	1	2	2	2	3	3
.15	1413	1416	1419	1422	1426	1429	1432	1435	1439	1442	0	1	1	1	2	2	2	3	3
.16	1445	1449	1452	1455	1459	1462	1466	1469	1472	1476	0	1	1	1	2	2	2	3	3
.17	1479	1483	1486	1489	1493	1496	1500	1503	1507	1510	0	1	1	1	2	2	2	3	3
.18	1514	1517	1521	1524	1528	1531	1535	1538	1542	1545	0	1	1	1	2	2	2	3	3
.19	1549	1552	1556	1560	1563	1567	1570	1574	1578	1581	0	1	1	1	2	2	3	3	3
.20	1585	1589	1592	1596	1600	1603	1607	1611	1614	1618	0	1	1	1	2	2	3	3	3
.21	1622	1626	1629	1633	1637	1641	1644	1648	1652	1656	0	1	1	2	2	2	3	3	3
.22	1660	1663	1667	1671	1675	1679	1683	1687	1690	1694	0	1	1	2	2	2	3	3	3
.23	1698	1702	1706	1710	1714	1718	1722	1726	1730	1734	0	1	1	2	2	2	3	3	4
.24	1738	1742	1746	1750	1754	1758	1762	1766	1770	1774	0	1	1	2	2	2	3	3	4
.25	1778	1782	1786	1791	1795	1799	1803	1807	1811	1816	0	1	1	2	2	2	3	3	4
.26	1820	1824	1828	1832	1837	1841	1845	1849	1854	1858	0	1	1	2	2	3	3	3	4
.27	1862	1866	1871	1875	1879	1884	1888	1892	1897	1901	0	1	1	2	2	3	3	3	4
.28	1905	1910	1914	1919	1923	1928	1932	1936	1941	1945	0	1	1	2	2	3	3	4	4
m	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
.29	1950	1954	1959	1963	1968	1972	1977	1982	1986	1991	0	1	1	2	2	3	3	4	4
.30	1995	2000	2004	2009	2014	2018	2023	2028	2032	2037	0	1	1	2	2	3	3	4	4
.31	2042	2046	2051	2056	2061	2065	2070	2075	2080	2084	0	1	1	2	2	3	3	4	4

.32	2089	2094	2099	2104	2109	2113	2118	2123	2128	2133	0	1	1	2	2	3	3	4	4
.33	2138	2143	2148	2153	2158	2163	2168	2173	2178	2183	0	1	1	2	2	3	3	4	4
.34	2188	2193	2198	2203	2208	2213	2218	2223	2228	2234	1	1	2	2	3	3	4	4	5
.35	2239	2244	2249	2254	2259	2265	2270	2275	2280	2286	1	1	2	2	3	3	4	4	5
.36	2291	2296	2301	2307	2312	2317	2323	2328	2333	2339	1	1	2	2	3	3	4	4	5
.37	2344	2350	2355	2360	2366	2371	2377	2382	2388	2393	1	1	2	2	3	3	4	4	5
.38	2399	2404	2410	2415	2421	2427	2432	2438	2443	2449	1	1	2	2	3	3	4	4	5
.39	2455	2460	2466	2472	2477	2483	2489	2495	2500	2506	1	1	2	2	3	3	4	5	5
.40	2512	2518	2523	2529	2535	2541	2547	2553	2559	2564	1	1	2	2	3	4	4	5	5
.41	2570	2576	2582	2588	2594	2600	2606	2612	2618	2624	1	1	2	2	3	4	4	5	5
.42	2630	2636	2642	2649	2655	2661	2667	2673	2679	2685	1	1	2	2	3	4	4	5	6
.43	2692	2698	2704	2710	2716	2723	2729	2735	2742	2748	1	1	2	3	3	4	4	5	6
.44	2754	2761	2767	2773	2780	2786	2793	2799	2805	2812	1	1	2	3	3	4	4	5	6
.45	2818	2825	2831	2838	2844	2851	2858	2864	2871	2877	1	1	2	3	3	4	5	5	6
.46	2884	2891	2897	2904	2911	2917	2924	2931	2938	2944	1	1	2	3	3	4	5	5	6
.47	2951	2958	2965	2972	2979	2985	2992	2999	3006	3013	1	1	2	3	3	4	5	5	6
.48	3020	3027	3034	3041	3048	3055	3062	3069	3076	3083	1	1	2	3	4	4	5	6	6
.49	3090	3097	3105	3112	3119	3126	3133	3141	3148	3155	1	1	2	3	4	4	5	6	6
.50	3162	3170	3177	3184	3192	3199	3206	3214	3221	3228	1	1	2	3	4	4	5	6	7
.51	3236	3243	3251	3258	3266	3273	3281	3289	3296	3304	1	2	2	3	4	5	5	6	7
.52	3311	3319	3327	3334	3342	3350	3357	3365	3373	3381	1	2	2	3	4	5	5	6	7
.53	3388	3396	3404	3412	3420	3428	3436	3443	3451	3459	1	2	2	3	4	5	6	6	7
.54	3467	3475	3483	3491	3499	3508	3516	3524	3532	3540	1	2	2	3	4	5	6	6	7
.55	3548	3556	3565	3573	3581	3589	3597	3606	3614	3622	1	2	2	3	4	5	6	7	7
.56	3631	3639	3648	3656	3664	3673	3681	3690	3698	3707	1	2	3	3	4	5	6	7	8
.57	3715	3724	3733	3741	3750	3758	3767	3776	3784	3793	1	2	3	3	4	5	6	7	8
.58	3802	3811	3819	3828	3837	3846	3855	3864	3873	3882	1	2	3	4	4	5	6	7	8
.59	3890	3899	3908	3917	3926	3936	3945	3954	3963	3972	1	2	3	4	5	5	6	7	8
.60	3981	3990	3999	4009	4018	4027	4036	4046	4055	4064	1	2	3	4	5	6	6	7	8
.61	4074	4083	4093	4102	4111	4121	4130	4140	4150	4159	1	2	3	4	5	6	7	8	9
.62	4169	4178	4188	4198	4207	4217	4227	4236	4246	4256	1	2	3	4	5	6	7	8	9
.63	4266	4276	4285	4295	4305	4315	4325	4335	4345	4355	1	2	3	4	5	6	7	8	9
.64	4365	4375	4385	4395	4406	4416	4426	4436	4446	4457	1	2	3	4	5	6	7	8	9
.65	4467	4477	4487	4498	4508	4519	4529	4539	4550	4560	1	2	3	4	5	6	7	8	9
.66	4571	4581	4592	4603	4613	4624	4634	4645	4656	4667	1	2	3	4	5	6	7	9	10
m	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

