МИНИСТЕРСТВО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ПРАКТИЧЕСКОЕ РУКОВОДСТВО ПО СТАТИСТИЧЕСКОМУ АНАЛИЗУ В ФИЗИОЛОГИИ

Для студентов дневного отделения
3 курса биолого-почвенного факультета
по специальности 011600 «Биология»

ВОРОНЕЖ
2003

Утверждено научно-методическим советом биолого-почвенного факультета 24 сентября 2002 г. (Протокол № 11).

Составители: Рецкий М.И., Чернов В.И., Семенов С.Н., Сереженко Н.П.

Практическое руководство подготовлено на кафедре физиологии человека и животных биолого-почвенного факультета Воронежского государственного университета

Рекомендуется для студентов 3 курса дневного отделения биолого-почвенного факультета, обучающихся по специальности 011600 «Биология».

ПРЕДИСЛОВИЕ

Методическое пособие по статистическому анализу исследований в физиологии предназначено для студентов старших курсов, чтобы помочь начинающим исследователям освоить основные понятия математической статистики и более полно представить диапазон применения статистических методов.

В пособии рассматриваются основные статистические понятия, даны краткие описания методов обработки и анализа эмпирических данных: построение статистических оценок, параметрические и непараметрические методы проверки статистических гипотез, дисперсионный, корреляционный и регрессионный анализ, некоторые вопросы планирования эксперимента.

Пособие начинается с обзора элементарных (основных) понятий математической статистики, а затем более подробно на конкретных примерах из различных областей физиологии, биологии и медицины, включая лабораторные и другие виды исследований, показаны отдельные возможности и области использования программного пакета статистической диалоговой системы «*STADIA*».

Представленный материал иллюстрирован 17 таблицами и 22 рисунками. В приложении приведены таблицы основных статистических критериев.

Настоящее методическое пособие подготовлено преподавателями кафедры физиологии человека и животных биолого-почвенного факультета Воронежского государственного университета (М.И. Рецкий) и кафедры информационных систем общественного здоровья и здравоохранения Воронежской государственной медицинской академии им. Н.Н. Бурденко (В.И. Чернов, С.Н. Семенов, Н.П. Сереженко) на основе опыта проведения практических занятий со студентами, выполнения ими курсовых и дипломных работ.

Рецензент – кандидат биологических наук, доцент кафедры биофизики и биотехнологии С.Г. Резван.

СОДЕРЖАНИЕ

ОБЩИЕ СВЕДЕНИЯ О СТАТИСТИЧЕСКОМ НАБЛЮДЕНИИ

Массовые явления, изучаемые статистикой, состоят из однородных в определенном отношении *единиц наблюдения.* Совокупность таких относительно однородных, но индивидуально различных единиц, объединенных в отношении некоторых общих условий для группового изучения, называется *статистической совокупностью,* а число единиц совокупности – *объемом совокупности.*

Статистическое наблюдение предусматривает сбор сведений по заранее разработанному плану. Фактический материал, содержащийся в первичных документах учета (история болезни, карта эпидемиологического обследования и т. д.), нуждается в упорядочении и систематизации собранных данных с тем, чтобы получить из них интересующую информацию. Этот процесс называется *группировкой.* Группировка предусматривает расчленение совокупности на группы, однородные по какому-либо одному признаку (простая группировка) или по нескольким признакам (комбинационная группировка). Процесс группировки не просто технический прием, а глубоко осмысленное действие, направленное на получение объективной и полноценной информации с учетом поставленной задачи. Наиболее приемлемой формой группировки являются статистические таблицы, статистические ряды и т. д., когда числовые значения признака расположены в определенном порядке.

Что такое переменные? Переменные – это то, что можно измерять, контролировать или что можно изменять в исследованиях. Переменные отличаются многими аспектами, особенно той ролью, которую они играют в исследованиях, шкалой измерения и т.д.

ИССЛЕДОВАНИЕ ЗАВИСИМОСТЕЙ В СРАВНЕНИИ С ЭКСПЕРИМЕНТАЛЬНЫМИ ИССЛЕДОВАНИЯМИ

Большинство эмпирических исследований данных можно отнести к одному из названных типов. В исследовании корреляций (зависимостей, связей) вы не влияете (или, по крайней мере, пытаетесь не влиять) на переменные, а только измеряете их и хотите найти зависимости (корреляции) между некоторыми измеренными переменными, например, между кровяным давлением и уровнем холестерина. В экспериментальных исследованиях, напротив, вы варьируете некоторые переменные и измеряете воздействия этих изменений на другие переменные. Например, исследователь может искусственно увеличивать кровяное давление, а затем на определенных уровнях давления измерить уровень холестерина. Анализ данных в экспериментальном исследовании также приходит к вычислению «корреляций» (зависимостей) между переменными, а именно между переменными, на которые воздействуют, и переменными, на которые влияет это воздействие. Тем не менее, экспериментальные данные потенциально снабжают нас более качественной информацией. Только экспериментально можно убедительно доказать причинную связь между переменными. Например, если обнаружено, что всякий раз, когда изменяется переменная A, изменяется и переменная B, то можно сделать вывод — «переменная A оказывает влияние на переменную B», т.е. между переменными A и B имеется причинная зависимость. Результаты корреляционного исследования могут быть проинтерпретированы в каузальных (причинных) терминах на основе некоторой теории, но сами по себе не могут отчетливо доказать причинность.

Зависимые и независимые переменные. Независимыми переменными называются переменные, которые варьируются исследователем, тогда как зависимые переменные – это переменные, которые измеряются или регистрируются. Может показаться, что проведение этого различия создает путаницу в терминологии, поскольку как говорят, что «все переменные зависят от чего-нибудь». Тем не менее, однажды отчетливо проведя это различие, вы поймете его необходимость. Термины зависимая и независимая переменная применяются в основном в экспериментальном исследовании, где экспериментатор манипулирует некоторыми переменными, и в этом смысле они «независимы» от реакций, свойств, намерений и т.д., присущих объектам исследования. Некоторые другие переменные, как предполагается, должны «зависеть» от действий экспериментатора или от экспериментальных условий. Иными словами, зависимость проявляется в ответной реакции исследуемого объекта на оказанное на него воздействие. Отчасти в противоречии с данным разграничением понятий находится использование их в исследованиях, где вы не варьируете независимые переменные, а только приписываете объекты к «экспериментальным группам», основываясь на некоторых их априорных свойствах. Например, если в эксперименте мужчины сравниваются с женщинами относительно числа лейкоцитов (WCC), содержащихся в крови, то «Пол» можно назвать независимой переменной, а WCC зависимой переменной.

Шкалы измерений. Переменные различаются также тем, «насколько хорошо» они могут быть измерены или, другими словами, как много измеряемой информации обеспечивает шкала их измерений. Очевидно, в каждом измерении присутствует некоторая ошибка, определяющая границы «количества информации», которое можно получить в данном измерении. Другим фактором, определяющим количество информации, содержащейся в переменной, является тип шкалы, в которой проведено измерение. Различают следующие типы шкал: a)номинальная; b)порядковая (ординальная); c)интервальная; d)относительная (шкала отношения).

Соответственно имеем четыре типа переменных: (a) номинальные, (b) порядковые (ординальные), (c) интервальные и (d) относительные.

Номинальные переменные используются только для качественной классификации. Это означает, что данные переменные могут быть измерены только в терминах принадлежности к некоторым, существенно различным классам; при этом вы не сможете определить количество или упорядочить эти классы. Например, 2 индивидуума различимы в терминах переменной А (принадлежат к разным национальностям). Типичные примеры номинальных переменных – пол, национальность, цвет, город и т.д. Часто номинальные переменные называют категориальными.

Порядковые переменные позволяют ранжировать (упорядочить) объекты, указав, какие из них в большей или меньшей степени обладают качеством, выраженным данной переменной. Однако они не позволяют сказать «на сколько больше» или «на сколько меньше». Порядковые переменные иногда также называют ординальными. Типичный пример порядковой переменной – социально-экономический статус семьи. Так, возможно, что уровень семьи соответствует верхнему среднему уровню выше среднего уровня, однако сказать, что разница между ними равна, скажем, 18% мы не сможем. Само расположение шкал в следующем порядке: номинальная, порядковая, интервальная является хорошим примером порядковой шкалы.

Интервальные переменные позволяют не только упорядочивать объекты измерения, но и численно выразить и сравнить различия между ними. Например, температура, измеренная в градусах Фаренгейта или Цельсия, образует интервальную шкалу: температура 40 градусов выше, чем температура 30 градусов, и увеличение температуры с 20 до 40 градусов вдвое больше увеличения температуры от 30 до 40 градусов.

Относительные переменные очень похожи на интервальные переменные. В дополнение ко всем свойствам переменных, измеренных в интервальной шкале, их характерной чертой является наличие определенной точки абсолютного нуля, таким образом, для этих переменных являются обоснованными предложения типа: x в два раза больше, чем y. Типичными примерами шкал отношений являются измерения времени или пространства. Например, температура по Кельвину образует шкалу отношения, и вы можете не только утверждать, что температура 200 градусов выше, чем 100 градусов, но и что она вдвое выше. Интервальные шкалы (например, шкала Цельсия) не обладают данным свойством шкалы отношения. Заметим, что в большинстве статистических процедур не делается различия между свойствами интервальных шкал и шкал отношения.

СВЯЗИ МЕЖДУ ПЕРЕМЕННЫМИ

Независимо от типа, две или более переменных связаны (зависимы) между собой, если наблюдаемые значения этих переменных распределены согласованным образом. Другими словами, переменные зависимы, если их значения систематическим образом согласованы друг с другом в имеющихся наблюдениях. Например, переменные «Пол» и «Число лейкоцитов» могли бы рассматриваться как зависимые, если бы большинство мужчин имело высокий уровень «Число лейкоцитов», а большинство женщин — низкий «Число лейкоцитов», или наоборот. «Рост» связан с «Весом», потому что обычно высокие индивиды тяжелее низких; «IQ» (коэффициент интеллекта) связан с «Количеством ошибок» в тесте, т.к. люди с высоким значением IQ делают меньше ошибок и т.д.

Статистическому наблюдению, представляющему собой начальный этап исследования, должно уделяться большое внимание, так как от полноты и качества собранных данных во многом зависят и выводы. Статистическое наблюдение осуществляется как посредством отчетности, так и посредством специально организованных исследований.

Результат или исход отдельного испытания (статистического наблюдения) называется *событием*. Реализация его, как правило, возможна при наличии комплекса условий, необходимых для того, чтобы данное событие произошло. Числовая мера возможности осуществления определенного события в некотором количестве случаев из общего числа возможных называется *вероятностью (р).*

Единицы совокупности обычно обладают многими признаками, которые имеют различное выражение у отдельных единиц. Так, например, больные различаются по полу, возрасту, профессии и т. д. Признаки, принимающие различные значения у отдельных единиц совокупности, называются *варьирующими,* а отдельные числовые значения варьирующего признака – *вариантами* (x_i). Варьирование – характерное свойство всего живого. Варьирующие признаки подразделяются на *атрибутивные (качественные)* и *количественные*. Признак называется атрибутивным, если отдельные его значения выражаются в виде состояния, свойств и т. д., присущих явлению (профессия больных, вид микроорганизмов, цвет мочи и т. д.). К количественным относят те признаки, отдельные значения (варианты) которых выражаются в виде чисел (количество лейкоцитов, титр антител, количество микроорганизмов и т. д.).

Расположение вариант в порядке возрастания (уменьшения) их числовых значений, показывающее закономерность распределения единиц изучаемой совокупности, называется *вариационным рядом.* Вариационные ряды, выраженные в виде целых чисел, называются *дискретными,* а при выражении количественных признаков в виде интервалов – *интервальными.*

Для обобщающей характеристики качественно однородных совокупностей используются *средние величины*, благодаря применению которых статистика, имея дело с массовыми явлениями, получает возможность переходить от единичного к общему, от случайного – к закономерному. Одной из таких величин является *средняя арифметическая* (обозначаемая как $\bar{x}$ или *М*), получающаяся при делении суммы однородных величин, характеризующих значение определенного признака, на число вариант.

Пример. Значения вариант составляют: 8, 16, 9, 3, 14, 30, 26. Сумма величин равна 105, а $\bar{x}$=105:7=15. Однако средняя арифметическая сама по себе ничего не говорит о том вариационном ряде, из которого она вычислена, так как колебания значений вариант внутри ряда могут быть различны.

Например: 1-й вариационный ряд: 8, 16, 9, 3, 14, 30, 26; $\bar{x}$=15;

2-й вариационный ряд: 2, 8, 3, 5, 7, 22, 58; $\bar{x}$=15.

В обоих случаях средние арифметические равны, однако они получены из рядов с различным размахом. *Размах* - разность между максимальным и минимальным значением вариант. В первом случае размах равен 30-3=27, во втором – 58-2=56. Чем больше индивидуальные значения вариант различаются между собой, тем больше они отличаются и от средней арифметической, являющейся центром группирования вариант данного ряда. Величина отклонения каждой варианты от средней арифметической называется *линейным отклонением* ($\bar{d}$), т. е. $\bar{d} = \bar{x} - x_i$. Однако при анализе степени отклонения (рассеяния) вариант используется *среднее квадратическое отклонение* (s), так как оно характеризует вариацию не отдельно взятых, а всех вариант ряда. Величина s^2 или D носит название *дисперсии.*

$$\sigma = \pm\sqrt{\frac{\sum \bar{d}^2}{n}}; \qquad \text{при } n<30 \quad \sigma = \pm\sqrt{\frac{\sum \bar{d}^2}{n-1}},$$

где s – среднее квадратическое отклонение; $\bar{d}$ – отклонение вариант от средней арифметической; n – число вариант.

Совокупность значений вариант (x_i) и их вероятностей (p) называется *распределением.* В статистике известен ряд законов распределения случайных величин: закон нормального распределения, закон биномиального распределения, закон Пуассона, распределение Максвелла и др.

Нормальное распределение занимает важное место в статистике. Многие медико-биологические признаки, характеризующиеся непрерывной вариацией, являются суммой большого числа независимых слагаемых и с достаточной степенью точности следуют закону нормального распределения. При графическом изображении распределения вариант получается симметричная куполообразная кривая, имеющая максимум в точке средней арифметической, называемая *кривой нормального распределения* (рис. 1).

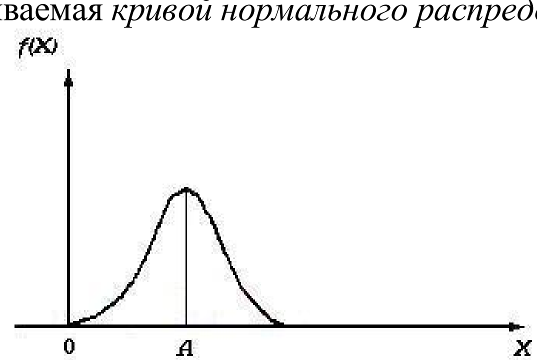


Рис.1. График плотности нормального распределения

Если площадь, ограниченную кривой нормального распределения, принять за 1 (или за 100 %), то можно рассчитать площадь, заключенную между кривой и любыми двумя ординатами.

Установлено, что площадь между ординатами, проведенными на расстоянии 1s с каждой стороны от $\bar{x}$, составляет 0,683 всей площади. Это означает, что 68,3 % всех исследованных единиц (частот) отклоняется от $\bar{x}$ не более чем на 1s, т. е. находится в пределах x±s. Площадь, заключенная между ординатами, проведенными на расстоянии 2σ от x, составляет 0,954, т. е. 95,4 % всех единиц совокупности находится в пределах x±2s.

И, наконец, 0,997 или 99,7 % всех единиц находятся в пределах x±3s. Таким образом, то, что находится в пределах 3s, относится к данному ряду, то, что за пределами 3s, вероятнее всего, к этому ряду уже не относится.

Для изучения закономерностей вариации при нормальном распределении признаков пользуются *нормированным отклонением (t),* представляющим собой отклонение той или иной варианты от средней арифметической, отнесенное к величине среднего квадратического отклонения.

$$t = \frac{\bar{x} - x_i}{\sigma} \text{, отсюда } \bar{x} - x_i = t\sigma$$

Величина t-критерия характеризует распределение выборочных средних в нормальной генеральной совокупности в зависимости от объема выборки, а t-распределение зависит только от двух величин: нормированного отклонения и числа степеней свободы n', т. е. числа свободно варьирующих признаков ($n'=n-1$). С увеличением числа наблюдений t-распределение быстро приближается к нормальному и уже при $n=30$ не отличается от него. Следовательно, для $n \geq 30$ величина t распределяется нормально, тогда как при $n<30$ (малая выборка) распределение t зависит от числа наблюдений. Для практического использования t-распределения имеется специальная таблица (приложение 1), в которой содержатся значения t для разных уровней значимости и объема выборки. Поскольку варианты в вариационном ряду распределяются в пределах трех сигм, то и значения t для отдельных вариант колеблются в пределах ±3.

При нормальном распределении признаков, чем ближе значения вариант к средней арифметической, тем чаще они встречаются; чем дальше от средней арифметической, тем реже частота их появления. Таким образом, вариационный ряд с характерным для него расположением большинства вариант вблизи его центральной части и рассеянием к краям ряда является в то же время и распределением вероятностей. Как отмечалось выше, вероятность (p) – это числовая мера объективной возможности осуществления определенного события в некотором количестве случаев из общего числа возможных. Обозначив абсолютную численность вариант через n, численность случаев появления интересующего признака через m, *доля* (p) вариант, обладающих данным признаком, выразится формулой:

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

Оставшаяся доля (q) вариант той же совокупности при альтернативной группировке данных (например, заболел – не заболел и т. д.) свидетельствует о частоте непоявления события. Поскольку p+q=l, то в числовом выражении вероятность представляет собой число, заключенное между 0 и 1 и, стало быть, выраженное в долях единицы, а при умножении доли на 100, 1000, 10000, 100000 выражается соответственно в процентах (%), промилле (%), продецимилле (%), просантимилле (%). При p=1 событие называется *достоверным*, т. е. единственно возможным при наличии комплекса условий, необходимых для его реализации. При p=0 событие считается *невозможным*. Если же событие в данных условиях может произойти и не произойти, а при многократных испытаниях обязательно наступает, то оно называется *возможным (случайным)*. Таким образом, количественной характеристикой вероятности того или иного явления может быть его относительная частота. Следовательно, в вариационном ряду отдельным значениям вариант (x_i) можно придать соответствующие вероятности (p). Существенно важны две вероятности, числовые значения которых в долях единицы составляют 0,95 и 0,99, или 95% и 99%. Они называются *доверительными вероятностями*. При доверительной вероятности 0,95 (95%) любая случайно взятая величина вариационного ряда в случае нормального распределения будет отклоняться от средней арифметической не более чем на $1{,}96s$ или, наоборот, с вероятностью 0,05 (5 %) она будет находиться за пределами $1{,}96s$. С вероятностью же, равной 0,99 (99%), она будет отклоняться от x не более чем на $2{,}58s$, или, наоборот, вероятность выхода за пределы $2{,}58s$ равна 0,01 (1 %).

Определенным значениям вероятностей соответствуют *уровни значимости (р)*, свидетельствующие о частоте получения случайного отклонения от установленных с определенной вероятностью результатов, т. е. в каком проценте случаев (или с какой вероятностью) все же возможна ошибка в результатах, выводах и т. д. Вероятности 0,95 (95 %) соответствует уровень значимости 0,05 (5%); вероятности 0,99 (99%) – 0,01 (1%). По отношению к закономерностям распределения признаков это означает, что выход за пределы принятых границ возможен в порядке случайности с вероятностью 0,05 (5%) и 0,01 (1 %), т. е. риск ошибки в выводах составляет 5 % и 1 % соответственно.

