АдВ
АДУКАЦЫЯ І ВЫХАВАННЕ
выдавецтва

Серыя «У дапамогу педагогу» заснавана ў 1995 годзе

Навукова-метадычны часопіс
Выдаецца з IV квартала 1995 года
Пасведчанне аб дзяржаўнай рэгістрацыі сродку масавай інфармацыі
№ 641 ад 04.09.2009 г., выдадзенае Міністэрствам інфармацыі Рэспублікі Беларусь
Выходзіць штомесячна з II паўгоддзя 2012 года

Геаграфія

Заснавальнік і выдавец —
Рэспубліканскае ўнітарнае прадпрыемства «Выдавецтва «Адукацыя і выхаванне»
Міністэрства адукацыі Рэспублікі Беларусь

Вул. Будзённага, 21, 220070, г. Мінск;
тэл.: 297-93-24 (адк. сакратар), 297-93-22 (аддзел маркетынгу),
факс: 297-91-49, e-mail: geography@aiv.by, http://www.aiv.by

2(147) люты 2018

ЗМЕСТ

Дасылаючы матэрыялы для публікацыі ў нашым часопісе, аўтары тым самым перадаюць выдаўцу невыключныя маёмасныя правы на ўзнаўленне, распаўсюджванне, паведамленне для ўсеагульнага ведама і іншыя магчымыя спосабы выкарыстання твора без абмежавання тэрыторыі распаўсюджвання (у тым ліку ў электроннай версіі часопіса).

Пераносы некаторых слоў зроблены не па правілах граматыкі, а паводле магчымасцей камп'ютара.

Рэдактар і карэктар *Т. К. Слаута*. Камп'ютарны набор, макет і вёрстка *В. Ю. Лагун*.
Выхад у свет 26.02.2018. Фармат 60×84 $^1/_8$. Друк афсетны. Папера афсетная.
Ум. друк. арк. 7,44. Ул.-выд. арк. 7,44.
Тыраж 604. Заказ 15. Цана свабодная.

Надрукавана ў таварыстве з абмежаванай адказнасцю «СУГАРТ».
ЛП № 02330/427 ад 17.12.2012. Вул. Валгаградская, 6, корп. 2, каб. 287, 220012, г. Мінск.

А. С. Соколов,
старший преподаватель кафедры экологии
Гомельского государственного университета им. Ф. Скорины

МЕТОДЫ СТАТИСТИЧЕСКОЙ ОБРАБОТКИ ДАННЫХ

Многолетний опыт участия автора в жюри естественно-научных секций городских и областных конкурсов научно-исследовательских работ школьников позволил выделить наиболее распространённые недостатки, отрицательно влияющие на оценку конкурсных работ. Ведущее место среди них занимает отсутствие, неправильное применение или явно формальный характер в работах статистических методов обработки полученных результатов наблюдений или экспериментов.

Статистическая обработка результатов полевых или иных исследований является важнейшим компонентом научно-исследовательской работы. Она позволяет существенно увеличить объём информации, извлекаемой из полевых и экспериментальных данных. Только используя статистический аппарат, можно делать выводы о наличии объективных закономерностей, связей между различными характеристиками объектов, степени их влиянии друг на друга, формулировать и доказывать гипотезы.

Несмотря на это в ряде случаев наблюдается неоднозначное отношение молодых исследователей к статистическим методам. Как пишут авторы одного из учебников [1], *«в настоящее время никто не оспаривает необходимости статистического анализа в исследовательской практике. Между тем, у многих исследователей сложилось отношение к нему, как к чему-то второстепенному. Большинство начинающих учёных считают, что главное — провести наблюдение, выполнить эксперимент, то есть, получить данные, а далее можно их, как говорят, статистически обработать. Иными словами, к статистике обращаются не потому, что испытывают в ней потребность, вытекающую из самой сути исследований, а потому, что «так принято».*

Научный работник должен безупречно владеть специальными методиками. Но только этого для научной работы недостаточно. Мало получить научные факты, их необходимо проанализировать. Первый этап — анализ статистический. Именно статистический анализ, а не субъективное ощущение».