.67	4677	4688	4699	4710	4721	4732	4742	4753	4764	4775	1	2	3	4	5	7	8	9	10
.68	4786	4797	4808	4819	4831	4842	4853	4864	4875	4887	1	2	3	4	6	7	8	9	10
.69	4898	4909	4920	4932	4943	4955	4966	4977	4989	5000	1	2	3	5	6	7	8	9	10
.70	5012	5023	5035	5047	5058	5070	5082	5093	5105	5117	1	2	4	5	6	7	8	9	11
.71	5129	5140	5152	5164	5176	5188	5200	5212	5224	5236	1	2	4	5	6	7	8	10	11
.72	5248	5260	5272	5284	5297	5309	5321	5333	5346	5358	1	2	4	5	6	7	9	10	11
.73	5370	5383	5395	5408	5420	5433	5445	5458	5470	5483	1	3	4	5	6	8	9	10	11
.74	5495	5508	5521	5534	5546	5559	5572	5585	5598	5610	1	3	4	5	6	8	9	10	12
.75	5623	5636	5649	5662	5675	5689	5702	5715	5728	5741	1	3	4	5	7	8	9	10	12
.76	5754	5768	5781	5794	5808	5821	5834	5848	5861	5875	1	3	4	5	7	8	9	11	12
.77	5888	5902	5916	5929	5943	5957	5970	5984	5998	6012	1	3	4	5	7	8	10	11	12
.78	6026	6039	6053	6067	6081	6095	6109	6124	6138	6152	1	3	4	6	7	8	10	11	13
.79	6166	6180	6194	6209	6223	6237	6252	6266	6281	6295	1	3	4	6	7	9	10	11	13
.80	6310	6324	6339	6353	6368	6383	6397	6412	6427	6442	1	3	4	6	7	9	10	12	13
.81	6457	6471	6486	6501	6516	6531	6546	6561	6577	6592	2	3	5	6	8	9	11	12	14
.82	6607	6622	6637	6653	6668	6683	6699	6714	6730	6745	2	3	5	6	8	9	11	12	14
.83	6761	6776	6792	6808	6823	6839	6855	6871	6887	6902	2	3	5	6	8	9	11	13	14
.84	6918	6934	6950	6966	6982	6998	7015	7031	7047	7063	2	3	5	6	8	10	11	13	15
.85	7079	7096	7112	7129	7145	7161	7178	7194	7211	7228	2	3	5	7	8	10	12	13	15
.86	7244	7261	7278	7295	7311	7328	7345	7362	7379	7396	2	3	5	7	8	10	12	13	15
.87	7413	7430	7447	7464	7482	7499	7516	7534	7551	7568	2	3	5	7	9	10	12	14	16
.88	7586	7603	7621	7638	7656	7674	7691	7709	7727	7745	2	4	5	7	9	11	12	14	16
.89	7762	7780	7798	7816	7834	7852	7870	7889	7907	7925	2	4	5	7	9	11	13	14	16
.90	7943	7962	7980	7998	8017	8035	8054	8072	8091	8110	2	4	6	7	9	11	13	15	17
.91	8128	8147	8166	8185	8204	8222	8241	8260	8279	8299	2	4	6	8	9	11	13	15	17
.92	8318	8337	8356	8375	8395	8414	8433	8453	8472	8492	2	4	6	8	10	12	14	15	17
.93	8511	8531	8551	8570	8590	8610	8630	8650	8670	8690	2	4	6	8	10	12	14	16	18
.94	8710	8730	8750	8770	8790	8810	8831	8851	8872	8892	2	4	6	8	10	12	14	16	18
.95	8913	8933	8954	8974	8995	9016	9036	9057	9078	9099	2	4	6	8	10	12	15	17	19
.96	9120	9141	9162	9183	9204	9226	9247	9268	9290	9311	2	4	6	8	11	13	15	17	19
.97	9333	9354	9376	9397	9419	9441	9462	9484	9506	9528	2	4	7	9	11	13	15	17	20
.98	9550	9572	9594	9616	9638	9661	9683	9705	9727	9750	2	4	7	9	11	13	16	18	20
.99	9772	9795	9817	9840	9863	9886	9908	9931	9954	9977	2	5	7	9	11	14	16	18	20
m	0	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9