Статистическое наблюдение может охватывать всех членов совокупности, а может ограничиваться обследованием лишь некоторой части их. Совокупность, из которой отбирается часть ее членов для изучения, называется *генеральной*, а отобранная часть – *выборочной*, или *выборкой*. Объем генеральной совокупности (N) теоретически мыслится как бесконечно большое множество относительно однородных единиц. Объем выборочной совокупности (n) может быть различным по величине, но не должен быть меньше двух единиц.

Выборочный метод является основным при изучении статистической совокупности, однако, он должен дать такую информацию, которая позволяла бы судить о состоянии генеральной совокупности, т. е. выборка должна быть достаточно *представительной (репрезентативной)*. Репрезентативность выборки зависит от ряда факторов, среди которых однородность исходной совокупности, объем выборки, способы отбора единиц и т. д. При выборочном обследовании обычно изучается доля единиц, обладающих тем или иным признаком (например, частота обнаружения стафилококков в воздухе), а также средний размер того или иного признака у единиц совокупности (например, количество стафилококков в воздухе). При этом между полученными выборочными средними возникают определенные расхождения по отношению к средним для генеральной совокупности, т. е. возникают *ошибки*.

Ошибки подразделяются на регистрационные (неправильный учет данных) и ошибки репрезентативности. *Ошибки репрезентативности* представляют собой расхождения между обобщающими показателями отобранной части совокупности и всей совокупности в целом при условии правильной регистрации данных. Ошибки репрезентативности могут быть систематическими и случайными. *Систематические ошибки* возникают при нарушении принципов проведения выборочного наблюдения, например вследствие произвольной замены попавших в выборку единиц другими. *Случайные ошибки (ошибки выборки)* - это расхождение между выборочной средней и генеральной средней при условии правильного отбора и регистрации данных. Они возникают в силу того, что выборочная совокупность не воспроизводит точно генеральную совокупность. О величине возможного отклонения выборочной средней от генеральной средней судят по *средней ошибке выборки* (s). Для установления границ, в которых находится генеральная средняя, используется предельная ошибка выборки (Δ); $\Delta = t \cdot s$.

Значение вероятности наступления события, с которой можно гарантировать надежность результатов выборки, находят по специальной таблице (приложение 1). В ней приводятся значения нормированного отклонения (t) и соответствующие им уровни значимости (P) при заданном объеме выборки (n).

ТИПЫ ОШИБОК И МОЩНОСТЬ КРИТЕРИЯ

Критическая область – это множество таких исходов некоторого статистического эксперимента, которые приводят нас к отклонению нулевой гипотезы. Если нулевая гипотеза справедлива и принят уровень значимости α, то с вероятностью α исход эксперимента попадет в критическую область; в этом случае нулевая гипотеза будет ошибочно отклонена и тем самым будет допущена *ошибка 1 рода*. Иногда возникает ситуация, когда нулевая гипотеза ложна, а исход эксперимента не попал в критическую область. В таком случае нулевая гипотеза ошибочно принимается и тем самым допускается *ошибка II рода*.

Естественно стремление минимизировать вероятности ошибок рассмотренных типов. Снижая уровень значимости α, можно легко сократить вероятность возникновения ошибки I рода, но в таком случае возрастет вероятность ошибки II рода. В связи с этим вводят понятие *мощности* критерия, которую определяют как вероятность отклонений нулевой гипотезы. Эта вероятность зависит от реального значения рассматриваемого параметра совокупности. Поскольку это значение заранее неизвестно, рассматривают *кривую мощности*, которая показывает соответствующее значение мощности критерия для каждого возможного значения параметра. Очевидно, что точки идеальной кривой мощности имеют ординату, равную единице, для всех значений параметра, кроме тех, которые соответствуют нулевой гипотезе. На практике получить подобный результат невозможно, но обычно удается повысить мощность критерия до любого желаемого уровня, соответственно увеличив объем выборки.

Следует предупредить об опасности, связанной с применением нескольких статистических критериев при анализе одних и тех же данных. Если к одним и тем же данным применяют два различных критерия для проверки одной и той же нулевой гипотезы (или двух сходных гипотез) и в каждом случае принимается уровень значимости, равный, например, 5%, то вероятность того, что хотя бы по одному из критериев нулевая гипотеза будет ошибочно отклонена, превосходит 5 %. Следует воспользоваться лишь одним критерием, желательно более мощным.

Иногда возникает необходимость проверки двух различающихся гипотез при использовании одних и тех же данных (например, гипотезы о значениях среднего и дисперсии некоторой нормально распределенной совокупности). Если для обоих критериев принимается 5 % уровень значимости и обе нулевые гипотезы справедливы, то вероятность ошибочного отклонения хотя бы одной из нулевых гипотез значительно превосходит 5 % и часто бывает близка к 10 %.

СТАТИСТИЧЕСКИЕ СРАВНЕНИЯ

Решение той или иной задачи не обходится, как правило, без сравнения статистических показателей, отображающих размеры и количественные соотношения анализируемых явлений. В статистике для этих целей применяется так называемая *нулевая гипотеза*, т. е. предположение о том, что разница между генеральными параметрами сравниваемых групп равна нулю, а различия, которые наблюдаются между выборочными показателями, носят случайный характер. Истинность принятой гипотезы проверяется с помощью *критериев значимости*, т. е. специально выработанных случайных величин, функции распределения которых известны. Обычно для каждого критерия составляются таблицы, в которых содержатся критические величины, отвечающие определенному объему выборки и принятым уровням значимости. В биологических исследованиях принимается 5 % уровень значимости, которому отвечает нормированное отклонение $t = 1{,}96$ ($\approx 2{,}0$) при объеме выборки больше 30 единиц в случае нормального распределения признаков. Например, если окажется, что $P>0{,}05$, то отвергнуть нулевую гипотезу нет оснований; при $P<0{,}05$ нулевая гипотеза отвергается, т. е. с вероятностью более 95 % разница между выборочными показателями считается статистически значимой (достоверной). Могут применяться 1 % или 0,1 % уровни значимости.

Для определения границ доверительного интервала выборочной доли (p) определяется ошибка выборки для доли (s_p) и предельная ошибка (Δ).

$$s_p = \pm\sqrt{\frac{p \cdot q}{n}}; \quad \text{при } n<30 \quad s_p = \pm\sqrt{\frac{p \cdot q}{n-1}},$$

где s_p – ошибка выборки для доли; p – доля вариант, обладающих изучаемым признаком; q – доля вариант, не обладающих изучаемым признаком; n – объем выборки.

Формула предельной ошибки: $\Delta = t \cdot s_p$.

Следовательно, границы доверительного интервала будут находиться в пределах: $p \pm t \cdot s_p$. Другими словами, выборка репрезентативна, если полученная величина доли превышает ошибку выборки (s_p) на величину t-критерия при уровне значимости P<0,05 (P<0,02; P<0,01; P<0,001) для данного числа степеней свободы *(n')*.

Пример. Из 52 исследований воздушной среды в послеродовых палатах родильного дома золотистые стафилококки выделены в 16 случаях ($p_{\text{посл.п.}}$ = 30,8%), а из 28 исследований воздушной среды в палатах новорожденных – в 16 ($p_{\text{п.нов.}}$= 57,1 %). Ошибки выборки составят:

$$s_{p_{\text{послоп}}} = \pm\sqrt{\frac{30{,}8 \cdot (100-30{,}8)}{52}} = \pm 6{,}4;$$

$$s_{p_{\text{п.нов.}}} = \pm\sqrt{\frac{57{,}1 \cdot (100-57{,}1)}{28-1}} = \pm 9{,}5.$$

Для P<0,05 и соответственно t=2,0 границы доверительного интервала для показателя частоты обнаружения стафилококков в детских палатах составят: нижняя – 30,8-2·6,4=18,0, верхняя – 30,8+2·6,4=43,6. Поскольку в палатах новорожденных выполнено меньше 30 исследований (28), величина нормированного отклонения t берется из приложения 1 при уровне значимости P<0,05 и соответствующем числе степеней свободы (n'). В приведенном примере при n = 28-1=27 и уровне значимости P<0,05 t=2,05. Следовательно, границы доверительного интервала будут составлять: нижняя – 57,1-2,05·9,5=37,6; верхняя – 57,1+2,05·9,5=76,6.

Т-КРИТЕРИЙ ДЛЯ НЕЗАВИСИМЫХ ВЫБОРОК

t-критерий (Стьюдента) является наиболее часто используемым методом обнаружения различия между средними двух выборок. Например, *t*-критерий можно использовать для сравнения средних показателей группы пациентов, принимавших определенное лекарство, с контрольной группой, где принималось безвредное лекарство. Теоретически, *t*-критерий может применяться, даже если размеры выборок очень небольшие (например, n=10, но можно исследовать выборки меньшего размера), если переменные нормально распределены (внутри групп), а дисперсии наблюдений в группах не слишком различны. Предположение о нормальности можно проверить, исследуя распределение (например, визуально с помощью гистограммы или критериев согласия) или применяя какой-либо критерий нормальности. Равенство дисперсий в двух группах можно проверить с помощью F критерия. Если условия применимости t-критерия не выполнены, следует использовать непараметрические альтернативы t-критерия.

P-уровень значимости t-критерия равен вероятности ошибочно отвергнуть гипотезу о равенстве средних двух выборок, когда в действительности эта гипотеза имеет место. Иными словами, он равен вероятности ошибки принять гипотезу о неравенстве средних, когда в действительности средние равны. Некоторые исследователи предлагают, в случае, когда рассматриваются отличия только в одном направлении (например, рассматривается альтернатива: среднее в первой группе больше (меньше), чем среднее во второй), использовать *одностороннее* t-распределение и делить *p-уровень* двустороннего *t-критерия* пополам. Другие предлагают всегда работать со стандартным двусторонним *t-критерием*.

Расположение данных. Чтобы применить t-критерий для независимых выборок, требуется, по крайней мере, одна независимая (*группирующая*) переменная (например, пол: *мужчина/женщина*) и одна зависимая переменная (например, тестовое значение некоторого показателя - кровяное давление, число лейкоцитов и т.д.). С помощью специальных значений независимой переменной (эти значения называются *кодами*, например, *мужчина* и *женщина*) данные разбиваются на две группы. Можно произвести анализ следующих данных с помощью t-критерия, сравнивающего средний уровень суточного выделения натрия с мочой для мужчин и женщин (табл.1).

Таблица 1

Наблюдения	Пол	Na, мМоль/сутки
1	мужчина	111
2	мужчина	110
3	мужчина	109
4	женщина	102
5	женщина	104
	среднее для мужчин = 110 среднее для женщин = 103	

БОЛЕЕ СЛОЖНЫЕ ГРУППОВЫЕ СРАВНЕНИЯ

На практике часто приходится сравнивать более двух групп данных (например, имеется *лекарство 1*, *лекарство 2* и *успокоительное лекарство*) или сравнивать группы, созданные более чем одной независимой переменной (например, *Пол*, тип *Лекарства* и *Доза*). В таких исследованиях следует использовать дисперсионный анализ, который можно рассматривать как обобщение t-критерия. Фактически в случае однофакторного сравнения двух групп, дисперсионный анализ дает результаты, идентичные t-критерию. Однако, если план существенно более сложный, дисперсионный анализ предпочтительнее t-критерия.

t-КРИТЕРИЙ ДЛЯ ЗАВИСИМЫХ ВЫБОРОК

Внутригрупповая вариация. Степень различия между средними в двух группах зависит от внутригрупповой вариации (дисперсии) переменных. В зависимости от того, насколько различны эти значения для каждой группы, «грубая разность» между групповыми средними показывает более сильную или более слабую степень зависимости между независимой (*группирующей*) и зависимой переменными. Например, если средний уровень выделения натрия с мочой равнялся 110 мМоль/сутки для мужчин и 103 мМоль/сутки для женщин, то разность внутригрупповых средних только на величину 7 будет чрезвычайно важной, когда все значения этого показателя для мужчин лежат в интервале от 111 до 109, а все значения для женщин - в интервале 102 - 104. В этом случае можно довольно хорошо предсказать средний уровень выделения натрия с мочой (значение зависимой переменной) исходя из пола субъекта (независимой переменной). Однако если та же разность 7 получена из сильно разбросанных данных (например, изменяющихся в пределах от 0 до 200), то этой разностью вполне можно пренебречь. Таким образом, можно сказать, что уменьшение *внутригрупповой вариации* увеличивает чувствительность критерия.

Для зависимых выборок *t*-критерий очень полезен в тех ситуациях, когда важный источник *внутригрупповой вариации* (или *ошибки*) может быть легко определен и исключен из анализа. Например, это относится к экспериментам, в которых две сравниваемые группы основываются на одной и той же совокупности наблюдений (субъектов), которые тестировались *дважды* (например, *до* и *после* воздействия, *до* и *после* приема лекарства). В подобных экспериментах значительная часть внутригрупповой изменчивости (вариации) в обеих группах может быть объяснена индивидуальными различиями субъектов. На самом деле, такая ситуация не слишком отличается от той, когда сравниваемые группы совершенно независимы, где индивидуальные отличия также вносят вклад в *дисперсию ошибки*. Однако в случае независимых выборок, вы ничего не сможете поделать с этим, т.к. не сможете определить (или «удалить») часть вариации, связанную с индивидуальными различиями субъектов. Если та же самая выборка тестируется дважды, то можно легко исключить эту часть вариации. Вместо исследования каждой группы отдельно и анализа исходных значений, можно рассматривать просто разности между двумя измерениями (например, «до приема лекарства» и «после приема лекарства») для каждого субъекта. Вычитая первые значения из вторых (для каждого субъекта) и анализируя затем только эти "чистые (парные) разности", вы исключите ту часть вариации, которая является результатом различия в исходных уровнях индивидуумов. Именно так и проводятся вычисления в *t*-критерии для зависимых выборок. В сравнении с *t*-критерием для независимых выборок, такой подход дает всегда «лучший» результат (критерий становится более чувствительным).

Теоретические предположения *t*-критерия для независимых выборок относятся также к критерию для зависимых выборок. Это означает, что попарные разности должны быть нормально распределены. Если это не выполняется, то можно воспользоваться одним из альтернативных непараметрических критериев.

Расположение данных. Вы можете применять *t*-критерий для зависимых выборок к любой паре переменных в наборе данных. Заметим, применение этого критерия мало оправдано, если значения двух переменных несопоставимы. Например, если вы сравниваете средний уровень выделения Na у людей в выборке до и после физической нагрузки, но используете различные методы вычисления количественного показателя или другие единицы во втором измерении, то высоко значимые значения *t*-критерия могут быть получены искусственно, именно за счет изменения единиц измерения. Следующий набор данных может быть проанализирован с помощью *t*-критерия для зависимых выборок (табл.2).

Таблица 2

Уровень выделения натрия с мочой (мМоль/сутки)
до и после физической нагрузки

Наблюдения	До нагрузки	После нагрузки
1	111,9	113
2	109	110
3	143	144
4	101	102
5	80	80,9
	средняя разность между «до» и «после» = 1	

Средняя разность между показателями в двух столбцах относительно мала (d=1) по сравнению с разбросом данных (от 80 до 143, в первой выборке). Тем не менее, *t*-критерий для зависимых выборок использует только парные разности, «игнорируя» исходные численные значения и их вариацию. Таким образом, величина этой разности *1* будет сравниваться не с разбросом исходных значений, а с разбросом *индивидуальных разностей*, который относительно мал: 0,2 (от 0,9 в наблюдении 5 до 1,1 в наблюдении 1). В этой ситуации разность 1 очень большая и может привести к значимому *t*-значению.

Матрицы t-критериев. *t*-критерий для зависимых выборок может быть вычислен для списков переменных и просмотрен далее как матрица. Пропущенные данные при этом обрабатываются либо построчно, либо попарно, точно так же, как при вычислении корреляционных матриц. Все те предостережения, которые относились к использованию этих методов обработки пропусков при вычислении матриц коэффициентов корреляций, остаются в силе при вычислении матриц *t*-критериев. Возможно:

1. Появление артефактов (искусственных результатов) из-за попарного удаления пропусков в *t*-критерии;

2. Возникновение чисто «случайно» значимых результатов.

ОЦЕНКА ДОСТОВЕРНОСТИ РАЗЛИЧИЙ МЕЖДУ РАЗМЕРАМИ ДОЛЕЙ

На практике, как правило, приходится сравнивать две или несколько долей между собой. Последовательность расчетов рассмотрим с использованием данных рассмотренного на стр.10 примера:

1. Определяется разность (d) между $p_{посл.п}$. и $p_{п.нов.}$:

$$d = p_1 - p_2;\quad d = 57{,}1 - 30{,}8 = 26{,}3.$$

2. Вычисляется средняя ошибка разности (s_d) по формуле:

$S_d = \pm\sqrt{s^2_{p_1} + s^2_{p_2}}$. В нашем примере $S_d = 11{,}5$.

3. Определяется величина t-критерия:

$$t = \frac{d}{s_d};\qquad t = \frac{26{,}3}{11{,}5} = 2{,}29.$$

4. Полученная величина t-критерия сравнивается с критическим значением при уровне значимости P<0,05 (<0,02; <0,01; <0,001) по приложению 1 с учетом числа степеней свободы (n’). Число степеней свободы показывает количество свободно варьирующих членов статистической совокупности, способных принимать любые произвольные значения.

$$n' = (n_1 - 1) + (n_2 - 1)$$

В нашем примере n'= (52-1) + (28-1) =78, т. е. полученную величину t=2,29 необходимо сравнить с критическим значением t-критерия для 78 варьирующих признаков (n>30). Так как полученная величина t=2,29 превышает 1,96, разность между сравниваемыми показателями признается достоверной. Таким образом, обсемененность золотистыми стафилококками воздуха в палатах новорожденных выше на 26,3 %, чем в послеродовых палатах (57,1 ±9,5% и 30,8 ±6,4 % соответственно, P<0,05).

Сравнительному анализу подлежат репрезентативные относительные показатели. Однако в исследовательской практике иногда приходится использовать для сравнительного анализа и единичные абсолютные величины, несмотря на то, что при переводе их в относительные величины последние оказываются непредставительными. Тем не менее, если явление имеет место, не считаться с ним нельзя, особенно в случаях, когда оно вообще не должно встречаться при определенных обстоятельствах (например, выделение патогенных микроорганизмов и др.).

Для характеристики относительной доли (удельного веса) части в целом используются *экстенсивные показатели.* Они выражаются, как правило, в процентах, т. е. вся совокупность принимается за 100 %, а составляющие ее статистические единицы определенную часть от 100 %. Экстенсивные показатели широко применяются для анализа структуры заболеваний, причин смерти, возрастно-половой структуры населения и т. д.

В качестве примера можно привести результаты обработки данных по структуре заболеваемости детей в зимне-весенний период года (табл.3).

Последовательность расчетов:

1. Вычисляем удельный вес (p_1) больных гриппом: 150:178·100 = 84,3%
2. Определяем ошибку по формуле:

$$s_p = \pm\sqrt{\frac{p \cdot q}{n}} = \sqrt{\frac{84{,}3 \cdot (100 - 84{,}3)}{100}} = 2{,}7.$$

Таблица 3

Структура заболеваемости детей в зимне-весенний период

Заболевание	Всего случаев	Удельный вес, %
Грипп	150	84,3
Сальмонеллез	10	5,6
Скарлатина	12	6,7
Коклюш	4	2,3
Корь	2	1,1
Всего	178	100,0

Оценка репрезентативности доли, равной 0 или 100%. В тех случаях, когда полученный при выборочных исследованиях показатель частоты явления равен 100 % (или 0, что также приравнивается к 100 %), это не означает, что данный результат можно отнести ко всей генеральной совокупности. Ошибка выборки в таких случаях определяется по формуле:

$$s_p = \frac{t^2 \cdot 100}{n + t^2},$$

где s_p – ошибка выборки; t – нормированное отклонение; n – объем выборочной совокупности.