Существует большое количество учебной литературы, посвящённой статистическим методам и их применению в научных, в том числе биологических и географических, исследованиях [2—4 и др.]. Поэтому в данном пособии освещаются только самые основы статистики, от которых можно отталкиваться при дальнейшем самостоятельном изучении данного вопроса.

Всякое множество идентифицируемых объектов, отличающихся друг от друга незначительно по конкретному признаку, но сохраняющих сходство по некоторым существенным характеристикам, называется ***совокупностью.*** Совокупностью можно называть стадо животных, популяцию растений, пробы воды из водоёма, группу анкетируемых, количество родившихся за определённый период и т. д.

В природе, как правило, невозможно исследовать все члены совокупности (все деревья в лесу, всю воду в водоёме, всех животных в популяции мышей, все растения на поле пшеницы и т. д.) в силу их огромного количества, поэтому из всей совокупности (которая называется ***генеральной совокупностью***) отбирают некоторую часть, называемую ***выборочной совокупностью***, или ***выборкой***, и изучают её характеристики, чтобы затем по этим характеристикам делать вывод обо всей генеральной совокупности. Количество членов выборки (число наблюдений) называют *объёмом выборки* и обозначают *n*. Изучаемый показатель (размер, концентрация вещества, вариант ответа, плотность и т. д.) называется *признаком*, величина изучаемого показателя называется *значением* признака и обозначается *x*.

Важным свойством выборки является её **случайный характер**. Это означает, что каждый член генеральной совокупности равновероятно может попасть в выборку для проведения эксперимента. То есть вероятность оказаться в выборке одинакова для всех членов генеральной совокупности. Случайность выборки обеспечивает исключение субъективизма, неосознанное стремление подогнать данные под заранее выдвинутую гипотезу. Если выборка не является случайной, то статистический аппарат для её анализа применять бессмысленно.

Случайный характер выборки обеспечивает её ***репрезентативность*** — важнейшее свойство, способность характеризовать всю генеральную совокупность. Если выборка не является репрезентативной, то исследователь может сделать ошибочные выводы обо всех объектах исследования (всей генеральной совокупности)

Любая выборка характеризуется ***распределением признака***. К примеру, изучая связь типа почв с высотой мятлика лугового, из каждой популяции мятлика, растущей на определённом типе почв, отбирается выборка и измеряется значение признака — высота всех растений, в неё входящих. Расположив все измеренные значения в порядке возрастания или убывания (создав ранжированный ряд или ***вариационный ряд***), можно увидеть, сколько раз то или иное значение высоты появилось в выборке. Этот показатель называется *частотой* и обозначается *f*. Построив график распределения, можно графически показать сколько раз то или иное значение появилось в ходе исследования (рис. 1)

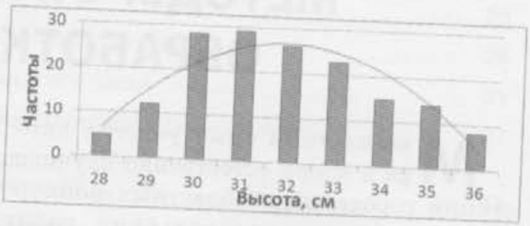


Рисунок 1 — **Пример распределения изучаемого признака**

Чаще же используют относительные частоты встречаемости, измеряемые в процентах или долях от единицы:

$$f = \frac{k}{n} \cdot 100\,\%,$$

где *k* — количество членов выборки с заданным конкретным значением, *n* — общий объём выборки. Сумма относительных частот должна равняться 100 % или 1.

Если кривая, отражающая распределение значений признака по частотам имеет приблизительно колоколообразный вид (как на рис. 1), то такое распределение называется ***нормальным***.

Нормальное распределение подразумевает, что большая часть значений признака находится в районе так называемого ***среднего значения***. При нормальном распределении наиболее часто в выборке встречаются значения, близкие по величине к среднему по выборке и располагающиеся симметрично ему (значений больше среднего и значений меньше среднего приблизительно одинаковое количество). Или если выражать в процентном соотношении (используя относительные частоты встречаемости), то можно говорить, что наибольший процент значений признака находится в районе среднего значения, тогда как всего несколько процентов — по краям кривой.