Таблица 5 – Площади и ординаты нормальной кривой распределения (из Миллса)

Нормированное отклонение $x = \frac{v - M}{\delta}$	Вторая функция нормированного оклонения $\varphi(x)$		Первая функция нормированного отклонения f (x) ордината y при значениях $x=\frac{v-M}{\delta}$ То есть вероятность y_x при отклонениях v от M на x
	Площадь между ординатами y_0 и y $\frac{v - M}{\delta}$	% числа наблюдений, заключенных между ординатами y_0 и y	
0,0	0,00000	0	0,39894
0,1	0,03983	3,983	0,39695
0,2	0,07926	7,926	0,39104
0,3	0,11791	11,791	0,38139
0,4	0,15542	15,542	0,36827
0,5	0,19146	19,146	0,35207
0,6	0,22575	22,575	0,33322
0,7	0,25804	25,804	0,31225
0,8	0,28814	28,814	0,28969
0,9	0,31594	31,594	0,26609
1,0	0,34134	34,134	0,24197
1,1	0,36433	36,433	0,21785
1,2	0,38493	38,493	0,19419
1,3	0,40320	40,320	0,17137
1,4	0,41924	41,924	0,14973
1,5	0,43319	43,319	0,12952
1,6	0,44520	44,520	0,11092
1,7	0,45543	45,543	0,09405
1,8	0,46407	46,407	0,07895
1,9	0,47128	47,128	0,06562
2,0	0,47725	47,725	0,05399
2,1	0,48214	48,214	0,43398
2,2	0,48610	48,610	0,03547
2,3	0,48928	48,928	0,02833
2,4	0,49180	49,180	0,02239
2,5	0,49379	49,379	0,01753
2,6	0,49534	49,534	0,01358
2,7	0,49653	49,653	0,01042
2,8	0,49744	49,744	0,00792
2,9	0,49813	49,813	0,00595
3,0	0,49865	49,865	0,00443
3,5	0,49977	49,977	0,00087
3,99	0,49997	49,997	0,00014

Таблица 6 – Значение критериев достоверности t при различных уровнях вероятности P и числа степеней свободы v, дающие достоверную величину средней арифметической и достоверность разности (M_1– M_2) при малом и большом числе наблюдений n (фрагмент)