Пример. При постановке серологической реакции 35 детям из 2500 данного возраста она оказалась у всех положительной (100%). Значит ли это, что среди остальных нет детей с отрицательной реакцией?

По приложению 1 при n>30 и уровне значимости 0,05 величина t=2,0. Следовательно:

$$s_p = \frac{2^2 \cdot 100}{35 + 2^2} = 10{,}3$$

Таким образом, среди остальных детей процент лиц с отрицательной реакцией может быть 2500·10,3 = 258 человек, а с положительной реакцией: (100 – 10,3) · 2500 = 2242 человека.

ОПРЕДЕЛЕНИЕ НЕОБХОДИМОГО ОБЪЕМА ВЫБОРКИ ПРИ АЛЬТЕРНАТИВНОМ ВАРЬИРОВАНИИ ПРИЗНАКОВ

Стремление к большому числу исследований не всегда оправдано, так как убедительные результаты можно получить и при минимально допустимом объеме наблюдений. Необходимый объем выборочной совокупности, обеспечивающий ее репрезентативность, может быть установлен заранее. При альтернативном распределении исходных данных, когда возможны только два исхода (например, заболел – не заболел и т. д.), необходимый объем исследований определяется по формуле:

$$n = \frac{t^2 \cdot p \cdot q}{\Delta^2},$$

где n – необходимый объем выборки; p – показатель частоты явления (доля),%; q = (100 – p),%; t – нормированное отклонение; Δ – предельная ошибка выборки, %.

Пример. За период проведения исследований 936 человек зарегистрировано 115 человек с нарушениями функции сердечно-сосудистой системы. Какое количество кардиограмм (n) при Δ=3% необходимо проанализировать для характеристики состояния сердечной деятельности у обследованного контингента людей?

По приложению 1 при n>30 и P<0,05 t=2,0. Следовательно:

$$p = \frac{115 \cdot 100\%}{936} = 12{,}3\%; \qquad n = \frac{4 \cdot 12{,}3 \cdot 87{,}7}{9} = 479.$$

Если p и q неизвестны, берется наибольшая возможная величина произведения, т. е. 250 (50 % ·50 %):

$$n = \frac{2^2 \cdot 50 \cdot 50}{3^2} = 333$$

СРЕДНИЕ ВЕЛИЧИНЫ

Наряду с показателями частоты явлений количественная сторона последних характеризуется *средними величинами: средней арифметической* ($\bar{x}$) и *средней геометрической* **(**$x_{геом}$**.)**. В ряде случаев для характеристики размеров признака применяют *моду* ***(Мо)*** и *медиану* ***(Ме).***

Отличительной особенностью средних величин является то, что в них взаимно погашаются индивидуальные различия признака у отдельных единиц изучаемой совокупности, в результате чего представляется возможность охарактеризовать общие свойства изучаемых явлений, закономерности их развития.

Средняя арифметическая получается при делении суммы однородных величин, характеризующих значение определенного признака, на число вариант. Однако прежде чем приступить к ее вычислению, необходимо убедиться в отсутствии в вариационном ряду так называемых «выскакивающих» вариант, которые в значительной степени могут искажать конечные результаты.

«Выскакивающие» варианты значительно отличаются по величине от других и нетипичны для данного ряда. Как правило, они являются следствием ошибочных записей или погрешностей в экспериментальных исследованиях. Если установить достоверно причину появления «выскакивающей» варианты невозможно, ее исключают из вариационного ряда путем расчетов. Методов определения «выскакивающих» вариант несколько.

Рассмотрим наиболее простой из них.

Пример. В шести пробах сыворотки крови у крыс была определена активность аланинаминотрансферазы (АлАТ) 8, 12, 16, 32, 140 и 32 Е/л.

В данном вариационном ряду обращает на себя внимание цифра 140, которая резко отличается по величине от других и существенно влияет на размер средней арифметической ($\bar{x}$=40 с учетом варианты 140 и $\bar{x}$=20 – без нее). Есть основания проверить варианту 140 на соответствие ее данному вариационному ряду. Для этого варианты располагаются в порядке возрастающих значений:

Активность АлАТ, Е/л	8	12	16	32	32	140
№ пробы	1	2	3	4	5	6

Цифровые данные подставляются в формулу:

$$\frac{x_6 - x_5}{x_6 - x_1} = \frac{140 - 32}{140 - 8} = 0{,}818,$$

т. е. в числителе – разность между максимальной (140) и предшествующей ей по величине (32) вариантами (порядковые номера 6 и 5), в знаменателе – разность между величинами максимальной (140) и минимальной (8) вариант (порядковые номера 6 и 1). Полученное значение 0,818 оценивается **по приложению 2.** Для числа наблюдений 6 и $P<0{,}05$ критический уровень составляет 0,560. Поскольку полученная величина 0,818 превышает 0,560, варианта 140 достоверно отличается от других и должна быть исключена из вариационного ряда.

Практически не приходится прибегать к исключению минимальных значений вариант, так как даже резкое отличие их от других существенно не влияет на значение средних величин. Однако при необходимости такие расчеты выполняются по формуле:

$$\frac{x_2 - x_1}{x_6 - x_1} = \frac{12 - 8}{140 - 8} = 0{,}030$$

Полученная в данном случае величина 0,030 значительно ниже критического уровня 0,560, следовательно, варианта исключению не подлежит.

Средняя геометрическая применяется в тех случаях, когда изменения вариант в вариационном ряду происходят в геометрической прогрессии, т. е. каждый последующий уровень ряда, характеризующий развитие явления, примерно равен предыдущему, умноженному на некоторое постоянное для данной прогрессии число, называемое знаменателем прогрессии. Изменение явлений в геометрической прогрессии имеет место в микробиологической и иммунологической практике (размножение микроорганизмов, нарастание титра антител и т. д.). Средняя геометрическая определяется по формуле: $\bar{x}_{геом} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \mathbf{K} \cdot x_n}$, где $\bar{x}_{геом}$ – средняя геометрическая; x_1, x_2, x_n – варианты; n – число членов вариационного ряда.

Медиана **(*Ме*) –** варианта, делящая ранжированный вариационный ряд на две части, одна из которых имеет значения вариант меньше Ме, другая – больше. Применение медианы особенно показано при несимметричном распределении вариант в ряду, а также при небольшом числе наблюдений, когда крайние варианты отличаются от остальных и в значительной степени могут определять уровень средней арифметической.

При небольшой выборке определить медиану довольно легко. Например, в вариационном ранжированном ряду (7, 9, 16, 24, 30, 32, 44), где число вариант нечетное, медианой является срединная варианта 24. В ряду 9, 13, 18, 24, 26, 29, 31, 36, т. е. при четном числе вариант, за медиану принимают среднюю арифметическую из двух центральных вариант, т. е. (24+26):2=25.

При распределении выборки в вариационный ряд медиана определяется по формуле:

$$Me = x_{Me} + i\frac{\frac{n}{2} - p_s}{p_{Me}},$$

где x_{Me} – нижняя граница интервала, в котором находится *Me* (при анализе интервального ряда), или полусумма соседних классовых вариант промежутка, в котором находится *Me* (при анализе дискретного ряда); p_s – число накопленных (кумулированных) частот, стоящее перед медианным интервалом (классом); p_{Me}– частота медианного интервала (класса); 1 – классовый промежуток (интервал); *n* – объем выборки.

Расчет медианы в дискретном и интервальном рядах рассмотрим на примерах (табл.4 и 5).

Пример. Найти медиану ряда распределения 146 штаммов золотистого стафилококка по спектру устойчивости к 9 антибиотикам.

Таблица 4

Исходные данные и параметры расчета Ме для дискретного ряда

Количество антибиотиков, к которым устойчивы стафилококки (x)	Число штаммов (p_i)	Накопленные частоты (p_s)
1	8	8
2	6	14 (8+6)
3	24	38 (14+24)
4	20	58 (38+20)
5	53	111 (58+53) и т.д.
6	12	
7	10	
8	5	
9	8	
	n=146	

1. Находится медианный интервал (класс), для чего объем выборки делится на 2 (146:2=73).

2. Для определения местоположения *Ме* кумулируются частоты ряда p_i (графа 3) до числа накопленных частот, стоящих перед медианным интервалом (p_s). p_s=111. Так как число 73 находится между p_s=58 и p_s=111, то величина нижней границы интервала, в котором находится *Ме*, будет между 4 и 5 антибиотиками:

$$x_{Me} = \frac{4+2}{5} = 4{,}5.$$

3. Частота медианного класса (p_{Me}) будет равна числу штаммов, находящихся в медианном классе. p_{Me}=53

4. Величина классового интервала (i) составляет 1, так как антибиотикограмма изучалась с интервалом в один препарат.

5. Полученные значения подставляются в формулу:

$$Me = 4{,}5 + 1 \cdot \frac{73-58}{53} = 4{,}8.$$

Незначительно отличается и определение *Ме* интервального вариационного ряда.

Пример. Определить наиболее подверженную заболеваемости острыми кишечными инфекциями возрастную группу населения согласно данным, представленным в таблице 5.

Таблица 5

Исходные данные и параметры расчета Ме для интервального ряда

Возраст больных, годы (x)	Число больных (p_i)	Накопленные частоты (p_s)
0-3	7	7
3-6	82	89 (7+82)
6-9	58	147 (89+58)
9-12	32	179 (147+32) и т.д.
12-15	25	
15-18	24	
18-21	19	
21-24	17	
24-27	6	
27-30	5	

n=275

1. Медианный интервал составит:

275:2=137,5 (138).

2. Накопленные частоты, между которыми находится медианный интервал, равны: p_s=89 и p_s=147. При p_s=147 границами интервала, в котором находится Ме, является возрастная группа 6–9 лет (нижняя граница x_{Me}=6) при частоте p_{Me}=58.

3. Величина классового интервала (i) составляет 3 года.

$$\mathrm{Me} = 6 + 3 \cdot \frac{138 - 89}{58} = 8{,}5$$

Мода (Мо) – величина, которая наиболее часто встречается в данной совокупности. Класс (интервал) с наибольшей частотой называется модальным. Мода определяется по формуле:

$$\mathrm{Mo} = x_{\text{н}} + i \cdot \frac{p_2 - p_1}{2p_2 - p_1 - p_3},$$

где $x_{\text{н}}$ – нижняя граница модального класса (интервала); p_2 – модальный класс; p_1 – частота класса, предшествовавшего модальному; p_3 – частота класса, следующего за модальным; i – величина классового интервала.

Определение Мо рассмотрим для данных, приведенных в таблице 5. В модальном классе (p_2=82) нижней границей будет возраст 3 года ($x_{\text{н}}$=3); частота предшествующего класса p_i =7, а следующего за модальным классом $p_з$=58; i=3 года.

$$\mathrm{Mo} = 3 + 3 \cdot \frac{82 - 7}{2 \cdot 82 - 7 - 58} = 5{,}3$$

Следовательно, применительно к условиям примера среди больных острыми кишечными инфекциями наиболее часто встречались дети в возрасте 5,3 года.

Кроме рассмотренных основных числовых характеристик случайной величины, в некоторых случаях вычисляют и другие числовые характеристики, из которых важнейшими являются *асимметрия As* и *эксцесс Ex*.

Для симметричного распределения каждому имеющемуся значению случайной величины слева от М соответствует такое значение случайной величины справа от М, которое дает с ним такую же (но с противоположным знаком) разность и наблюдается такое же число раз. Поэтому сумма разностей *(M–x_i)*, умноженных на P_i, для симметричного распределения равна нулю. Этот результат не меняется, если возвести все разности *(M–x_i)* в любую нечетную степень. Именно поэтому в качестве показателя асимметрии применяется математическое ожидание куба отклонения случайной величины от среднего значения. Этот показатель делят на куб среднего квадратического отклонения, чтобы получить безразмерную величину.

Показатель E_x характеризует крутизну спадания распределения в области его математического ожидания.

Для одного из важнейших законов распределения – закона Гаусса, плотность распределения которого изображается кривой колоколообразной формы (см. рис. 1), имеет место соотношение:

$$\frac{\sum_{i=1}^{i=n} (x_i - M)^4}{\sigma^4} = 3,$$

т. е. для него E_x=0. Все остальные симметричные распределения, таким образом, как бы сравниваются с распределением Гаусса: для более островершинных E_x>0, для более плосковершинных E_x<0.

ОПРЕДЕЛЕНИЕ ОБЪЕМА ВЫБОРКИ ДЛЯ ПОЛУЧЕНИЯ РЕПРЕЗЕНТАТИВНЫХ СРЕДНИХ ВЕЛИЧИН

Для обнаружения закономерностей изучаемых явлений необходим достаточно большой объем исследований, однако он не может возрастать бесконечно, а должен находиться в рациональных границах. Последние зависят от желаемой точности наблюдения (допустимой ошибки выборки), а также от заданного уровня значимости, который не должен быть менее 95%. Необходимый объем исследований рассчитывается по формуле:

$$n = \frac{t^2 \cdot \sigma^2}{\Delta^2},$$

где n – число необходимых исследований (объем выборки); Δ – предельная ошибка выборки; s – среднее квадратическое отклонение; t – нормированное отклонение (2,0 или 3,0).

По указанной формуле возможно определение объема выборки только для расчета средней арифметической, т. е. в случаях, когда рассеяние вариант подчиняется закону нормального распределения.

ОБРАБОТКА ДАННЫХ ВЫБОРОЧНОГО ИССЛЕДОВАНИЯ МЕТОДАМИ ОПИСАТЕЛЬНОЙ СТАТИСТИКИ В СЛУЧАЕ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ ДАННЫХ С ПРИМЕНЕНИЕМ ПРОГРАММНОГО ПАКЕТА STADIA

В пакете *STADIA* довольно полно представлены методы описательной статистики, все они собраны воедино в разделе пакета «Параметрические тесты» блока «Статистики». Проиллюстрируем их работу на примерах.

Пример. Был измерен в мкм диаметр 50 эритроцитов (табл.6).

Таблица 6

Диаметр эритроцитов человека, мкм				
7,49	7,7	7,55	7,64	7,62
7,51	7,53	7,55	7,56	7,58
7,53	7,55	7,58	7,53	7,55
7,55	7,52	7,58	7,55	7,55
7,54	7,57	7,55	7,54	7,54
7,55	7,55	7,55	7,56	7,55
7,55	7,53	7,52	7,55	7,55
7,53	7,55	7,55	7,53	7,54
7,55	7,55	7,51	7,58	7,56
7,55	7,55	7,57	7,55	7,55

Для выборки диаметров эритроцитов вычислим среднее значение, медиану, дисперсию, нижнюю и верхнюю квартили, а также минимальный и максимальный элементы

Подготовка данных. Находясь в редакторе базы данных пакета, следует ввести данные таблицы с клавиатуры. Далее необходимо ввести имя файла для сохранения данных, например «erytr». Результаты ввода частично представлены на рисунке 2.

Файл: erytr **Переменных=1** **Измерений=50**

Var/Cases	1/50	2/0	3/0	4/0	5/0	6/0	7/0
Varname	x1	x2	x3	x4	x5	x6	x7
37	7.55						
38	7.51						
39	7.58						
40	7.55						
41	7.55						
42	7.55						
43	7.51						
44	7.58						
45	7.56						
46	7.55						
47	7.55						
48	7.57						
49	7.55						
50	7.55						

БЛОК РЕДАКТОРА ДАННЫХ

F1 **Помощь** F2 **ПечЭкр** F3 **Чтение** F4 **Запись** F5 **Архив** F6 **Рисунок** F7 **Очист** F8 **Преобр** F9 **Статис** F10 **Выход**

Вводите в матрицу числа + Enter (работают также: Enter/ Ins/ Del/ Tab и F-ключи

Рис.2. Пакет ***STADIA.*** Экран блока редактора данных
с загруженной выборкой

На запрос системы «Выберите метод или нажмите его ключ» следует нажать ключ процедуры «1=Описательная статистика» (рис.3). На следующий запрос «Укажите номер анализируемой переменной (Enter = Все)» надо ввести «1». В данном случае, когда в файле только одна переменная, можно нажать «Enter».

Файл: **Переменных=1** **Измерений=20**

СТАТИСТИЧЕСКИЕ МЕТОДЫ

ПАРАМЕТРИЧЕСКИЕ ТЕСТЫ
1 = Описательная статистика
2 = Гистограмма и нормальность
3 = Корреляция
4 = Тесты Стьюдента и Фишера
НЕПАРАМЕТРИЧЕСКИЕ ТЕСТЫ
5 = Хи-квадрат
6 = Сдвига(положения)
7 = Масштаба (рассеяния)
8 = Произвольных альтернатив
9 = Для парных выборок
A = Корреляция (независимость)
B = Кросстабуляция
АНАЛИЗ ВРЕМЕННЫХ РЯДОВ
C = Корреляционный анализ
D = Спектральный анализ
E = Сглаживание и фильтрация

ДИСПЕРСИОННЫЙ АНАЛИЗ
G = 1-факторный параметрический
H = 2-факторный параметрический
I = 1-факторный Крускала -Уоллиса
J = 2-факторный Фридмана
РЕГРЕССИОННЫЙ АНАЛИЗ
K = Сравнение двух регрессий
L = Простая регрессия (тренд)
M = Множественная линейная регрессия
N = Пошаговая регрессия
O = Общая (+ нелинейная) регрессия
МНОГОМЕРНЫЕ МЕТОДЫ
P = Дискриминантный анализ
Q = Кластерный анализ
R = Факторный анализ
S = Шкалирование

СТАТИСТИЧЕСКИЕ МЕТОДЫ

f10 Выход f1 Информация f2 Экран на печать / в файл Esc Выход с прерыванием

Выберите метод или нажмите его ключ >>

Рис.3. Меню блока статистических методов

Результаты. На экране появятся значения основных описательных статистик и запрос системы «Выдать дополнительную статистику». В ответ на запрос можно нажать «Y», и тогда программа выведет остальные описательные статистики (рис.4).

Файл: **Переменных=1** **Измерений=50**

ОПИСАТЕЛЬНАЯ СТАТИСТИКА **Переменные: x1**

Размер	Среднее	Ош. средн	Довер. инт	Медиана		
50	7.553	4.579E-3	9.082E-3	7.55		
Выбор. дисп	**Ст. откл**	**Ош. ст. от**	**Генер. дисп**	**Ст. откл**	**Доверит. интервал**	
1.049 E-3	3.238 E-2	3.238 E-3	1.028 E-3	3.206 E-2	7.317 E-4	1.629 E-3
Сумма	**Сум. квадр**	**Дисп.сум.кв**	**m2**	**m3**	**m4**	
377.6	2852	5.243 E-2	1.028E-3	7.319 E-5	3.815 E-6	
Ассиммерия	**Значим**	**Эксцесс**	**Значим**			
2.222	0	3.613	0.1768			

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием

Нажмите Enter=продолжить или F2=печать экрана >>

Рис.4. Результат обработки с использованием описательной статистики

Во встроенном справочнике программы имеются определения и сведения о назначениях всех этих описательных статистик. Для вывода данной информации на экран (рис.5) следует нажать клавишу «F1».

РУБРИКАТОР **Стр.1**

Переместите рамку на требуемую рубрику из нижеследующего списка с помощью клавиш со стрелками и нажмите клавишу <Enter>. Для выхода из справки нажмите клавишу <f10>.Сначала желательно ознакомится с порядком работы со справочником в разделе <Работа со справочником> и узнать <Новости>.