Именно в зависимости от типа распределения выбираются методы статистического анализа. Если распределение является нормальным, то применяют методы так называемой ***параметрической статистики***.

Показатели вариационного ряда

К главным статистическим характеристикам вариационного ряда можно отнести среднее значение признака и количественную меру варьирования значения признака.

Для характеристики среднего значения можно использовать несколько показателей:

— *средняя арифметическая*. Она представляет собой то значение признака, которое имел бы каждый объект, если бы все значения признака были одинаковы. Находят среднюю арифметическую суммированием всех значений признака и делением получившейся суммы на число объектов (наблюдений). В приведённом выше примере с высотой мятлика среднее арифметическое равно 31,9 см;

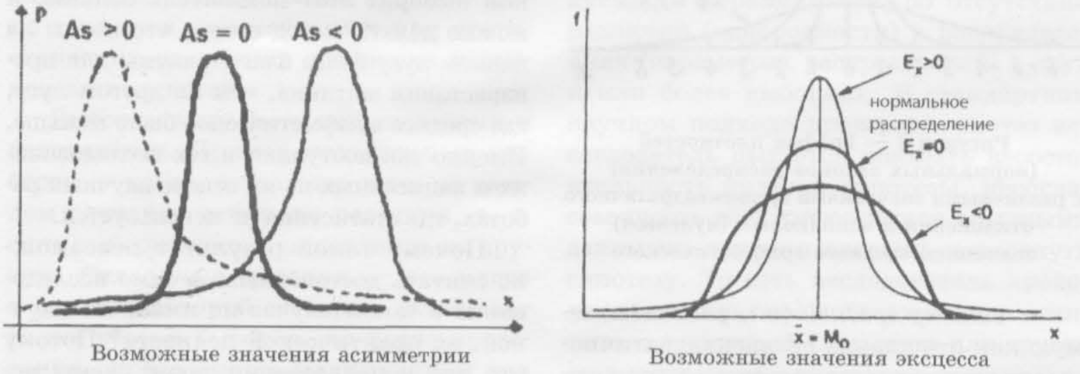
— *мода* — наиболее часто встречающееся значение. Преимуществом моды перед средним арифметическим является то, что она не зависит от крайних значений признака, которые подвержены большим колебанием из-за случайных факторов. В приведённом примере мода равна 31;

— *медиана* — это центральное значение в ранжированном ряду данных. То есть это такое значение, меньше которого 50 % значений признака в выборке и больше которого также 50 % значений. В приведённом примере медиана равна 32.

Ещё два показателя, которые характеризуют особенности распределения случайной величины:

— коэффициент асимметрии (A_s) — показатель, характеризующий симметричность распределения признака относительно среднего арифметического. Если распределение полностью симметрично, то коэффициент асимметрии равен 0. Если он положителен, то правый хвост на графике распределения длиннее левого, т. е. наблюдается смещение значений в сторону больше среднего арифметического, если он отрицателен, то наоборот, левый хвост распределения длиннее правого, и наблюдается смещение значений в сторону меньше среднего арифметического (рис. 2). Величина коэффициента показывает степень смещения. В рассматриваемом примере $A_s = 0{,}23$;

— коэффициент эксцесса (E_x) (коэффициент островершинности) — мера остроты пика распределения случайной величины. Он положителен, если пик распределения около среднего арифметического острый, и отрицателен, если пик очень гладкий (см. рис. 2). В примере $E_x = -0{,}72$.



Возможные значения асимметрии

Возможные значения эксцесса

***Рисунок 2* — Возможные значения асимметрии и эксцесса**

К мерам, показывающим степень варьирования значения признака можно отнести:

— *среднеквадратичное отклонение* (σ). Даёт представление о том, как сильно могут отклоняться от своего центра (среднего арифметического) значения исследуемой случайной величины. Чем больше её величина, тем больше степень её «размазанности» по всему диапазону при одинаковых значениях средней. На рисунке 3 показаны графики распределения значений, характеризующиеся одинаковыми значениями средней и различными значениями среднеквадратичного отклонения. В рассматриваемом нами примере с высотой мятлика $\sigma = 2{,}04$;