Число степеней свободы v	Уровень вероятности P		
	0,95	0,99	0,999
	Значение t		
1	12,71	63,7	637,0
2	4,3	9,9	31,6
3	3,2	5,8	12,9
4	2,88	4,6	8,6
5	2,6	4,0	6,9
6	2,4	3,7	6,0
7	2,4	3,5	5,3
8	2,3	3,4	5,0
9	2,3	3,3	4,8
10	2,2	3,2	4,6
11	2,2	3,1	4,4
12	2,2	3,1	4,3
13	2,2	3,0	4,1
14	2,15	3,0	4,1
15	2,1	3,0	4,1
16	2,1	2,9	4,0
17	2,1	2,9	4,0
18	2,1	2,9	3,9
19-20	2,1	2,9	3,9
21-24	2,1	2,8	3,8
25-28	2,1	2,8	3,7
29-31	2,0	2,8	3,7
32-34	2,0	2,7	3,7
35-42	2,0	2,7	3,6
43-62	2,0	2,7	3,5
63-175	2,0	2,6	3,4
176 и больше	2,0	2,6	3,3

Таблица 7 – Критические значения критерия соответствия χ^2(хи-квадрат)

Число степеней свободы	Уровни значимости			Число степеней свободы	Уровни значимости		
	p=0,05	p=0,01	p=0,001		p=0,05	p=0,01	*p*
1	3,84	6,63	10.83	21	32,67	38,93	46,80
2	5,99	9,21	13,82	22	33,92	40,29	48,27

3	7,81	11,07	16,27	23	35,17	41,64	49,73
4	9,49	13,28	18,47	24	36,42	42,98	51,18
5	11,07	15,09	20,51	25	37,65	44,31	52.62
6	12,59	16,81	22.46	26	38,89	45,64	54,05
7	14,07	18,48	24,32	27	40,11	46,96	55,48
8	15,51	20,09	26,12	28	41,34	48,28	56,89
9	16,92	21,67	27,88	29	42,56	49,59	58,30
10	18,31	33,21	29,59	30	43,77	50,89	59,70
11	19,68	34,73	31,26	31	44,99	52.19	61,10
12	21,03	26,22	32,91	32	46,19	53,49	62.49
13	22,36	27,69	34,53	33	47,40	54,78	63,87
14	23,68	29,14	36,12	34	48,60	56,06	65,25
15	25,00	30,58	37,70	35	49,80	57,34	66,62
16	26,30	32,00	39,25	36	51,00	58,62	67,98
17	27,59	33,41	40,79	37	52,19	59,89	69,35
18	28,87	34,81	42,31	38	53,38	61,16	70,70
19	30,14	36,19	43,82	39	54,57	62,43	72,06
20	31,41	37,57	45,31	40	55,76	63,69	73,40

Таблица 8 – Краткая таблица критических значений критерия *F* Фишера
(P_t = 0,95)

K_1 / K_2	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**14**	**16**	**20**	**24**	**30**
2	19,0	19,2	19,3	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5
3	9,6	9,3	9,1	9,0	8,9	8,9	8,8	8,8	8,8	8,7	8,7	8,7	8,7	8,6	8,6
4	6,9	6,6	6,4	6,3	6,2	6,1	6,0	6,0	6,0	5,9	5,9	5,8	5,8	5,8	5,8
5	5,8	5,4	5,2	5,1	5,0	4,9	4,8	4,8	4,7	4,7	4,6	4,6	4,6	4,5	4,5
6	5,1	4,8	4,5	4,4	4,3	4,2	4,2	4,1	4,1	4,0	4,0	3,9	3,9	3,8	3,8
7	4,7	4,4	4,1	4,0	3,9	3,8	3,7	3,7	3,6	3,6	3,5	3,5	3,4	3,4	3,4
8	4,5	4,1	3,8	3,7	3,6	3,5	3,4	3,4	3,3	3,3	3,2	3,2	3,2	3,1	3,1
9	4,3	3,9	3,6	3,5	3,4	3,3	3,2	3,2	3,1	3,1	3,0	3,0	2,9	2,9	2,8
10	4,1	3,7	3,5	3,3	3,2	3,1	3,1	3,0	3,0	2,9	2,9	2,8	2,8	2,7	2,7
12	3,9	3,5	3,3	3,1	3,0	2,9	2,9	2,8	2,8	2,7	2,6	2,6	2,5	2,5	2,5
14	3,7	3,3	3,1	3,0	2,9	2,8	2,7	2,7	2,6	2,5	2,5	2,4	2,4	2,4	2,3
16	3,6	3,2	3,0	2,9	2,7	2,7	2,6	2,5	2,5	2,4	2,4	2,3	2,3	2,2	2,2
20	3,5	3,1	2,9	2,7	2,6	2,5	2,5	2,4	2,4	2,3	2,2	2,2	2.1	2,1	2,0
24	3,4	3,0	2,8	2,6	2,5	2,4	2,4	2,3	2,3	2,2	2,1	2,1	2,0	2,0	1,9
30	3,3	2,9	2,7	2,5	2,4	2,3	2,3	2,2	2,2	2,1	2,0	2,0	1,9	1,9	1,8