Состав системы	**Настройка и копирование системы**
Это вы можете	**Порядок диалога**
Что может статистика	**Ввод и изменение данных**
Какой метод анализа выбрать	**Команда блока редактора данных**
Статистические данные	**Построение графиков данных**
Пропущенные значения	**Преобразование данных**
Статистические методы	**Средства ввода / вывода и файлы**
Диагностика ошибок	

СПРАВОЧНИК ПО СИСТЕМЕ

f1 **Оглавление** f2 **Печать экрана** f3 **Выбор метода анализа данных** f9 **Предыд. раздел** f10 **Выход**

Используйте: PgUp/PgDw=листование; стрелк=выбор раздела; Enter=вход в раздел.

Рис.5. Главное окно справочной системы пакета ***STADIA***

Комментарии. 1.Часть описательных статистик, вычисляемых этой процедурой, относятся только к выборкам из нормального распределения. Это касается размера доверительного интервала для среднего и значений концов доверительного интервала для дисперсии.

2. Если в блок редактора данных загружено несколько переменных и на запрос системы «Укажите номер анализируемой переменной (Enter =

все)» Вы нажали «Enter», то будут вычислены описательные статистики для всех этих переменных.

Возможности графического анализа данных в пакете STADIA

Пример. Сгруппируем данные предыдущего примера в диапазоне от 7,50 мкм до 7,60 мкм с шагом группировки 0,01 мкм и вычислим частоты попадания в полученные интервалы группировки. Подготовка данных осуществляется так же, как в предыдущем примере.

Выбор процедуры. В блоке статистических методов следует нажать клавишу «2», чтобы выбрать процедуру «2=Гистограмма и нормальность».

Заполнение полей ввода данных. На запрос системы «Укажите число интервалов и диапазон гистограммы (Enter=вычисл.)» надо ввести требуемые в примере значения: 7,5, 7,6, 0,01, и затем нажать «Enter».

Результаты. На экране появится результаты расчетов, включающие таблицу табуляции частот, значения статистик Колмогорова и хи-квадрат (χ^2), а также заключение системы «Гипотеза 0: Распределение не отличается от нормального» (рис.6).

Файл: **Переменных=1** **Измерений=50**

ГИСТОГРАММА И ТЕСТ НОРМАЛЬНОСТИ. **Переменные: x1**

x	x-станд	Частота	%	Накопл.	%
7.49	-1.965	4	8	4	8
7.518	-1.092	10	20	14	28
7.546	-0.2183	29	58	43	86
7.574	0.655	4	8	47	94
7.602	1.528	1	2	48	96
7.63	2.402	1	2	49	98
7.658	3.275	0	0	49	98
7.686	4.148				

Колмогоров-Смирнов=0.2964 Значимость=1.48E-10 степ.своб=50
Хи-квадрат=23.18 Значимость=0.4382 степ.своб=5.5

Гипотеза 0 : Распределение не отличается от нормального

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием

Нажмите Enter=продолжить >>

Рис.6. Результаты процедуры «Гистограмма и нормальность»

В первом столбце таблицы указан правый конец интервала группировки, во втором значения первого столбца трансформированы следующим образом: из каждого элемента первого столбца вычитается среднее значение выборки и полученная разность делится на стандартное отклонение выборки. Следующие четыре столбца содержат частоту, относительную частоту, накопленную частоту и относительную накопленную частоту соответственно.

После нажатия «Enter» появится запрос системы запрос системы «Вывести график». При ответе «Y» программа выводит гистограмму и подобранную по выборке кривую плотности нормального распределения. На экран выводится запрос вида оформления графика и выбора устройства вывода (рис.7).

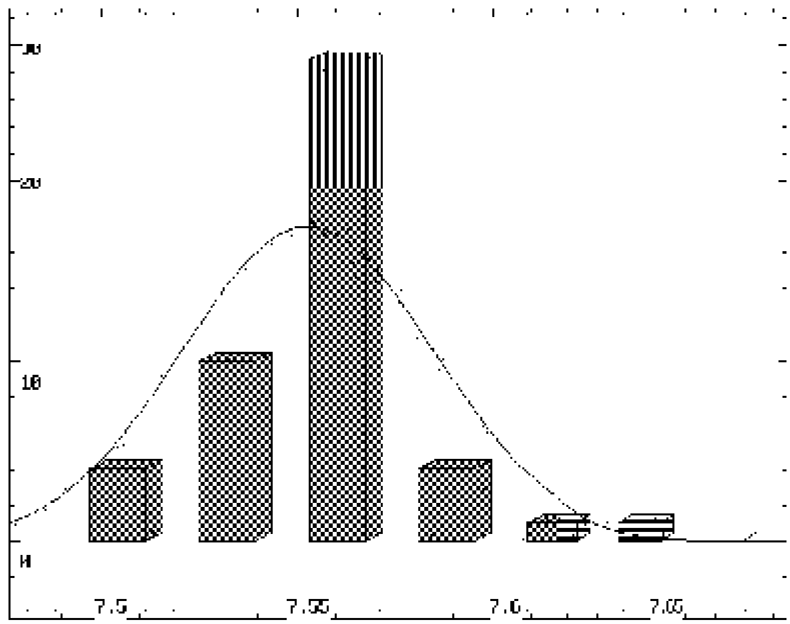


Рис.7. Гистограмма с наложенным графиком нормальной кривой

Прочие возможности. Из других графических методов описательной статистики в пакете ***STADIA*** представлен также матричный график, в котором значения каждой переменной, находящейся в текущий момент в блоке редактора данных, сгруппированы в отдельном столбце графика с указанием положения среднего значения и интервала стандартного отклонения.

Сравнение данных двух выборок с использованием пакета STADIA с применением t-критерия.

В качестве данных для анализа используем данные о числе лейкоцитов (10^9/л) крови у двух групп обследуемых лиц.

Группа 1

7,97		9,47	7,41	14,82	3,61	4,59	4,49	9,98	14,89	5,24

Группа 2

7,40	9,23	7,23	15,41	4,69	6,28	3,70	10,87	14,26	5,54

Подготовим данные для анализа, введя их с клавиатуры в базу данных пакета. Сохраним введенные данные в файле с именем «leu».

Выбор процедуры. В меню статистических методов выберем пункт «4 = Тесты Стьюдента и Фишера», далее следует нажать клавишу «Enter».

Результаты. Результаты анализа представлены на рисунке 8.

Приводимые данные свидетельствуют об отсутствии различий между выборками как по их дисперсиям, так и по данным t-критерия для парных данных.

Файл: Переменных=2 Измерений=20

КРИТЕРИЙ ФИШЕРА И СТЬЮДЕНТА. Переменные: x1 x2

Статистика Фишера=1.067, Значимость=0.4621, степ.своб=9, 9

Статистика Стьюдента=0.1189, Значимость=0.9025, степ.своб=18

Стьюдент для парных данных=0.8134, Значимость=0.5581, степ.своб=9

Гипотеза 0 : Нет различий между выборочными средними

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием
Нажмите Enter=продолжить или F2=печать экрана >>

Рис. 8. Результаты сравнения двух выборок с использованием пакета ***STADIA*** с применением t-критерия.

КРАТКИЙ ОБЗОР ПОНЯТИЯ «КРИТЕРИЙ ЗНАЧИМОСТИ»

Если известно распределение наблюдаемой переменной, то можно предсказать, как в повторных выборках равного объема будет «вести себя» используемая статистика – т.е. каким образом она будет распределена. Пусть, например, имеется 100 случайных выборок, из одной популяции по 100 взрослых человек в каждой. Вычислим средний рост субъектов в каждой выборке, т.е. найдем выборочное среднее.

Тогда распределение выборочных средних можно хорошо аппроксимировать нормальным распределением (более точно, *t* распределением Стьюдента с 99 степенями свободы). Теперь представим, что случайным образом извлечена еще одна выборка из жителей некоего города «N», где, по нашим представлениям, проживают люди с ростом выше среднего. Если средний рост людей в этой выборке попадает в верхнюю 95% критическую область *t* распределения, то можно сделать обоснованный вывод, что жители города «N», действительно, в среднем более высокие, чем в целом в популяции, т.е. это действительно город высоких людей.

ИМЕЮТ ЛИ БОЛЬШИНСТВО ПЕРЕМЕННЫХ НОРМАЛЬНОЕ РАСПРЕДЕЛЕНИЕ?

В рассмотренном примере использовался тот факт, что в повторных выборках равного объемы средние значения (роста людей) будут иметь *t* распределение (с определенным средним и дисперсией). Однако это верно, если рассматриваемая переменная (рост) имеет нормальное распределение, т.е. распределение людей определенного роста нормально распределено.

Для многих изучаемых переменных невозможно сказать с уверенностью, что это действительно так. Случаи редких болезней не являются нормально распределенными в популяции, число автомобильных аварий также не является нормально распределенным, как и многие другие переменные, интересующие исследователя.

ОБЪЕМ ВЫБОРКИ

Другим фактором, часто ограничивающим применимость критериев, основанных на предположении нормальности, является объем или размер выборки, доступной для анализа. До тех пор пока выборка достаточно большая (например, 100 или больше наблюдений), можно считать, что выборочное распределение нормально, даже если вы не уверены, что распределение переменной в популяции, действительно, является нормальным. Тем не менее, если выборка очень мала, то критерии, основанные на нормальности, следует использовать только при наличии уверенности, что переменная действительно имеет нормальное распределение. Однако нет способа проверить это предположение на малой выборке.

ПРОБЛЕМЫ ИЗМЕРЕНИЯ

Использование критериев, основанных на предположении нормальности, кроме того, ограничено точностью измерений. Например, рассмотрим исследование, в котором средний балл успеваемости (СБУ) является основной переменной. Можно ли сказать, что средняя успеваемость студента A в два раза выше, чем успеваемость студента C? Является ли различие между средним баллом студентов B и A сравнимым с различием между студентами D и C? Индекс СБУ является грубой мерой, позволяющей только ранжировать студентов в порядке «хороший»–«плохой». Эта общая задача измерений обычно обсуждается в учебниках по статистике в терминах *типов измерений* или *шкалы измерения*.

Наиболее общие статистические методы, такие как дисперсионный анализ (*t-критерий*), регрессия и т.д. предполагают, что исходные измерения выполнены, по крайней мере, в интервальной шкале, в которой интервалы можно разумным образом сравнивать между собой (например, B минус A равняется D минус C). Тем не менее, как в данном примере, такие предположения часто неестественны, и данные скорее просто упорядочены (измерены в порядковой шкале), чем измерены точно.

ПАРАМЕТРИЧЕСКИЕ И НЕПАРАМЕТРИЧЕСКИЕ МЕТОДЫ

Непараметрические методы разработаны для достаточно часто возникающих на практике ситуаций, когда исследователь ничего не знает о параметрах исследуемой выборки (отсюда и название методов - *непараметрические*). Говоря более специальным языком, непараметрические методы не основываются на оценке параметров (таких как среднее или стандартное отклонение) при описании выборочного распределения интересующей величины. Поэтому эти методы иногда также называются *свободными от параметров* или *свободно распределенными*.

КРАТКИЙ ОБЗОР НЕПАРАМЕТРИЧЕСКИХ ПРОЦЕДУР

По существу, для каждого параметрического критерия имеется, по крайней мере, один непараметрический аналог. Эти критерии можно отнести к одной из следующих групп:

- критерии различия между группами (независимые выборки);
- критерии различия между группами (зависимые выборки);
- критерии зависимости между переменными.

Различия между независимыми группами. Обычно, когда имеются две выборки (например, мужчины и женщины), которые вы хотите сравнить относительно среднего значения некоторой изучаемой переменной, вы используете t-критерий для независимых выборок. Непараметрическими альтернативами этому критерию являются: *критерий серий Вальда-Вольфовица*, *U-критерий Манна-Уитни* и *двухвыборочный критерий Колмогорова-Смирнова*. Если анализируется несколько групп, то возможно использование дисперсионный анализ. Его непараметрическими аналогами являются: ранговый дисперсионный анализ *Краскела-Уоллиса* и *медианный* тест.

Различия между зависимыми группами. При сравнении двух переменных, относящиеся к одной и той же выборке (например, успехи студентов в начале и в конце семестра), то обычно используется *t-критерий для зависимых выборок*. Альтернативными непараметрическими тестами являются: *критерий знаков* и *критерий Вилкоксона парных сравнений*. Если рассматриваемые переменные по природе своей категориальны или являются категоризованными (т.е. представлены в виде частот попавших в определенные категории), то подходящим будет *критерий c(хи)-квадрат*. Если рассматривается более двух переменных, относящихся к одной и той же выборке, то обычно используется дисперсионный анализ (*ANOVA*) с повторными измерениями. Альтернативным непараметрическим методом является *ранговый дисперсионный анализ Фридмана* или *Q-критерий Кохрена* (последний применяется, например, если переменная измерена в номинальной шкале). Q-критерий Кохрена используется также для оценки изменений частот (долей).

Зависимости между переменными. Для того, чтобы оценить зависимость (связь) между двумя переменными, обычно вычисляют коэффициент корреляции. Непараметрическими аналогами стандартного коэффициента корреляции Пирсона являются статистики *Спирмена R*, *t(тау) Кендалла* и *коэффициент Гамма*. Если две рассматриваемые переменные по природе своей категориальны, подходящими непараметрическими критериями для тестирования зависимости будут: *хи-квадрат (c^2)*, *Фи-коэффициент*, *точный критерий Фишера (F)*. Дополнительно доступен критерий зависимости между несколькими переменными так называемый *коэффициент конкордации Кендалла*. Последний часто используется для оценки согласованности мнений независимых экспертов.

ОПИСАТЕЛЬНЫЕ СТАТИСТИКИ

Если данные не являются нормально распределенными, а измерения, в лучшем случае, содержат ранжированную информацию, то вычисление обычных описательных статистик (например, среднего, стандартного отклонения) не слишком информативно. Например, в психометрии хорошо известно, что воспринимаемая интенсивность стимулов (например, воспринимаемая яркость света) представляет собой логарифмическую функцию реальной интенсивности (яркости, измеренной в объективных единицах - *люксах*). В данном примере обычная оценка среднего (сумма значений, деленная на число стимулов) не дает верного представления о среднем значении действительной интенсивности стимула (в обсуждаемом примере скорее следует вычислить среднее геометрческое). При использовании методов *непараметрической статистики* вычисляется разнообразный набор мер положения (среднее, медиана мода и т.д.) и рассеяния (дисперсия, гармоническое среднее, квантильный размах и т.д.), позволяющий представить более «полную картину» данных.

НЕПАРАМЕТРИЧЕСКАЯ СТАТИСТИКА

Зачем нужны непараметрические тесты? Ниже перечислены несколько причин, не позволяющих в определенных случаях использовать тесты, основанные на нормальном распределении:

1. Не все переменные имеют нормальное распределение

Существует множество тестов, которые основываются на предположении, что анализируемые переменные нормально распределены. Но бывают случаи, когда встречаются переменные, распределение которых не нормальное. Например, число пожаров не является нормально распределенной величиной. Существует множество переменных, которые не имеют нормального распределения, но представляют большое интерес для ученых.

2. Маленький размер выборки

Если объем выборки достаточно большой (например, 100 наблюдений), то вы можете использовать тесты основанные на предположении нормальности, даже если вы не уверены, что исследуемая переменная нормально распределена в популяции. Если же размер выборки невелик, то прежде чем использовать «нормальные» тесты, необходимо убедится, что переменные распределены нормально, поскольку в обратном случае результаты полученные с помощью теста будут некорректными. Но проблема в том, что как раз маленький объем выборки и не позволяет убедиться в том, что переменная в действительности имеет нормальное распределение.

3. Недостаточно информативная шкала измерения

Основные тесты, основанные на предположении о нормальности, требуют, чтобы анализируемы переменные были измерены как минимум в интервальной шкале. Но как быть в случаях, когда переменная измерена в порядковой шкале и имеет значения типа: «высокий», «средний», «низкий» (например, при измерении уровня владения иностранным языком)?

Исходя из этих причин, вытекает потребность в тестах, с помощью которых можно было бы проанализировать переменные, обладающие всеми теми недостатками, что были перечислены выше. Термин «непараметрические» происходит оттого, что мы практически не знаем ни каких параметров исследуемой популяции. При описании выборочного распределения, непараметрические тесты не основываются на таких параметрах как среднее или стандартное отклонение.

КАКОЙ МЕТОД ИСПОЛЬЗОВАТЬ?

Каждая непараметрическая процедура в модуле имеет свои достоинства и свои недостатки. Например, двухвыборочный *критерий Колмогорова-Смирнова* чувствителен не только к различию в положении двух распределений, например, к различиям средних, но также чувствителен и к форме распределения. *Критерий Вилкоксона* парных сравнений предполагает, что можно ранжировать различия между сравниваемыми наблюдениями. Если это не так, лучше использовать *критерий знаков*. В общем, если результат исследования является важным (например, оказывает ли людям помощь очень дорогостоящее и болезненное лечение?), то всегда целесообразно применить различные непараметрические тесты. Возможно, результаты проверки (разными тестами) будут различны. В таком случае следует попытаться понять, почему разные тесты дали разные результаты. С другой стороны, непараметрические тесты имеют меньшую статистическую мощность (менее чувствительны), чем параметрические критерии, и если важно обнаружить даже слабые отклонения (например, является ли данная пищевая добавка опасной для людей), следует особенно внимательно выбирать статистику критерия.

Большие массивы данных и непараметрические методы. Непараметрические методы наиболее приемлемы, когда объем выборок мал. Если данных много (например, n>100), то не имеет смысла использовать непараметрические статистики. Главное состоит в том, что когда выборки становятся очень большими, то выборочные средние подчиняются нормальному закону, даже если исходная переменная не является нормальной или измерена с погрешностью. Таким образом, параметрические методы, являющиеся более чувствительными (имеют большую статистическую мощность), всегда подходят для больших выборок. Большинство критериев значимости многих непараметрических статистик, описанных далее, основываются на асимптотической теории (больших выборок), поэтому соответствующие тесты часто не выполняются, если размер выборки становится слишком малым.

Подгонка распределения. В некоторых исследовательских проектах можно сформулировать гипотезы относительно распределения рассматриваемой переменной. Например, переменные, значения которых определяются бесконечным числом независимых факторов, распределены по нормальному закону: можно предположить, что рост индивидуума является результатом воздействия многих независимых факторов, таких как различные генетические предрасположенности, болезни, перенесенные в раннем возрасте и т.д. Как следствие, рост имеет тенденцию к нормальному распределению в населении. С другой стороны, если наблюдаемые значения переменной являются результатом очень редких событий, то переменная будет иметь распределение *Пуассона* (которое иногда называется распределением редких событий). Например, несчастные случаи на производстве можно рассматривать как результат пересечения ряда неудачных событий (на житейском языке стечением маловероятных обстоятельств), поэтому их частота приближенно описывается распределением Пуассона.

КРИТЕРИЙ СООТВЕТСТВИЯ ПИРСОНА (ХИ-КВАДРАТ)

Критерий соответствия (согласия) хи-квадрат (χ^2) предложен К. Пирсоном в 1900 г. для проверки предположения о наличии (отсутствии) связи между явлениями. Хи-квадрат – критерий, устанавливающий соответствие между теоретическими и эмпирическими частотами распределения. Его вычисление проводится по формуле:

$$\chi^2 = \sum \frac{(x - y)^2}{y},$$

где x – реальный ряд; y – теоретически ожидаемый ряд.

Таким образом, математически хи-квадрат представляет сумму частных от деления квадратов отклонений фактически полученных данных от «ожидаемых» на число «ожидаемых». Значения хи-квадрат могут возрастать от 0 до ∞. Если $x = y$, т. е. «ожидаемые» числа соответствуют фактическим, то $\chi^2 = 0$. Такое положение подтверждает правильность нулевой гипотезы и свидетельствует об отсутствии различий между сравниваемыми явлениями.

При исчислении критерия хи-квадрат используются только абсолютные показатели, причем группа должна состоять не менее чем из 5 наблюдений. В противном случае надлежит изменить группировку исследуемого материала, уменьшив число групп и получив большие числа первоначальных наблюдений.

Пример. Имеются ли различия в характере эпидемических процессов при дизентерии и прочих острых кишечных инфекциях (ОКИ), если помесячное число больных было, как представлено в таблице 7.

1. Суммируется помесячное число больных обеими формами заболеваний (графа 4).

2. Определяется удельный вес помесячного числа больных (графа 5).

3. В таблицу 8 заносятся данные реального динамического ряда, т. е. зарегистрированное число больных (графа 3).

4. Рассчитывается «ожидаемое» число больных (теоретический динамический ряд) на основании данных таблицы.

Для января: 109 больных дизентерией (всего) составляют 100 %,
x больных в январе – 11,2 %.
$x = 12,2$

Для февраля: 109 больных дизентерией (всего) составляют 100 %,
x больных в феврале – 3,8%
$x = 4,1$

Расчет данных для прочих острых кишечных инфекций в январе:

940 больных ОКИ (всего) составляют 100 %,

x больных ОКИ в январе – 11,2%

x = 105,3

Таблица 7

Месяцы	Число больных		Всего	Процент
	Дизентерия	Прочие ОКИ		
1	***2***	***3***	***4***	***5***
Январь	8	110	118	11,2
Февраль	10	30	40	3,8
Март	7	90	97	9,2
Апрель	10	40	50	4,8
Май	8	20	28	2,7
Июнь	8	20	28	2,7
Июль	12	120	132	12,7
Август	10	100	110	10,5
Сентябрь	8	120	128	12,2
Октябрь	15	100	115	10,9
Ноябрь	7	140	147	14,0
Декабрь	6	50	56	5,3
Итого:	109	940	1049	100,0

Полученные данные заносятся в графу 4 таблицы 8.

5. Определяется разность (x–y) между реальным и теоретически ожидаемым числом больных (графа 5). Поскольку полученные отклонения в последующем возводятся в квадрат, знаки разности не учитываются.

6. Разность (x–y) возводится в квадрат (графа 6).

7. Квадрат отклонений теоретических данных $(x–y)^2$ делится на «ожидаемые» числа (y). Полученные данные суммируются, и эта сумма представляет собой величину χ^2 (графа 7). В нашем примере χ^2 = 44,9.

Чем больше разность между x и y, тем больше и значение χ^2. Оценка достоверности величины χ^2 проводится **по приложению 5** с учетом числа степеней свободы (n'), представляющих собой число клеток таблицы, данные в которых могут быть свободно изменены без существенных изменений конечных результатов:

$$n = (s-1) \cdot (r-1),$$

где s – число граф в таблице 7; r – число строк в таблице 7.

В нашем примере полученная величина χ^2 выше критического уровня 19,68 при $P<0,05$ и n = 11 **в приложении 5.** Следовательно, есть основания отвергнуть нулевую гипотезу и говорить о существенности различий, а стало быть, и независимости характера эпидемического процесса при дизентерии и прочих острых кишечных инфекциях.

Таблица 8

Нозологические формы	Месяцы	Зарегистрировано больных (реальный динамический ряд) (x)	«Ожидаемое» число больных (теоретический динамический ряд) (y)	(x-y)	$(x\text{-}y)^2$	$\frac{(x-y)^2}{y}$
1	**2**	**3**	**4**	**5**	**6**	**7**
Дизентерия	I	8	12,2	4,2	17,6	1,4
	II	10	4,1	5,9	34,8	8,5
	III	7	10,0	3,0	9,0	0,9
	IV	10	5,2	4,8	23,0	4,4
	V	8	2,9	5,1	26,0	8,9
	VI	8	2,9	5,1	26,0	8,9
	VII	12	13,8	1,8	3,2	0,2
	VIII	10	11,4	1,4	1,9	0,2
	IX	8	13,3	5,3	28,1	2,1
	X	15	11,9	3,1	9,6	0,8
	XI	7	15,3	8,3	68,9	4,5
	XII	6	5,8	0,2	0,04	0
Прочие острые кишечные инфекции (ОКИ)	**I**	110	105,3	4,7	22,1	0,2
	II	30	35,7	5,7	32,5	0,9
	III	90	86,4	3,6	12,9	0,1
	IV	40	45,1	5,1	26,0	0,6
	V	20	25,4	5,4	29,2	1,1
	VI	20	25,4	5,4	29,2	1,1
	VII	120	119,4	0,6	0,4	0
	VIII	100	98,7	1,3	1,7	0
	IX	120	114,7	5,3	28,1	0,3
	X	100	102,5	2,5	6,3	0
	XI	140	131,6	8,4	70,6	0,6
	XII	50	49,8	0,2	0,04	0
						χ^2=44,9

Критерий хи-квадрат можно применять и для оценки влияния различных факторов на те или иные процессы и явления. В этих случаях исходной нулевой гипотезой, которая должна быть или отвергнута после определения величины χ^2, или, наоборот, сохранена, является отсутствие влияния тех или иных факторов.

КРИТЕРИЙ ВИЛКОКСОНА ДЛЯ СВЯЗАННЫХ СОВОКУПНОСТЕЙ

Критерий Вилкоксона для связанных совокупностей – это непараметрический метод, который используется для оценки значимости различий двух связанных совокупностей количественных признаков.

Практический расчет критерия включает следующие этапы:

1. Найти разности парных вариант.

2. Определить ранги полученных разностей (без учета знаков, пары наблюдений, разности которых оказались равными нулю, из дальнейшей оценки исключаются).

3.Определить сумму рангов полученных разностей, имеющих одинаковые алгебраические знаки, и взять меньшую из них (T).

4.Установить достоверность различий.

При количестве наблюдений меньше 26 сравнивают найденную сумму с критическими значениями из таблицы, в противном случае рассчитывают по специальной формуле случайную переменную (u).

Критерий Вилкоксона является более мощным, чем критерий знаков и максимум-критерий. Его следует использовать при наличии у сравниваемых совокупностей количественных признаков значительного числа разностей с противоположенными знаками.

КРИТЕРИЙ ЗНАКОВ

Критерий знаков - это непараметрический критерий, который основан на оценке разности попарно сопряженных вариант (например, до и после лечения). Учитывается не величина, а направленность сдвигов. Применение критерия знаков не зависит от характера распределения данных. Изменения оценивают в альтернативной форме (увеличение-уменьшение и т.п., что обозначают знаками "+" и "-", откуда и произошло название критерия). Случаи, когда парные наблюдения не имеют разницы, в расчет не принимаются. Следует стремиться, чтобы количество нулевых разностей было минимальным. Для этого необходимо повышать точность измерения показателей, что обеспечивает непрерывность выборочных данных.

Практическое применение критерия знаков включает следующие этапы:

1) Определяется направленность изменений в сравниваемых наблюдениях.

2) Подсчитывается общее число парных наблюдений, имеющих различия (n).

3) Подсчитывается меньшее число однозначных результатов сравнения, обозначаемых как Z**.**

4) **Z** сравнивается по специальной таблице с критическими значениями для данного n.

Мощность критерия знаков ограничена и составляет примерно 2/3 мощности критерия Стьюдента.

Критерий знаков приложим как к совокупностям непрерывных признаков, так и для оценки различия полуколичественных признаков (баллы и т.п.) при достаточном числе их градаций.

КРИТЕРИЙ МАННА – УИТНИ

Область применения критерия Манна-Уитни – анализ двух независимых выборок. Размеры этих выборок могут различаться.

Назначение критерия – проверка гипотезы о статистической однородности двух выборок. Иногда эту гипотезу называют гипотезой об отсутствии эффекта обработки (имея в виду, что одна из выборок содержит характеристики объектов, подвергшихся некоему воздействию а другая – характеристики контрольных объектов).

Данные. Рассматриваются две выборки объемов m и n. Закон распределения первой выборки F, а второй – G.

Допущения. 1. Выборки должны быть независимы. 2. Законы распределений F и G непрерывны.

Гипотеза. Утверждение об однородности выборок в введенных выше обозначениях можно записать в виде H : F = G.

Альтернативы. В качестве альтернатив к H могут выступать все возможности $F \neq G$. Однако критерий Манна-Уитни способен обнаруживать отнюдь не все возможные отступления от H : F = G. Этот критерий предназначен, в первую очередь, для проверки H против альтернативы $F \leq G$ (правосторонняя альтернатива) или альтернативы $F \geq G$ (левосторонняя альтернатива). Можно рассматривать и объединение обеих возможностей (двусторонняя альтернатива).

Метод. Критерий Манна-Уитни повторяет основные идеи критерия знаков и в определенном смысле является его продолжением. Он основан на попарном сравнении результатов из первой и второй выборок.

ТЕСТ КОЛМОГОРОВА – СМИРНОВА

Этот тест решает две задачи. В случае, если анализируется одна выборка, то этот тест определяет к какой плотности вероятности относится данная выборка. То есть функция плотности, полученная для данной выборки, сравнивается с одной или несколькими теоретическими функциями плотности, которые могут быть нормальными, и (или) экспоненциальными, и (или) Коши, и (или) равномерными. В случае, когда анализируется две выборки, то функция плотности одной выборки сравнивается с функцией плотности другой выборки. Из указанного выше сравнения выводится статистика. В случае одной выборки эта статистика оценивает вероятность того, что статистика будет такой же, как полученная величина, или больше нее, если верна гипотеза, что действительная (выборочная) и теоретическая функции распределения совпадают. Другими словами, если вероятность получена, например, равной 0,40, то отвержение гипотезы о равенстве функций распределения будет ошибочным в 40 случаях из 100. При двух выборках проверяется гипотеза о равенстве двух действительных (выборочных) функций распределений.

КРИТЕРИЙ КОЛМОГОРОВА – СМИРНОВА ДЛЯ ОДНОЙ ВЫБОРКИ

Критерий Колмогорова-Смирнова является другим непараметрическим критерием. Его преимущество перед критерием χ^2 связано с тем, что он принимает во внимание порядок наблюдений. Критерий предназначен для проверки согласия эмпирической и теоретической функций распределений.

Критерий Колмогорова-Смирнова основывается на статистической модели, которая предполагает непрерывность распределения, так что вероятность совпадения выборочных значений равна нулю. Однако на практике критерий часто применяется к сгруппированным данным и данным выборок из дискретных распределений. В обоих этих случаях учитывается возможность появления равных значений наблюдений. В данном случае уровень значимости критерия ниже номинального, и вероятность ошибки второго рода возрастает.

Как критерий χ^2, так и критерий Колмогорова-Смирнова предполагают, что распределение, фигурирующее в нулевой гипотезе, должно быть полностью определено заранее (например, оно может быть нормальным с нулевым средним и единичной дисперсией). При работе с критерием согласия χ^2 необходимо заранее определить тип распределения, а параметры оцениваются по выборочным данным. Критерий χ^2 легко модифицируется при помощи уменьшения числа степеней свободы, но неизвестно, какие изменения должны быть внесены в процедуру применения критерия Колмогорова-Смирнова. Несмотря на это он иногда применяется в качестве критерия при проверке гипотез о законе распределения. Следует учитывать, что в этом случае истинный уровень значимости будет несколько ниже номинального и возрастет вероятность ошибок второго рода. По всей видимости, этот эффект не будет велик, если число оцениваемых параметров мало по сравнению с объемом выборки.

Данные: выборка x_1, x_2, x_3, …, x_n из n независимых наблюдений, упорядоченных так, что $x_1 < x_2 < x_3 < \ldots < x_n$.

Статистическая модель: наблюдения независимы и берутся из генеральной совокупности, распределение которой предполагается непрерывным.

Гипотезы: нулевая гипотеза – функция распределения равна $F(x)$.

Вычисление критериальной статистики:

1. Вычисляются кумулятивные разности: $D_1=1-nF(x_1)$, $D_2=1-nF(x_2)$ и т.д.
2. Находим $|D_i|_{max}$, наибольшее абсолютное значение кумулятивных разностей.
3. Вычисляем значение критерия $D=(|D_i|_{max})/n$

Пример. Даны 20 наблюдений (табл.9) из неизвестной генеральной совокупности.

Таблица 9

0,33	-0,52	-2,41	-1,93	0,46
-0,44	-0,97	-0,38	0,48	1,29
-1,82	-1,23	-0,21	2,66	-1,22
-0,41	-0,95	1,47	-0,83	-0,43

Необходимо проверить нулевую гипотезу, предполагающую, что это нормальное распределение с нулевым средним значением и единичной дисперсией.

Для этого надо первоначально упорядочить наблюдения по возрастающей, а затем – вычислить наблюдаемые и ожидаемые значения, которые меньше каждого наблюдения или равны ему. Необходимо также вычислить абсолютное значение кумулятивной разницы во всех наблюдениях.

Пример. Зафиксированное число наблюдений, которые не больше 0,97 равно 6. Ожидаемое значение равно: 20F (-0,97) = 3,32. Абсолютное значение кумулятивной разности при x=-0,97 равно 6-3,32=2,68.

По данным составленной таблицы 10 находим максимальное значение $|D_i|$ – 5,96. Значение критериальной статистики равно:

$$D = \frac{5{,}96}{20} = 0{,}298$$

При помощи таблиц, содержащих сведения о табличных значениях распределения, известно, что критическое значение p=0,05 равно 0,294. Таким образом, необходимо отклонить нулевую гипотезу о том, что данное распределение является нормальным с нулевым средним и единичной дисперсией.

Таблица 10

Наблюдение x_i	Число наблюдений, меньших или равных x_i		Разность $\lvert D_i\rvert$
	Наблюдаемое	Ожидаемое	
-2,41	1	0,16	0,84
-1,93	2	0,54	1,46
-1,82	3	0,69	2,31
-1,23	4	2,19	1,81
-1,22	5	2,22	2,78
-0,97	6	3,32	2,68
-0,95	7	3,42	3,58
-0,83	8	4,07	3,93
-0,52	9	6,03	2,97
-0,44	10	6,60	3,40
-0,43	11	6,67	4,33
-0,41	12	6,82	5,18
-0,38	13	7,04	5,96
-0,21	14	8,34	5,66
0,33	15	12,59	2,41
0,46	16	13,54	2,46
0,48	17	13,69	3,31
1,29	18	18,03	3,03
1,47	19	18,58	0,42
2,66	20	19,92	0,08

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ В ПАКЕТЕ STADIA С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ НЕПАРАМЕТРИЧЕСКОЙ СТАТИСТИКИ

В качестве примера используем данные о содержании холестерина (мМоль/л) в сыворотке крови здоровых мужчин различных возрастных групп: 1 группа – лица до 19 лет, вторая группа – лица старше 30 лет (табл.11).

Таблица 11

Группа 1										
5,50	5,53	5,51	5,46	5,53	5,86	5,73	5,55	5,49	5,48	5,44
5,49	5,71	4,88	5,74	5,44	5,48	5,41	5,77	5,79	5,53	
Группа 2										
5,84	6,20	5,93	5,99	5,97	6,31	6,22	6,24	6,33	6,11	6,12
6,18	6,05	5,88	5,97	5,94	6,13	6,21	6,18	6,22	6,24	

Приводимый ниже рисунок 9 иллюстрирует введенные данные в окне редактора данных пакета *STADIA.*

Файл:			Переменных=2			Измерений=42	
Var/Cases	1/21	2/21	3/0	4/0	5/0	6/0	7/0
Varname	x1	x2	x3	x4	x5	x6	x7
1	5.5	5.84					
2	5.53	6.2					
3	5.51	5.93					
4	5.46	5.99					
5	5.55	5.97					
6	5.86	6.31					
7	5.75	6.77					
8	5.55	6.24					
9	5.49	6.39					
10	5.48	6.11					
11	5.44	6.17					
12	5.49	6.18					
13	5.71	6.05					
14	4.88	5.88					
15	5.74	5.97					
16	5.44	5.94					
17	5.48	6.13					
18	5.41	6.21					
БЛОК РЕДАКТОРА ДАННЫХ							
F1 Помощь F2 ПечЭкр F3 Чтение F4 Запись F5 Архив F6 Рисунок F7 Очист F8 Преобр F9 Статис F0 Выход							
Вводите в матрицу числа + Enter (работают также: Enter/ Ins/ Del/ Tab и F-ключи							

Рис.9. Введенные данные по содержанию холестерина в сыворотке крови у двух групп мужчин

Аналогично анализу выборки с использованием методов описательной статистики для выхода в меню основных методов анализа нажмем клавишу F9 и выберем пункт «5 = хи-квадрат». При выборе пункта «0 эмпирич. (однородность)» на экране появляются результаты анализа (рис.10), свидетельствующие об отсутствии различий в распределении выборочных величин.

При выборе пункта «1 = теоретич. (согласия)» производится анализ согласия выборочного и предполагаемого статистического распределения. Результаты анализа приводятся на рисунке 11.

Файл:	Переменных=2	Измерений=42
КРИТЕРИЙ ХИ-КВАДРАТ.		**Переменные: x1 x2**
Хи-квадрат=6.739 Е-2	**Значимость=1**	**степ.своб=20**
Гипотеза 0: Нет различий между двумя распределениями		
	СТАТИСТИЧЕСКИЕ МЕТОДЫ	
F10 Выход F1 Информация	**F2 Экран на печать / в файл**	**Esc Выход с прерыванием**
Нажмите Enter=продолжить или F2=печать экрана >>		

Рис.10. Результаты анализа различий в распределении выборочных величин при выборе пункта «0 эмпирич. (однородность)»

Файл:	Переменных=2	Измерений=42
КРИТЕРИЙ ХИ-КВАДРАТ.		**Переменные: x1 x2**
Хи-квадрат=1.227	**Значимость=0.9999**	**степ.своб=20**
Гипотеза 0: Нет различий между двумя распределениями		
	СТАТИСТИЧЕСКИЕ МЕТОДЫ	
F10 Выход F1 Информация	**F2 Экран на печать / в файл**	**Esc Выход с прерыванием**
Нажмите Enter=продолжить или F2=печать экрана >>		

Рис.11. Результаты анализа различий в распределении выборочных величин при выборе пункта «1 = теоретич. (согласия)»

Далее проанализируем используемые данные с применением критерия Вилкоксона. Для этого выберем пункт «6 = сдвига (положения)» раздела непараметрических тестов. Результаты анализа приведены на рисунке 12.

Файл:	Переменных=2	Измерений=42
КРИТЕРИЙ СДВИГА (ПОЛОЖЕНИЯ)	**Переменные: x1 x2**	
Вилкоксон=232 Z=5.522	**Значимость=0, степ.своб=21, 21**	
Гипотеза 1: Есть различия между медианами выборок		
	СТАТИСТИЧЕСКИЕ МЕТОДЫ	
F10 Выход F1 Информация	**F2 Экран на печать / в файл**	**Esc Выход с прерыванием**
Нажмите Enter=продолжить >>		

Рис.12. Результаты анализа с применением критерия Вилкоксона

По данным анализа с применением критерия Вилкоксона мы должны отвергнуть гипотезу о равенстве медиан выборок.

Для анализа данных примера с помощью критерия Колмогорова-Смирнова для двух выборок необходимо в главном меню методов анализа избрать пункт «8 = произвольных альтернатив». На рисунке 13 приведены результаты применения данного критерия к анализируемым выборкам.