Таблица 9 – Значений F-критерия Фишера при уровне значимости α =0,05
k1 – число степеней свободы большей дисперсии, k2 – число степеней свободы меньшей дисперсии

k2 \ k1	1	2	3	4	5	6	8	12	24	∞
1	161,5	199,5	215,7	224,6	230,2	233,9	238,9	243,9	249,0	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,41	19,45	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,74	8,64	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,91	5,77	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,68	4,53	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,00	3,84	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,57	3,41	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,28	3,12	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,07	2,90	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,91	2,74	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,79	2,61	2,40
12	4,75	3,88	3,49	3,26	3,11	3,00	2,85	2,69	2,50	2,30
13	4,67	3,80	3,41	3,18	3,02	2,92	2,77	2,60	2,42	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,53	2,35	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,48	2,29	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,42	2,24	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,38	2,19	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,34	2,15	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,31	2,11	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,28	2,08	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,42	2,25	2,05	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,23	2,03	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,38	2,20	2,00	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,18	1,98	1,73
25	4,24	3,38	2,99	2,76	2,60	2,49	2,34	2,16	1,96	1,71
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,15	1,95	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,30	2,13	1,93	1,67
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,12	1,91	1,65
29	4,18	3,33	2,93	2,70	2,54	2,43	2,28	2,10	1,90	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,09	1,89	1,62
35	4,12	3,26	2,87	2,64	2,48	2,37	2,22	2,04	1,83	1,57
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,00	1,79	1,51
45	4,06	3,21	2,81	2,58	2,42	2,31	2,15	1,97	1,76	1,48
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	1,95	1,74	1,44

60	4,00	3,15	2,76	2,52	2,37	2,25	2,10	1,92	1,70	1,39
70	3,98	3,13	2,74	2,50	2,35	2,23	2,07	1,89	1,67	1,35
80	3,96	3,11	2,72	2,49	2,33	2,21	2,06	1,88	1,65	1,31
90	3,95	3,10	2,71	2,47	2,32	2,20	2,04	1,86	1,64	1,28
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,85	1,63	1,26
125	3,92	3,07	2,68	2,44	2,29	2,17	2,01	1,83	1,60	1,21
150	3,90	3,06	2,66	2,43	2,27	2,16	2,00	1,82	1,59	1,18
200	3,89	3,04	2,65	2,42	2,26	2,14	1,98	1,80	1,57	1,14
300	3,87	3,03	2,64	2,41	2,25	2,13	1,97	1,79	1,55	1,10
400	3,86	3,02	2,63	2,40	2,24	2,12	1,96	1,78	1,54	1,07
500	3,86	3,01	2,62	2,39	2,23	2,11	1,96	1,77	1,54	1,06
1000	3,85	3,00	2,61	2,38	2,22	2,10	1,95	1,76	1,53	1,03
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,75	1,52	1

Таблица 10 – Стандартные коэффициенты корреляции, считающиеся достоверными (по Л.С. Каминскому)

Число степеней свободы п' = n — 2	Уровень вероятности наличия связи p (%)			Число степеней свободы	Уровень вероятности наличия связи p (%)		
	95,0	98,0	99,0		95.0	98,0	99,0
1	0,997	0,999	0,999	12	0,532	0,612	0,661
2	0,950	0,980	0,990	13	0,514	0,592	0,641
3	0,878	0,934	0,959	14	0,497	0,574	0,623
4	0,811	0,882	0,917	15	0,482	0,558	0,606
5	0,754	0,833	0,874	!6	0,468	0,542	0,590
6	0,707	0,789	0,834	17	0,456	0,528	0,575
7	0,666	0,750	0,798	18	0,444	0,516	0,561
8	0,632	0,716	0,765	19	0,433	0,503	0,549
9	0,602	0,685	0,735	20	0,423	0,492	0,537
10	0,576	0,658	0,708	25	0,381	0,445	0,487
11	0,553	0,634	0,684	30	0,349	0,409	0,449

Чудновская Г.В., Саловаров В.О., Демидович А.П.

Биометрия в ихтиологии

Учебное пособие

Лицензия на издательскую деятельность
ЛР № 070444 от 11.03.98 г.
Подписано в печать 14.05.2018 г.
Тираж 20 экз.

Издательство Иркутского государственного
аграрного университета им. А.А. Ежевского
664038, Иркутская обл., Иркутский р-н,
пос. Молодежный