Файл: **Переменных=2** **Измерений=42**

КРИТЕРИЙ КОЛМОГОРОВА-СМИРНОВА **Переменные: x1 x2**

Колмогоров-Смирнов=0.9524, Значимость=1.068E-8, степ.своб=21, 21

Гипотеза 1: Есть интегральные различия между выборками

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием
Нажмите Enter=продолжить или F2=печать на экран >>

Рис.13. Результаты анализа с помощью критерия Колмогорова-Смирнова

Результаты анализа свидетельствуют о наличии интегральных различий между двумя анализируемыми выборками.

Применение непараметрических тестов для парных выборок – пункт меню «9 = для парных выборок» позволяет проанализировать выборочные значения для выборок, полученных при парном выборочном исследовании с использованием как критерия Вилкоксона, так и критерия знаков. Результаты анализа приведены на рисунке 14.

Файл: **Переменных=2** **Измерений=42**

ПАРНЫЕ КРИТЕРИИ СДВИГА **Переменные: x1 x2**

Вилкоксон=0 Z=-4.015, Значимость=0, степ.своб=21
Знаков=0 Z=-4.583, Значимость=0, степ.своб=21

Гипотеза 1: Есть различия между медианами выборок

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F10 Выход F1 Информация F2 Экран на печать/в файл Esc Выход с прерыванием

Нажмите Enter=продолжить или F2=печать на экран

Рис.14. Результаты анализа при парном выборочном исследовании с использованием критерия Вилкоксона и критерия знаков.

Данные критерии также подтверждают наличие различий между медианами выборок.

В заключение данного раздела необходимо повторить приводившееся ранее предупреждение о применении к одним данным нескольких методов анализа. Если к одним и тем же данным применяют два различных критерия для проверки одной и той же нулевой гипотезы (или двух сходных гипотез) и в каждом случае принимается уровень значимости, равный, например, 5%, то вероятность того, что хотя бы по одному из критериев нулевая гипотеза будет ошибочно отклонена, превосходит 5 %. Следует воспользоваться лишь одним критерием, желательно более мощным.

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ

Между явлениями существуют две категории причинно-следственных связей: *функциональные* и *корреляционные*.

Функциональная связь имеет отношение к каждому отдельному наблюдению, является обязательной и зависит от физических и химических законов. Например, при повышении давления насыщенного пара увеличивается его температура; нагревание металлов приводит к их расширению и т. п.

Связь между явлениями, отмечающаяся не в каждом отдельном случае, а при массовом сопоставлении рассматриваемых признаков, называется *корреляционной*. В этом случае какой-либо признак (уровень заболеваемости острыми кишечными инфекциями) может быть связан главным образом (но не обязательно) с изменениями другого признака (качество реализуемых молочных продуктов), хотя на интенсивность течения эпидемического процесса при острых кишечных инфекциях влияют и другие факторы: микробная обсемененность потребляемой населением воды, кулинарных изделий, миграционные процессы и др. Корреляционная связь проявляется в среднем для всей совокупности наблюдений, а в отношении единичных случаев она неполна и неточна.

Корреляционная зависимость между явлением может быть прямой, когда с повышением одного показателя отмечается рост другого, и обратной, когда с увеличением одного признака другой уменьшается (например, чем больше населения охвачено прививками, тем ниже уровень заболеваемости).

Прежде чем измерять величину коэффициента корреляции, необходимо решить вопрос о возможности причинно-следственной связи между изучаемыми явлениями. Параллельное изменение статистических показателей само по себе еще не говорит о наличии связи, так как может быть обусловлено случайным совпадением многих обстоятельств, не связанных друг с другом.

Цифровые данные, подвергающиеся корреляционному анализу, должны быть сгруппированы с учетом особенностей изучаемых явлений. В противном случае значение полученного коэффициента будет заведомо ошибочным. Это особенно касается области инфекционной патологии, когда от момента инфицирования, т. е. действия причины, до клинического проявления процесса, т. е. следствия, проходит инкубационный период, длительность которого различна при отдельных заболеваниях. Продолжительность инкубационного периода и является оптимальным сроком временной группировки цифровых данных. При неинфекционной патологии в основу группировки кладут другие временные признаки, интересующие исследователя. Например, при изучении частоты развития инфаркта миокарда у лиц одного и того же возраста в зависимости от длительности курения материал можно группировать с учетом интересующей давности курения: 1, 2, 3 и т. д. лет.

Метод ранговой корреляции наиболее простой, требующий незначительных затрат времени на проведение вычислений. В то же время он менее точен, дает приблизительное представление о характере и тесноте связи между явлениями.

Пример. Определить возможную связь между уровнем заболеваемости острыми кишечными инфекциями (ОКИ) и качеством реализуемых кулинарных изделий, определяемого по микробной обсемененности (табл.12).

Таблица 12

Неделя года	Заболеваемость ОКИ	Процент нестандартных проб кулинарных изделий по микробной обсемененности	Порядковые номера (ранги)		$\overline{d}$	$\overline{d^2}$	$\overline{d^2}$ с нарастающим итогом
			Заболеваемости	% нестандартной продукции			
1	***2***	***3***	***4***	***5***	***6***	***7***	***8***
1	12,8	8,3	6	2	4	16	16
2	11,6	5,1	2	1	1	1	17
3	15,4	18,0	9	10	1	1	18
4	17,8	14,0	11	7	4	16	34
5	12,3	11,6	4	4,5	0,5	0,25	34,25
6	12,5	14,2	5	8	3	9	43,25
7	14,0	24,4	7	11	4	16	59,25
8	20,3	25,4	12	12	0	0	59,25
9	15,6	11,6	10	4,5	5,5	30,25	89,50
10	14,8	12,9	8	6	2	4	93,50
11	11,6	14,6	3	9	6	36	139,50
12	9,3	10,0	1	3	2	4	133,50

1. Определяются порядковые номера (ранги) показателей заболеваемости (графа 4) и процента нестандартной продукции (графа 5). Ранг 1 присваивается наименьшим показателям. При наличии нескольких, равных по величине показателей (11,6 в графе 3) их порядковые номера (ранги) суммируются (4+5=9), а сумма делится на число одинаковых показателей (в нашем примере на 2). Оба показателя имеют один и тот же ранг (4,5), а следующему по величине показателю (12,9) присваивается 6-й ранг.

2. Вычисляется разность (d) между рангами в отдельные недели года. Результаты заносятся в графу 6.

3. Разность между рангами возводится в квадрат, и полученные данные суммируются (графа 7). В нашем примере $\sum d^2 = 133{,}5$.

4. Полученные значения подставляются в формулу:

$$\rho = 1 - \frac{6 \cdot \sum d^2}{n(n^2 - 1)},$$

где ρ – коэффициент ранговой корреляции; *n* – число пар коррелируемых рядов (в нашем примере 12); $\sum d^2$ – сумма разности между рангами двух коррелируемых рядов; 6 – постоянный коэффициент. Для нашего примера: ρ=0,53.

Прежде чем делать заключение о наличии корреляционной связи, необходимо убедиться в репрезентативности коэффициента, так как он получен на основе выборочного исследования. Представительность коэффициента ранговой корреляции определяется путем сопоставления его величины с критическими значениями, приведенными в приложении 3. Если полученная величина при заданном числе исследований окажется больше критического уровня, будут основания говорить о наличии корреляционной связи. В нашем примере коэффициент равен 0,53, а критическая величина при $P<0,05$ составляет 0,58 для n=12. Следовательно, полученное значение коэффициента ниже того, при котором можно было бы говорить о наличии корреляционной связи.

Метод линейной корреляции используется при наличии прямолинейной связи между взаимосвязанными признаками. Последовательность расчета коэффициента линейной корреляции (r_{xy}) рассмотрим на предыдущем примере. Для этого заносим данные граф 1, 2, 3 таблицы 12 в таблицу 13.

Таблица 13

Неделя года	Заболеваемость ОКИ, (x)	Процент нестандартных проб кулинарных изделий (y)	Отклонение от средней арифметической		Квадрат отклонения от средней арифметической		$\overline{d_x^2} \cdot \overline{d_y^2}$
			$\overline{d_x}$	$\overline{d_y}$	$\overline{d_x^2}$	$\overline{d_y^2}$	
1	***2***	***3***	***4***	***5***	***6***	***7***	***8***
1	12,8	8,3	-1,2	-5,9	1,44	34,81	7,08
2	11,6	5,1	-2,4	-9,1	5,76	82,81	21,84
3	15,4	18,0	1,4	3,8	1,96	14,4	5,32
4	17,8	14,0	3,8	-0,2	14,44	0,04	-0,76
5	12,3	11,6	-1,7	-2,6	2,89	6,76	4,42
6	12,5	14,2	-1,5	0	2,25	0	0
7	14,0	24,4	0	10,2	0	104,04	0
8	20,3	25,4	6,3	11,2	39,69	125,44	70,56
9	15,6	11,6	1,6	-2,6	2,56	6,76	-4,16
10	14,8	12,9	0,8	-1,3	0,64	1,69	-1,04
11	11,6	14,6	-2,4	0,4	5,76	0,16	-0,96
12	9,3	10,0	-4,7	-4,2	22,09	17,64	19,74

1. Суммируются данные ряда x (графа 2), и сумма делится на число наблюдений.

2. Суммируются данные ряда y (графа 3), и сумма делится на число наблюдений.

3. Определяется разность (d) каждого показателя ряда x от средней арифметической ($\bar{x}$=14,0) и ряда y от своей средней арифметической ($\bar{y}$=14,2). Полученные значения заносятся соответственно в графы 4 и 5. Результаты будут иметь как положительные, так и отрицательные значения.

4. Полученные величины $\overline{d_x}$ и $\overline{d_y}$ возводятся в квадрат, а результаты заносятся в графы 6 и 7, после чего суммируются данные.

5. Перемножаются показатели $\overline{d_x}$ и $\overline{d_y}$ с учетом положительных и отрицательных значений (графа 8). Полученные произведения суммируются.

6. Цифровые данные подставляются в формулу:

$$r_{xy} = \pm \frac{\sum(\overline{d_x} \cdot \overline{d_y})}{\sqrt{\sum \overline{d_x^2} \cdot \sum \overline{d_y^2}}}$$, где r_{xy} – коэффициент линейной корреляции.

Для нашего примера: $r_{xy} = \pm 0{,}616$

Поскольку полученный коэффициент определен в результате выборочного исследования, необходимо убедиться в степени его надежности. Представительность r_{xy} определяется **по приложению 4**. Коэффициент считается представительным, если полученная величина превышает критическое значение при P<0,05 и заданном числе степеней свободы (n'=n-2). В нашем примере полученная величина 0,616 превышает критический уровень 0,576 при n'= 10 и P<0,05.

При отсутствии таблицы критических значений коэффициента корреляции его представительность может быть определена по величине средней ошибки $(S_{r_{xy}})$:

$$S_{r_{xy}} = \pm \sqrt{\frac{1 - r_{xy}^2}{n}}$$, при числе парных наблюдений больше 100 или

$$S_{r_{xy}} = \pm \sqrt{\frac{1 - r_{xy}^2}{n - 2}}$$, при числе парных наблюдении меньше 100.

Отношение величины коэффициента корреляции к величине средней ошибки позволяет найти значение t:

$$t = \frac{r_{xy} \cdot \sqrt{n}}{1 - r_{xy}^2}$$ при n>100, или $$t = r_{xy} \frac{\sqrt{n-2}}{1 - r_{xy}^2}$$ при n<100.

В нашем примере: t=2,471. По таблице значений t-критерия при P<0,05 и числе степеней свободы 10 t =2,23.

После установления представительности полученного значения коэффициента определяется теснота связи, выражающаяся величиной от 0 до 1 (табл.14).

Таблица 14

Степень связи	Величина коэффициента
Малая (слабая)	0,01 – 0,30
Средняя (умеренная)	0,31 – 0,70
Большая (сильная)	0,71 – 1,00

Степень «связанности» в вариации двух или нескольких изучаемых явлений более точно измеряется квадратом коэффициента корреляции (r_{xy}). Это значит, что при r_{xy} = 0,5 25% изменений одного признака объясняется вариацией другого. При r_{xy} = 0,3 изменчивость одного явления закономерно связана с изменением другого в 10%, при r_{xy} = 0,7 – в 50%, при r_{xy} = 0,9 –в 81% случаев. Таким образом, хотя коэффициент корреляции и указывает на общность элементов в коррелируемых рядах, но не вся эта общность объясняется закономерной связью в вариации признаков. Из сказанного ясно, что о тесной корреляции можно говорить лишь в случаях, когда r_{xy} не ниже 0,7.

Правильная трактовка коэффициента корреляции предполагает нормальное распределение сопряженных величин коррелируемых рядов x и y. Однако при малом числе наблюдений и сравнительно сильной корреляции (r_{xy}=0,5) распределение коэффициента r_{xy} отличается значительной асимметрией от нормального распределения (рис.15).

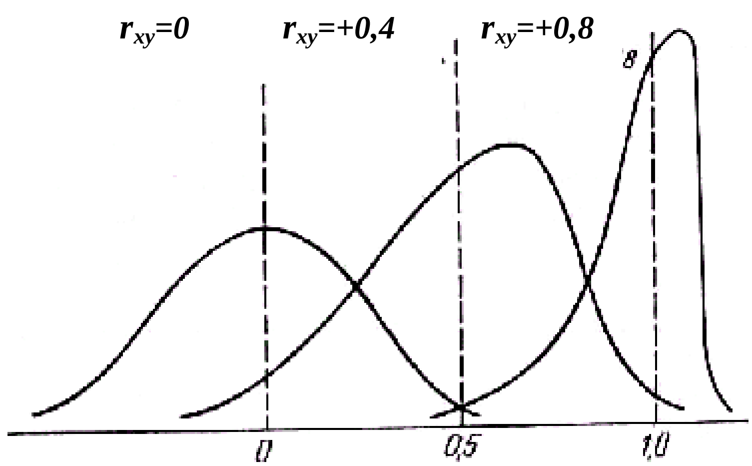


Рис.15. Кривые распределения эмпирического коэффициента корреляции при n=12 для различных значений генерального параметра r_{xy} (по А.К. Митропольскому, 1971)

Следовательно, эмпирический коэффициент корреляции не будет точной оценкой для всей генеральной совокупности, если он определен на малочисленной выборке и его величина больше 0,5. В этих случаях коэффициент r_{xy} целесообразно заменить преобразованной величиной z (z — преобразование Фишера), связанной с коэффициентом корреляции следующим образом:

$$z = 1{,}15129 \lg \frac{1 + r_{xy}}{1 - r_{xy}}.$$

Распределение величины z является почти неизменным по форме, так как она мало зависит от численности выборки и значения r_{xy} в генеральной совокупности. Преобразование r_{xy} в z проводится по **приложению 4**, в котором содержатся величины z, соответствующие значениям r_{xy}. Достоверность z определяется по формуле:

$$t_z = z \cdot \sqrt{n - 3}$$

Полученная величина t_z сравнивается с t-критерием Стьюдента при n'=n–2 и P<0,05. В нашем примере: r_{xy} = 0,616 соответствует z = 0,72; t_z = 2,16. t-критерий Стьюдента при P<0,05 и n=10 составляет 2,23.

Значение числа z заключается еще в том, что только с его помощью можно определить достоверность различий между двумя коэффициентами корреляции, а также объединить данные по нескольким корреляциям.

Пример. Между помесячными данными числа родов и числа заболевших гнойно-воспалительными инфекциями новорожденных в двух акушерских стационарах получены следующие коэффициенты: r_{xy} = 0,525, r_{xy}= 0,750. Число парных коррелируемых величин составляло по 1-му акушерскому стационару 60, по 2-му – 72. Имеются ли различия между полученными коэффициентами?

По приложению 4 переводим значения r_{xy} в z.

$$z_1=0{,}59;\ z_2=0{,}97;\ d_z=z_1-z_2$$

Средняя ошибка для разности между z_1 и z_2 определяется по формуле:

$$S_{d_z} = \pm\sqrt{s_{z_1}^2 + s_{z_2}^2}$$

Поскольку $S_z = \pm\sqrt{\dfrac{1}{n-3}}$, то $S_{z_1} = \pm 0{,}13$; $S_{z_2} = \pm 0{,}12$; $S_{d_z} = \pm 0{,}18$.

Отсюда $t = \dfrac{d}{S_{d_z}} = \dfrac{0{,}38}{0{,}18} = 2{,}11.$

По таблице значений t-критерия при P<0,05 и n'=(n_1–2)+(n_2–2)=128 t=1,96. Поскольку полученная величина 2,11 выше критического значения, разность между коэффициентами признается существенной.

ОПРЕДЕЛЕНИЕ ОБЪЕМА ВЫБОРКИ ДЛЯ ПОЛУЧЕНИЯ РЕПРЕЗЕНТАТИВНОГО КОЭФФИЦИЕНТА КОРРЕЛЯЦИИ

Необходимый объем выборочных исследований для заданной точности коэффициента корреляции рассчитывается по формуле:

$$n = \frac{t^2}{z^2} + 3,$$

где n – искомый объем выборки; t – нормированное отклонение; z – преобразованная величина эмпирического (основанного на опыте) коэффициента корреляции.

Пример. При корреляционном анализе двух рядов из 12 парных наблюдений получен r_{xy}=0,35. Коэффициент непредставителен, так как критическая величина при P<0,05 и n'= 12 – 2 = 10 составляет 0,58. Сколько же необходимо наблюдений, чтобы при такой зависимости между признаками получить репрезентативный коэффициент корреляции?

1. Величина r_{xy} переводится в z, т. е. r_{xy} = 0,35 соответствует z = 0,365.

$$t_Z = 0{,}365 \cdot \sqrt{12-3} = 1{,}095$$

2. По таблице значений t-критерия при P<0,05 и n' = 10 t = 2,23. Поскольку полученная величина t_z = 1,095 ниже 2,23, следовательно, r_{xy} = 0,35 непредставителен и необходимо увеличить объем выборки.

3. Определяется необходимое число выборочных исследований при P<0,05 и t = 1,96 (число наблюдений больше 30). Тогда:

$$n = \frac{1{,}96^2}{0{,}365^2} + 3 = 32$$

Таким образом, число парных наблюдений (n) нужно довести по крайней мере до 32, чтобы с вероятностью 95 % можно было считать выборочный коэффициент репрезентативным.

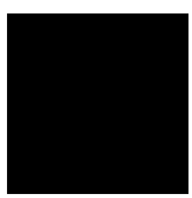
РЕГРЕССИОННЫЙ АНАЛИЗ

Регрессионный анализ, в отличие от корреляционного, указывающего лишь на степень связи в вариации двух или нескольких переменных величин, дает возможность судить о том, как количественно меняется один показатель по мере изменения другого на единицу.

Последовательность расчетов коэффициента регрессии (R) рассмотрим на данных, приведенных в таблице 15.

В графе 3 процент нестандартных проб кулинарных изделий заменен количеством микроорганизмов в 1 грамме продукции.

Поскольку изменчивых величин две (x, y) и регрессия является двусторонней, соответственно будут и два коэффициента: R_{xy}, R_{yx}, которые вычисляются по формулам:



Таким образом, для определения значения коэффициента регрессии необходимо знать:

1)σ ряда x;

2)σ ряда y;

3) величину коэффициента корреляции (r_{xy}).

Таблица 15

Неделя года	Заболеваемость ОКИ, (x)	Среднее количество бактерий в 1 г кулинарных изделий (·100), y	Отклонение от средней арифметической		Квадрат отклонения от средней арифметической		$\overline{d_x^2} \cdot \overline{d_y^2}$
			$\overline{d_x}$	$\overline{d_y}$	$\overline{d_x^2}$	$\overline{d_y^2}$	
1	12,8	17	1,2	3,8	1,44	14,44	4,56
2	11,6	7	2,4	13,8	5,76	190,44	33,12
3	15,4	20	-1,4	0,8	1,96	0,64	-1,12
4	17,8	25	-3,8	-4,2	14,44	17,64	15,96
5	12,3	15	1,7	5,8	2,89	33,64	9,86
6	12,5	18	1,5	2,8	2,25	7,84	4,20
7	14,0	25	0	-4,2	0	17,64	0
8	20,3	54	-6,3	-34,8	39,69	1211,04	219,24
9	15,6	26	-1,6	-5,2	2,56	27,02	8,32
10	14,8	20	-0,8	0,8	0,64	0,64	-0,64
11	11,6	12	2,4	8,8	5,76	77,44	21,12
12	9,3	10	4,7	10,8	22,09	116,64	50,76

Однако величины R могут быть вычислены и в случае отсутствия готовых значений σ и r_{xy}.

[redacted]; [redacted]

Для нашего примера: $R_{xy}=0,21$; $R_{yx}=3,57$.

Поскольку коэффициент регрессии, как и другие статистические показатели, получается в результате выборочных исследований, необходимо убедиться в репрезентативности выборки. С этой целью определяется ошибка выборки для R:

[redacted]

Для нашего примера $S_{R_{xy}} = \pm 0,08$, $S_{R_{yx}} = \pm 1,30$.

Степень представительности устанавливается по t-критерию при n’=n–2 и уровне значимости 0,05.

$$t_{xy} = \frac{R_{xy}}{S_{R_{xy}}}; t_{yx} = \frac{R_{yx}}{S_{R_{yx}}}.$$

Для нашего примера: t_{xy}=2,63; t_{yx}=2,82.

Поскольку полученные значения t превышают критический уровень 2,23 при P<0,05 и n'=10, выборка признается репрезентативной. Применительно к описанной ситуации это дает основание говорить о том, что увеличение микробной обсемененности кулинарных изделий на 1·100 бактерий в 1 г продукта способствует подъему заболеваемости на 0,21 $^{0}/_{00}$. В свою очередь, увеличение заболеваемости на 1 $^{0}/_{00}$ происходит при увеличении числа микроорганизмов в 1 г кулинарных изделий на 3,67·100.

Коэффициент регрессии характеризует только линейную зависимость и имеет знак «+» при положительной или «–» отрицательной связи.

Регрессионный анализ находит широкое применение в связи с тем, что дает возможность оценить количественное изменение одного показателя по мере изменения количественной характеристики другого, в то время как коэффициент корреляции служит общим мерилом сопряженной вариации признаков. В приведенном выше примере процент проб кулинарных изделий, не соответствующих ГОСТу, является более обобщенной величиной, чем массивность микробной обсемененности. Например, число бактерий более 1000 в 1 г продукта уже не соответствует ГОСТу, однако это может быть как 1500, так и 15000 микроорганизмов. Другими словами, несоответствие несоответствию рознь. И если в первом случае пищевой продукт может и не привести к подъему заболеваемости, то во втором случае такая вероятность очень высока.

Между коэффициентом корреляции и коэффициентом регрессии имеется определенная связь, выражающаяся формулой:

$$r_{xy} = \pm\sqrt{R_{xy} \cdot R_{yx}}.$$

Следовательно, зная коэффициенты регрессии, легко определить коэффициент корреляции. Для нашего примера: r_{xy}=±0,88.

РЕГРЕССИВНЫЙ АНАЛИЗ С ИСПОЛЬЗОВАНИЕМ ПАКЕТА STADIA

В пакете широко представлены различные методы регрессионного анализа, включая простую, множественную, пошаговую, нелинейную регрессию и др. Следует сразу обратить внимание на не совсем традиционную классификацию регрессионных моделей в пакете.

Для общего обозначения моделей данных, обрабатываемых методами регрессионного анализа, в справочнике пакета используется термин «Экспериментальные зависимости». Последние делятся в пакете на однопараметрические и многопараметрические, линейные и нелинейные по параметрам. При этом под однопараметрической зависимостью понимается произвольная функция y = f(x), где x – простая действительная переменная. Это определение может привести к путанице, так как число параметров в подобной зависимости может быть любое. В частности, все полиномиальные модели при этом попадают в процедуру «Простой регрессии». Скорее, эти зависимости следовало бы назвать одномерными или однофакторными.

Методом наименьших квадратов вычислим оценки параметров в модели простой линейной регрессии для данных калибровочного эксперимента. Построим 95% доверительную трубку для среднего значения отклика.

Подготовка данных. Введем в редакторе данных пакета данные таблицы 16.

Таблица 16

Var 1	0	0,2	0,4	0,6	0,8	1
Var 2	29,38	110,3	188,8	268,9	348,5	426,4

Выбор процедуры. При нажатии клавиши F9 пакет выведет на экран меню статистических методов. Нажав «L», выберем в нем пункт «L = Простая регрессия (тренд)».

Заполнение полей ввода данных. На запрос пакета «Укажите номер одной или номера двух перем.» укажем номера независимой и зависимой переменных, находящихся в текущий момент в блоке редактора данных. (Задание номера только одной переменной пакет интерпретирует как ввод зависимой переменной, а в качестве независимой переменной использует значения натурального ряда. Подобное представление часто используется при анализе временных рядов.) Далее программа выдает меню зависимостей, отнесенных в пакете к однопараметрическим (рис.16).

Файл:		Переменных=2	Измерений=12
ПРОСТАЯ РЕГРЕССИЯ		Переменные: x1 x2	
Однопараметрическая регрессия			
0 = линейная	Y=a+b*x	C = гипербола	Y=a+b/x
1 = парабола	Y=a+b*x+c*x^2	D = -“-	Y=1/(a+b/x)
2 = полином	Y=сум{ai*x^i}	E = -“-	Y=1/(a+b/x)
3 = степени . 5	Y=a+b*SQR(x)	F = -“-	Y=1/(a+b*SQR(x))
4 = логарифмич	Y=a+b*LN(x)	G = -“-	Y=1/(a+b*LN(x))
5 = степенная	Y=a*x^b	H = -“-	Y=a+1/(b+c*x)
6 = -“-	Y=a+b*x^c	I =оптимума	Y=1/(a+b*x+c*x^2)
7 = экспонента	Y=e^(a+b*x)	J = -“-	Y=X/(a+b*x+c*x^2)
8 = -“-	Y=e^(a+b/x)	K = логистич.	Y=a+b/(1+e^(c+d*x))
9 = -“-	Y=e^(a+b*SQR(x))	L = синусоид.	Y=a+b*x+c*SIN(d+e*x)
A = -“-	Y=e^(a+b*x+c*x^2)	M = задаваемая формулой	
B = -“-	Y=a+b*e^(c*x)	N = робастое сглаживание Хубера	
СТАТИСТИЧЕСКИЕ МЕТОДЫ			
F10 Выход	F1 Информация	F2 Экран на печать / в файл	Esc Выход с прерыванием
Выберите модель или нажмите ее ключ >>			

Рис.16. Меню зависимостей, отнесенных к однопараметрическим

Для выбора простой линейной регрессии надо нажать клавишу «0».

Результаты. Экран вывода результатов этой процедуры содержит четыре блока информации.

В первом из них представлены оценки коэффициентов модели, их стандартные ошибки и уровни значимости t-отношений для проверки гипотез об отличии соответствующих коэффициентов от нуля.

Второй блок информации содержит базовую таблицу дисперсионного анализа, показывающую, как общая вариация отклика распределяется между вариацией, обусловленной введенной моделью, и вариацией остатков.

Третий блок информации содержит абсолютную величину коэффициента множественной корреляции R, коэффициент детерминации R^2, несмещенную оценку коэффициента детерминации R^2прив, а также F-отношение и его уровень значимости для проверки гипотезы о соответствии выбранной модели наблюденным данным. Сравнивая полученный уровень значимости с пятипроцентным, система делает заключение об адекватности модели. Обратим внимание, что при выдаче информации используется округление до четвертой значащей цифры, поэтому близкие к единице значения указанных коэффициентов представлены равными единице.

Четвертый блок информации включает введенные пользователем с клавиатуры значения независимой переменной и соответствующие им значения прогноза, вычисляемые с помощью подобранной модели.

Так, введя значение X = 0,25 на запрос системы: Введите X-значение для предсказания Y, будет получено соответствующее значение Y.

Далее процедура предлагает построить график экспериментальных точек и регрессионной кривой.

Дополнительные возможности. Затем пользователю предлагается меню дополнительных возможностей процедуры. Кроме значений экспериментальных данных они содержат подобранные значения модели, остатки и их стандартизированные значения, а также стандартные ошибки остатков и доверительные интервалы для них (в виде допустимого отклонения для 95% уровня доверия). Процедура также позволяет вывести график остатков и сохранить остатки в отдельной переменной базы данных пакета.

Комментарии. 1. Количества наблюдений в зависимой и независимой переменных должны быть одинаковыми.

2. Большинство функций из списка функциональных зависимостей, обрабатываемых процедурой «L = Простая регрессия (тренд)», являются нелинейными относительно входящих в них параметров. В таких случаях для решения задачи регрессии возможны два подхода. Наиболее общий из них сводится к применению нелинейного метода наименьших квадратов для нахождения оценок неизвестных параметров в модели с аддитивной ошибкой.

Другой, частный метод, основан на преобразовании векторов зависимой и независимой переменных таким образом, чтобы преобразованная функциональная зависимость была линейной относительно параметров. Например, для функции $y=1/(a + b/\sqrt{x})$ преобразование вектора y вида $u_i=1/y_i$ переводит ее в линейную относительно параметров функцию. Аналогичные преобразования допустимы для большинства функций, указанных в списке. Для этих функций процедура «Простая регрессия» сначала осуществляет необходимые преобразования векторов независимой и зависимой переменных *x* и *y*, а затем применяет к преобразованной модели стандартный метод наименьших квадратов для нахождения оценок параметров. Но в связи с этим следует помнить, что требования аддитивности, одинаковой распределенности и нормальности случайной ошибки относятся к преобразованной модели, а не к первоначальной. Более того, за исключением нескольких частных случаев, сформулировать статистические требования к характеру случайной ошибки в исходной модели крайне трудно. Поэтому последующий анализ остатков для первоначальной модели не имеет смысла и необходимо исследовать остатки преобразованной модели. Мы специально обращаем на это внимание, так как процедура «Простая регрессия» пакета выводит в дальнейшем на график и предусматривает возможность сохранения остатков только для первоначальной модели.

3. Еще одной особенностью приведенного выше списка функций является возможность задания пользователем довольно широкого набора алгебраических функций. Как отмечалось, подбор конкретного вида функциональной зависимости – наиболее трудная и творческая часть задачи регрессии. На этой стадии, кроме представлений о физической сущности взаимосвязи, весьма полезно представлять поведение различных функций на различных участках их области определения. Достоинством встроенного справочника пакета *STADIA* является краткая классификация различных функциональных зависимостей, представленных в процедуре простой регрессии, с точки зрения скорости изменения (поведения производных), максимумов, асимптот, периодичности и т.п. При этом пакет позволяет быстро построить графики производных, используя процедуры пункта «Сглаживание и фильтрация».

ДИСПЕРСИОННЫЙ АНАЛИЗ

Изменение (вариация) случайных величин вызывается одновременным действием целого ряда причин (факторов). Например, различные параметры организма (температура, артериальное давление) изменяются в зависимости от времени суток, при назначении какого-либо фармакологического препарата в различных дозах и т. д. Влияние тех или иных факторов на изучаемый признак невозможно выделить в чистом виде. Несмотря на то что при проведении опытов стараются сохранить условия максимально однородными, все же различные опыты дают несколько неодинаковые результаты. Объясняется это тем, что на них влияют многочисленные случайные обстоятельства, многие другие неконтролируемые факторы. Поэтому возникает важная задача разложения общей изменчивости признака на составные части, с одной стороны, определяемыми конкретными факторами, а с другой – вызываемыми случайными, неконтролируемыми причинами.

Раздел статистики, изучающий влияние факторов на изменчивость случайной величины, называется *дисперсионным анализом***.** Задача дисперсионного анализа – выделить те факторы и их сочетания, которые оказывают влияние на изменение случайной величины. В зависимости от количества учитываемых факторов различают однофакторный и многофакторный дисперсионный анализ.

Для того чтобы выводы при дисперсионном анализе были обоснованы, необходимо выполнение следующих условий:

1. Изучаемые факторы должны быть независимыми.
2. Распределение выборочных данных должно соответствовать номальному распределению или сводиться к нему путем соответствующих преобразований.

Основной *целью дисперсионного анализа* является исследование значимости различия между средними. Если вы просто сравниваете средние в двух выборках, дисперсионный анализ даст тот же результат, что и обычный ***t***-критерий для независимых выборок (если сравниваются две независимые группы объектов или наблюдений) или ***t***-критерий для зависимых выборок (если сравниваются две переменные на одном и том же множестве объектов или наблюдений).

ПРОВЕРКА ЗНАЧИМОСТИ

Проверка значимости в дисперсионном анализе основана на сравнении компоненты дисперсии, обусловленной межгрупповым разбросом (называемой *средним квадратом эффекта* или *MS эффект*), и компоненты дисперсии, обусловленной внутригрупповым разбросом (называемой *средним квадратом ошибки* или *MS ошибка*; эти термины были впервые использованы в работе Edgeworth, 1885). Если верна нулевая гипотеза (равенство средних в двух популяциях), то можно ожидать сравнительно небольшое различие выборочных средних из-за чисто случайной изменчивости. Поэтому, при нулевой гипотезе, внутригрупповая дисперсия будет практически совпадать с общей дисперсией, подсчитанной без учета групповой принадлежности. Полученные внутригрупповые дисперсии можно сравнить с помощью F-критерия, проверяющего, действительно ли отношение дисперсий значимо больше 1.

ОСНОВНАЯ ЛОГИКА ДИСПЕРСИОННОГО АНАЛИЗА

Целью дисперсионного анализа является проверка статистической значимости различия между средними (для групп или переменных). Эта проверка проводится с помощью разбиения суммы квадратов на компоненты, т.е. с помощью разбиения общей дисперсии (вариации) на части, одна из которых обусловлена случайной ошибкой (то есть внутригрупповой изменчивостью), а вторая связана с различием средних значений. Последняя компонента дисперсии затем используется для анализа статистической значимости различия между средними значениями. Если это различие *значимо*, нулевая гипотеза *отвергается* и принимается альтернативная гипотеза о существовании различия между средними.

МНОГОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ

Переменные, значения которых определяются с помощью измерений в ходе эксперимента (например, балл, набранный при тестировании), называются *зависимыми* переменными. Переменные, которыми можно управлять при проведении эксперимента (например, методы обучения или другие критерии, позволяющие разделить наблюдения на группы или классифицировать), называются *факторами* или *независимыми* переменными.

Ситуации, когда некоторое явление полностью описывается одной переменной, чрезвычайно редки. Например, если мы пытаемся научиться выращивать большие помидоры, следует рассматривать факторы, связанные с генетической структурой растений, типом почвы, освещенностью, температурой и т.д. Таким образом, при проведении типичного эксперимента приходится иметь дело с большим количеством факторов. Основная причина, по которой использование дисперсионного анализа предпочтительнее повторного сравнения двух выборок при разных уровнях факторов с помощью серий *t*-критерия, заключается в том, что дисперсионный анализ существенно более *эффективен* и, для малых выборок, более информативен.

ГЛАВНЫЕ ЭФФЕКТЫ. ПОПАРНЫЕ (ДВУХФАКТОРНЫЕ) ЭФФЕКТЫ ВЗАИМОДЕЙСТВИЯ

Существует еще одно преимущество дисперсионного анализа перед обычным *t*-критерием: дисперсионный анализ позволяет обнаружить эффекты *взаимодействия* между факторами и поэтому позволяет проверять более сложные гипотезы. Рассмотрим еще один пример, иллюстрирующий только что сказанное. (Термин *взаимодействие* впервые был использован Фишером в работе Fisher, 1926)

Предположим, что имеется две группы студентов, причем психологически студенты первой группы настроены на выполнение поставленных задач и более целеустремленны, чем студенты второй группы, состоящей из более ленивых студентов. Разобьем каждую группу случайным образом пополам и предложим одной половине в каждой группе сложное задание, а другой – легкое. После этого измерим, насколько напряженно студенты работают над этими заданиями. Средние значения для этого (вымышленного) исследования показаны в таблице 17.

Таблица 17

Степень трудности задания	Целеустремленные	Ленивые
Сложное	10	5
Легкое	5	10

Какой вывод можно сделать из этих результатов? Можно ли заключить, что: (1) над сложным заданием студенты трудятся более напряженно; (2) честолюбивые студенты работают упорнее, чем ленивые? Ни одно из этих утверждений не отражает сущность систематического характера средних, приведенных в таблице. Анализируя результаты, правильнее было бы сказать, что над сложными заданиями работают упорнее только честолюбивые студенты, в то время как над легкими заданиями только ленивые работают упорнее. Другими словами, характер студентов и сложность задания *взаимодействуя* между собой влияют на затрачиваемое усилие. Это является примером *попарного взаимодействия* между характером студентов и сложностью задания. Заметим, что утверждения 1 и 2 описывают *главные эффекты*.

ОБЩИЙ СПОСОБ ОПИСАНИЯ ВЗАИМОДЕЙСТВИЙ

В общем случае взаимодействие между факторами описывается в виде изменения одного эффекта под воздействием другого. В рассмотренном выше примере двухфакторное взаимодействие можно описать как изменение главного эффекта фактора, характеризующего сложность задачи, под воздействием фактора, описывающего характер студента. Для взаимодействия трех факторов из предыдущего параграфа можно сказать, что на взаимодействие двух факторов (сложности задачи и характера студента) оказывает влияние, например, и половая принадлежность испытуемых. Если изучается взаимодействие четырех факторов, можно сказать, что взаимодействие трех факторов, изменяется под воздействием четвертого фактора, т.е. существуют различные типы взаимодействий на разных уровнях четвертого фактора. Оказалось, что во многих областях взаимодействие пяти или даже большего количества факторов не является чем-то необычным.

Если, например, измеряются показатели состояния здоровья студентов в начале семестра и в конце семестра, то это и есть повторные измерения. Изучение критерия значимости в таких планах это логическое развитие одномерного случая. Заметим, что методы многомерного дисперсионного анализа обычно также используются для исследования значимости *одномерных* факторов повторных измерений, имеющих более чем два уровня. Соответствующие применения будут рассмотрены позднее в этой части.

Суммы значений переменных и дисперсионный анализ. Даже опытные пользователи одномерного и многомерного дисперсионного анализа часто приходят в затруднение, получая разные результаты при применении многомерного дисперсионного анализа, например, для трех переменных, и при применении одномерного дисперсионного анализа к сумме этих трех переменных, как к одной переменной. Идея *суммирования* переменных состоит в том, что каждая переменная содержит в себе некоторую истинную переменную, которая и исследуется, а также случайную ошибку измерения. Поэтому при усреднении значений переменных, ошибка измерения будет ближе к 0 для всех измерений и усредненное значений будет более надежным. На самом деле, в этом случае применение дисперсионного анализа к сумме переменных разумно и является мощным методом. Однако, если зависимые переменные по своей природе многомерны, то суммирование неуместно.

Например, пусть зависимые переменные состоят из четырех показателей *успеха в обществе*. Каждый показатель характеризует совершенно независимую сторону человеческой деятельности (например, профессиональный успех, преуспевание в бизнесе, семейное благополучие и т.д.). Сложение этих переменных подобно сложению яблока и апельсина. Сумма этих переменных не будет подходящим одномерным показателем. Поэтому с такими данными нужно обходится как с многомерными показателями в *многомерном дисперсионном анализе*.

ОДНОФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ В ПАКЕТЕ *STADIA*

В пакете представлены следующие методы однофакторного анализа: непараметрические критерии Крускала-Уоллиса, а также методы дисперсионного анализа. Обращение к ним осуществляется из раздела «Дисперсионный анализ» блока «Статистики».

Пример. Довольно часто в физиологии труда оценивается влияние тех или иных факторов на производительность труда, например, влияние поощрения за выполненную работу. В качестве примера рассмотрим влияние финансового поощрения на производительность труда. С этой целью шести однородным группам исследуемых были предложены задачи одинаковой сложности. Все задачи предоставлялись каждому члену группы независимо от остальных респондентов. Группы отличались между собой величиной денежного вознаграждения за число правильно решенных задач.

Проверим гипотезу об отсутствии эффектов обработки с помощью критерия Крускала-Уоллиса для данных о влиянии данного фактора стимулирования на производительность труда.

Подготовка данных. В редакторе базы данных пакета введем данные первого столбца таблицы в переменную x1, второго – в переменную x2 и так далее, как это показано в рисунке 17.

Файл: stimul			Переменных=6			Измерений=30	
Var/Cases	1/5	2/5	3/5	4/5	5/5	6/5	7/0
Varname	x1	x2	x3	x4	x5	x6	x7
1	10	8	12	12	24	19	
2	11	10	17	15	16	18	
3	9	16	14	16	22	27	
4	13	13	9	16	18	25	
5	7	12	16	19	20	24	
БЛОК РЕДАКТОРА ДАННЫХ							
F1 Помощь F2 ПечЭкр F3 Чтение F4 Запись F5 Архив F6 Рисунок F7 Очист F8 Преобр F9 Статис F0 Выход							
Вводите в матрицу числа + Enter (работают также: Enter/ Ins/ Del/ Tab и F-ключи							

Рис.17. Ввод полученных результатов в «Блок редактора данных»

Замечание. Процедуры однофакторного анализа пакета STADIA требуют, чтобы данные, отвечающие различным способам обработки (уровням фактора), находились в отдельных переменных. При этом в файле данных недопустимо наличие посторонних переменных. Отсюда следует, что если мы хотим провести факторный анализ только для части способов обработки, либо или объединить несколько способов обработки в один, следует завести новый файл данных и осуществить в нем требуемые преобразования.

Выбор процедуры. После нажатия клавиши «F9» программа выведет на экран меню статистических методов. Нажав клавишу «1», мы выберем пункт этого меню «1-факторный Крускала-Уоллеса».

Результаты. На экране появятся значения статистики Крускала-Уоллиса (рис.18), минимального уровня значимости и числа степеней свободы распределения хи-квадрат, которое используется в качестве асимптотического приближения распределения статистики Крускала-Уоллиса.

Файл:	**Переменных=6**	**Измерений=30**
1-ФАКТОРНЫЙ ДИСПЕРС. АНАЛИЗ КРУСКАЛА-УОЛЛИСА.		
H=21.08, Значимость=7E-4, степ.своб=5		
Гипотеза 1: Есть влияние фактора на отклик		
СТАТИСТИЧЕСКИЕ МЕТОДЫ		
F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием		
Нажмите Enter=продолжить или F2=печать экрана >>		

Рис.18. Значения статистики Крускала-Уоллиса

Сравнение минимального уровня значимости статистики с фиксированным уровнем значимости 0,05 позволяет системе сделать заключение «Есть влияние фактора на отклик».

Комментарии. 1. Процедуры непараметрического однофакторного анализа в пакете допускают также ввод в виде таблицы рангов данных (при ранжировании по всей совокупности). То есть можно было использовать для ввода и данные таблицы.

2. В пакете ***STADIA*** отсутствует процедура оценки эффектов обработки непараметрическими методами.

Пример. Проведем однофакторный дисперсионный анализ для данных примера на стр. 58: проверим нулевую гипотезу об отсутствии эффектов обработки и построим 95% доверительные интервалы для эффектов обработки.

Подготовка данных. См. предыдущий пример.

Выбор процедуры. В меню блока статистических методов нажатием клавиши «G» выберем пункт «G = 1-факторный параметрический».

Результаты. На рисунке 19 приведена базовая таблица результатов дисперсионного анализа.

Файл: **Переменных=6** **Измерений=30**

1-ФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ.

Источник	Сумма квадр.	Степ.своб	Средн.квадр.	Сила влияния
факт.1	590.8	5	118.2	0.1496
остат.	224.4	24	9.35	
общая.	815.2	29	28.11	

F(фактор)=12.64, Значимость=0, степ.своб=5, 24

Гипотеза 1: Есть влияние фактора на отклик	
СТАТИСТИЧЕСКИЕ МЕТОДЫ	
F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием	
Нажмите Enter=продолжить или F2=печать экрана >>	

Рис.19. Базовая таблица результатов дисперсионного анализа.

Назначение этого рисунка – дать ответ на вопрос о наличии значимого влияния уровней факторов на исследуемый отклик или, другими словами, о присутствии эффектов обработки. Дадим определения величин, приведенных в этой таблице. Сначала рассмотрим столбец «Сумма квадр.».

В строке «Общая» указана общая сумма квадратов разностей наблюдений и их среднего значения.

В строке «Факт.1» приведен вклад в общую сумму квадратов, обусловленный различиями в уровнях фактора a_j. Часто эту величину называют *суммой квадратов между группами*.

В строке «Остат.» указан вклад в общую сумму квадратов, вызванный случайной изменчивостью данных внутри групп. Его часто называют *суммой квадратов внутри групп*.

Легко видеть, что сумма величин первой и второй строк столбца «Сумма квадр.» таблицы дисперсионного анализа дает величину в третьей строке этого столбца.

Таким образом, смысл анализа вариации данных сводится к выяснению разложения общей суммы квадратов отклонений на две части. Первая из них интерпретируется как вариация, обусловленная введенной моделью, а вторая – как случайная изменчивость данных внутри самой модели.

В случае справедливости нулевой гипотезы каждая из величин в первом столбце таблицы имеет распределение $\sigma^2\chi^2$ со своим числом степеней свободы. Это число степеней свободы указывает второй столбец «Ст. своб». таблицы дисперсионного анализа.

Наконец, в третьем столбце таблицы «Ср.квадр.» находятся частные от деления величин первого столбца на соответствующие величины второго столбца. Отношение этих оценок (нормированные средние квадраты между группами и средние квадраты внутри групп) носит название *F-отношения* и его значение, приведенное снизу от таблицы дисперсионного анализа, как раз и используется для проверки нулевой гипотезы.

Справа от F-отношения указывается минимальный уровень значимости указанной F-статистики (здесь он практически равен нулю), и числа степеней свободы соответствующего F-распределения. Как обычно, если значимость F-статистики близка к нулю, есть основание отвергнуть нулевую гипотезу. Система сравнивает уровень значимости F-статистики с 0,05 и на основе этого сравнения выводит на экран заключение «Есть влияние фактора на отклик».

В четвертой строке таблицы выводится *сила влияния фактора (по Снедекору)*, т.е. величина. Величина силы влияния показывает, какую долю вариации данных определяет модель.

Для данных примера мы получили, что доля фактора стимулирования составляет 14,9% в производительности.

Пакет STADIA содержит также методы для множественных сравнений, но оценки эффектов обработки он не дает. Впрочем, для получения этих оценок можно воспользоваться процедурой Описательной статистики для каждой из переменных, отвечающих различным уровням фактора. Однако следует помнить, что эта процедура строит границы доверительных интервалов, опираясь на оценки дисперсии по каждой группе наблюдений, а не по всей их совокупности. Из-за отсутствия в пакете STADIA возможности непосредственной проверки правомерности применения методов дисперсионного анализа, мы рекомендуем пользователям по крайней мере проводить сравнение результатов критерия Крускала-Уоллиса и однофакторного анализа.

ДВУХФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ В ПАКЕТЕ *STADIA*

Пример. С помощью критерия Фридмана проверим нулевую гипотезу об отсутствии эффектов обработки для данных о зависимости частоты самопроизвольного дрожания мышц рук (тремор) от тяжести специального браслета, одеваемого на запястье. Испытуемым (5 групп по 6 человек) предлагалось провести щуп через лабиринт прибора «Тремометр», число касаний оценивалось в баллах по специальной шкале, при этом на кисть испытуемых надевались браслеты различной массы.

Подготовка данных. В редакторе базы данных пакета введем результаты тестирования испытуемых 1 группы в переменную x1, второй – в переменную x2 и т.д.(рис.20).

Файл: tremor			Переменных=5			Измерений=30	
Var/Cases	1/6	2/6	3/6	4/6	5/6	6/0	7/0
Varname	x1	x2	x3	x4	x5	x6	x7
1	3.01	2.85	2.62	2.63	2.58		
2	3.47	3.43	3.15	2.83	2.70		
3	3.35	3.14	3.02	2.71	2.78		
4	3.10	2.86	2.58	2.49	2.36		
5	3.41	3.32	3.08	2.96	2.67		
6	3.07	3.06	2.85	2.50	2.43		
БЛОК РЕДАКТОРА ДАННЫХ							
F1 Помощь F2 ПечЭкр F3 Чтение F4 Запись F5 Архив F6 Рисунок F7 Очист F8 Преобр F9 Статис F0 Выход							
Вводите в матрицу числа + Enter (работают также: Enter/ Ins/ Del/ Tab и F-ключи							

Рис.20. Ввод результатов тестирования в «Блок редактора данных»

Процедуры непараметрического двухфакторного анализа пакета STADIA однозначно требуют, чтобы данные, отвечающие различным способам обработки (уровням фактора), находились в отдельных переменных. Число наблюдений в каждой переменной должно быть одинаковым, наблюдения, соответствующие одному блоку, должны стоять в I одной строке. Как и в процедурах однофакторного анализа, недопустимо наличие посторонних переменных в файле данных.

Выбор процедуры. Нажав «F9», перейдем в блок статистических методов и выберем в нем пункт «J = 2-факторный Фридмана». Это можно сделать нажатием клавиши «I».

Результаты. Экран вывода результатов этой процедуры и содержит значение статистики Фридмана (рис.21), ее уровень значимости, вычисленный с помощью асимптотического распределения хи-квадрат, и число степеней свободы этого распределения.

Сравнивая полученный уровень значимости с фиксированным (равным 0,05), система выдает сообщение о наличии или отсутствии влияния фактора на отклик.

Файл: Переменных=5 Измерений=30
2-ФАКТОРНЫЙ ДИСПЕРС. АНАЛИЗ ФРИДМАНА. S=22.53 Значимость=1E-4, степ.своб=4 Гипотеза 1: Есть влияние фактора на отклик
СТАТИСТИЧЕСКИЕ МЕТОДЫ
F10 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием
Нажмите Enter=продолжить или F2=печать экрана >>

Рис. 21. Двухфакторный дисперсионный анализ. Значение статистики Фридмана

Пример. Проведем двухфакторный дисперсионный анализ для данных примера 1: проверим нулевую гипотезу об отсутствии эффектов обработки, оценим значения этих эффектов и построим для них 95% доверительные интервалы.

Подготовка данных. Смотри пример 1.

Выбор процедуры. В меню статистических методов выберем пункт «H = 2-факторный параметрический» (клавиша «H»). Появится запрос системы «План: 0=неповторяемый; 1=повторяемый». Под повторяемым (неповторяемым) планом эксперимента в пакете подразумевается план содержащий (не содержащий) повторные измерения при каждом сочетании значений двух исследуемых факторов.

Каждый из этих планов предусматривает свою форму ввода данных. В частности, исходные данные эксперимента без повторных измерений должны представлять собой матрицу размером n x *k*, в которой столбцы отвечают различным способам обработки (*k* уровням первого фактора), строки отвечают различным блокам (n уровням второго фактора), а каждый элемент есть отклик, измеренный пои соответствующем сочетании уровней исследуемых факторов. Этим требованиям в точности соответствуют исследуемые данные.

В рассматриваемом примере в ходе эксперимента проводились повторные наблюдения при каждом сочетании уровней фактора. Однако в таблице приведены только результаты усреднения повторных наблюдений, а исходная информация нам не доступна. То есть на стадии фиксации данных был осуществлен переход от повторяемого плана эксперимента к неповторяемому. Поэтому на запрос системы о плане эксперимента следует указать неповторяемый план.

Результаты. Экран вывода результатов этой процедуры содержит базовую таблицу дисперсионного анализа (рис.22).

Значения F-статистик для каждого из факторов F (фактор 1) и F (фактор 2) и их уровни значимости показывают, что имеется влияние каждого из факторов на отклик.

Файл: **Переменных=5** **Измерений=30**

2-ФАКТОРНЫЙ ДИСПЕРСИОННЫЙ АНАЛИЗ.

Факторный план: неповторяемый

Источник	Сумма квадр.	Степ.своб	Средн.квадр.	Сила влияния
факт.1	1.801	4	0.4504	0.9835
факт.2	0.9074	5	0.1815	0.9364
остат.	0.1645	20	8.223E-3	
общая.	2.873	29	9.908E-2	

F(фактор1)=54.77, Значимость=0, степ.своб=4, 20
F(фактор2)=22.07, Значимость=0, степ.своб=5, 20

Гипотеза 1: Есть влияние фактора на отклик

СТАТИСТИЧЕСКИЕ МЕТОДЫ

F0 Выход F1 Информация F2 Экран на печать / в файл Esc Выход с прерыванием

Нажмите Enter=продолжить >>

Рис.22. Базовая таблица результатов дисперсионного анализа.

Комментарии. К сожалению, процедура не вычисляет оценок параметров двухфакторной модели и их характеристик, а также не дает возможности провести анализ остатков модели на предмет проверки адекватности используемого метода. Эти обстоятельства уменьшают ценность этой процедуры, так как во многих факторных задачах важна именно оценка вклада уровней факторов в наблюдаемый отклик.

ПРИЛОЖЕНИЯ

Приложение 1

Значения t-критерия (по Стьюденту)

Число степеней свободы (n')	Уровни значимости (P)					
	0,2 (20,0%)	0,1 (10,0%)	0,05 (5,0%)	0,02 (2,0 %)	0,01 (1.0%)	0,001 (0,1%)
4	1,533	2,132	2,776	3,747	4,604	8,61
5	1,476	2,014	2,571	3,365	4,032	6,86
6	1,440	1,943	2,447	3,143	3,707	5,96
7	1,415	1,895	2,365	2,998	3,449	5,31
8	1,397	1,860	2,306	2,896	3,355	5,04
9	1,383	1,833	2,262	2,821	3,250	4,78
10	1,372	1,812	2,228	2,764	3,169	4,59
11	1,363	1,796	2,201	2,718	3,106	4,44
12	1,356	1,782	2,179	2,681	3,055	4,32
13	1,530	1,771	2,160	2,650	3,012	4,22
14	1,345	1,761	2,145	2,624	2,997	4,14
15	1,341	1,753	2,131	2,602	2,947	4,07
16	1,337	1,746	2,120	2,583	2,921	4,02
17	1,333	1,740	2,110	2,567	2,898	3,96
18	1,330	1,734	2,101	2,552	2,878	3,92
19	1,328	1,729	2,093	2,539	2,861	3,88
20	1,325	1,725	2,086	2,528	2,845	3,85
21	1,323	1,721	2,080	2,518	2,831	3,82
22	1,321	1,717	2,074	2,508	2,819	3,79
23	1,319	1,714	2,069	2,500	2,807	3,77
24	1,318	1,711	2,064	2,492	2,797	3,75
25	1,316	1,708	2,060	2,485	2,787	3,73
26	1,315	1,706	2,056	2,497	2,779	3,71
27	1,314	1,703	2,052	2,473	2,771	3,69
28	1,313	1,701	2,048	2,467	2,763	3,67
29	1,311	1,699	2,045	2,462	2,756	3,66

30	1,310	1,697	2,042	2,457	2,750	3,64
∞	1,281	1,644	1,957	2,326	2,575	3,29

Приложение 2

Критерии для исключения «выскакивающих» вариант
(по Н. Большову и Н.В. Смирнову, 1965)

n	Уровень достоверности	
	95%	99%
3	0,941	0,988
4	0,765	0,899
5	0,642	0,780
6	0,560	0,698
7	0,507	0,637
8	0,468	0,590
9	0,437	0,555
10	0,412	0,527
11	0,392	0,502
12	0,376	0,482
15	0,338	0,438
20	0,300	0,391
25	0,277	0,362
30	0,260	0,341

Приложение 3

Критические значения коэффициента корреляции рангов (ρ)

n	P		n	P	
	0,05	0,01		0,05	0,01
7	0,78	0,94	24	0,41	0,52
8	0,72	0,88	25	0,40	0,51
9	0,68	0,83	26	0,39	0,50
10	0,64	0,79	27	0,38	0,49
11	0,61	0,76	28	0,38	0,48
12	0,58	0,73	29	0,37	0,48
13	0,56	0,70	30	0,36	0,47
14	0,54	0,68	31	0,36	0,46
15	0,52	0,66	32	0,36	0,45
16	0,50	0,64	33	0,34	0,45

17	0,48	0,62	34	0,34	0,44
18	0,47	0,60	35	0,33	0,43
19	0,46	0,58	36	0,33	0,43
20	0,45	0,57	37	0,33	0,42
21	0,44	0,56	38	0,32	0,41
22	0,43	0,54	39	0,31	0,41
23	0,42	0,53	40	0,31	0,40

Приложение 4

Критические значения коэффициента корреляции (r_{xy})

n'=n-2	Уровни значимости		n' = n-2	Уровни значимости	
	5%	1%		5%	1%
5	0,75	0,87	27	0,37	0,47
6	0,71	0,83	28	0,36	0,46
7	0,67	0,80	29	0,36	0,46
8	0,63	0,77	30	0,35	0,45
9	0,60	0,74	35	0,33	0,42
10	0,58	0,71	40	0,30	0,39
11	0,55	0,68	45	0,29	0,37
12	0,53	0,66	50	0,27	0,35
13	0,51	0,64	60	0,25	0,33
14	0,50	0,62	70	0,23	0,30
15	0,48	0,61	80	0,22	0,28
16	0,47	0,59	90	0,21	0,27
17	0,46	0,58	100	0,20	0,25
18	0,44	0,56	125	0,17	0,23
19	0,43	0,55	150	0,16	0,21
20	0,42	0,54	200	0,14	0,18
21	0,41	0,53	300	0,11	0,15
22	0,40	0,52	400	0,10	0,13
23	0,40	0,51	500	0,09	0,12

24	0,39	0,50	700	0,07	0,10
25	0,38	0,49	900	0,06	0,09
26	0,37	0,48	1000	0,06	0,09

Приложение 5

Критические значения критерия χ^2

n'	P		n'	P	
	0,05	0,01		0,05	0,01
1	3,8	6,6	18	28,9	34,8
2	6,0	9,2	19	30,1	36,2
3	7,8	11,3	20	31,4	37,6
4	9,5	13,3	21	32,7	38,9
5	11,1	15,1	22	33,9	40,3
6	12,6	16,8	23	35,2	41,6
7	14,1	18,5	24	36,4	43,0
8	15,5	20,1	25	37,7	44,3
9	16,9	21,7	26	38,9	45,6
10	18,3	23,2	27	40,1	47,0
11	19,7	24,7	28	41,3	48,3
12	21,0	26,2	29	42,6	49,6
13	22,4	27,7	30	43,8	50,9
14	23,7	29,1	40	55,8	63,7
15	25,0	30,6	50	67,5	76,2
16	26,3	32,0	60	79,1	88,4
17	27,6	33,4	70	90,5	100,4

Составители:	Рецкий Михаил Исаакович Чернов Виктор Иванович Семенов Сергей Николаевич Сереженко Николай Петрович
Редактор	Тихомирова О.А.