— *коэффициент вариации* (V). Среднеквадратичное отклонение даёт абсолютную оценку меры разброса. Поэтому, чтобы понять, насколько разброс велик относительно самих значений (т. е. независимо от их масштаба), требуется относительный показатель. Такой показатель называется коэффициентом вариации и равен отношению дисперсии к средней арифметической, умноженному на 100 %. В нашем примере $V = 6{,}4$ %;

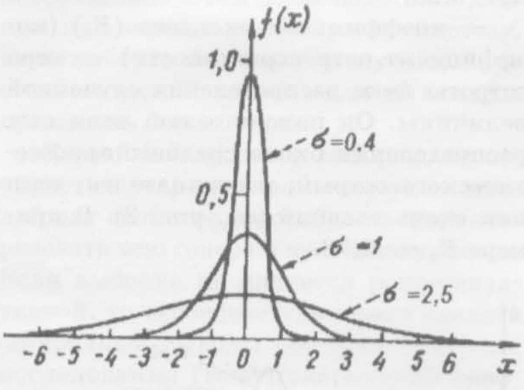


Рисунок 3 — График плотностей (нормальных законов распределения) с различными значениями среднеквадратичного отклонения и одинаковым (нулевым) значением среднего арифметического

— *ошибку средней* (m), рассчитываемую как отношение среднеквадратичного отклонения и квадратного корня величины объёма выборки. На графиках ошибка средней может быть показана с помощью планок погрешностей ⊥ T. Математическая статистика даёт доказательство того, что с вероятностью 95 % среднее значение всей генеральной совокупности находится в пределах от $x - 2m$ до $x + 2m$. В нашем примере $m = 0{,}26$.

Сравнение количественных показателей

После того как исследователь вычислил показатели отдельного вариационного ряда, описанные в предыдущем подразделе, он переходит к следующим шагам и сравнивает эти показатели для различных вариационных рядов. Именно эти сравнения лежат в основе формулировки результатов большей части научно-исследовательских конкурсных работ.

К примеру, исследователь имеет рабочую гипотезу, что высота мятлика лугового не будет одинакова на всех лугах, а будет зависеть от типа почв — на каких-то типах почв он будет более высоким, на каких-то — более низким. Для проверки своей гипотезы он, изучая зависимость высоты мятлика от типа почв, изучил два луга, расположенных на различных почвах. На каждом из этих лугов он отобрал случайным образом по нескольку десятков растений, то есть создал две выборки, составил два вариационных ряда, у каждого из которых определил среднее значение.

Казалось бы, достаточно сравнить два средних арифметических, выявить, в какой выборке этот показатель больше, и можно делать вывод о том, что почвы на данном лугу более благоприятны для произрастания мятлика, чем на другом лугу, где среднее арифметическое было меньше. Именно так поступают в тех исследованиях и написанных на их основе научных работах, где статистика не используется.

Почему такой результат невозможно считать достоверным, а само исследование в таком случае не имеет ни научной, на практической ценности? Потому что при использовании такой схемы исследовательского процесса ***отсутству-***

ет механизм отделения случайностей от закономерностей. На высоту мятлика в различных лугах оказывает влияние не только тип почв, но и огромное количество случайных факторов. Даже если взять две случайные выборки с одного и того же луга и определить их средние значения, они не будут абсолютно идентичными. Всегда какое-то значение будет большим, а какое-то меньшим в силу чисто случайных факторов. И такое (влияние случайных факторов) происходит абсолютно во всех исследованиях, связанных с изучением реальной природной среды. Даже если посадить в одинаковые цветочные горшки два генетически идентичных цветка и поместить их в одни и те же условия влажности, освещённости и др., всё равно их высоты через некоторое время не будут абсолютно идентичны, так как даже в описанном случае влияние случайных факторов всё равно проявляется (хотя и не в такой степени, как в природных экосистемах).

Поэтому после определения основных характеристик двух выборочных совокупностей, взятых с двух лугов с разными типами почв, необходимо проверить, ***существует ли между ними статистически доказанное различие, или обе этих выборки принадлежат к одной и той же генеральной совокупности*** (т. е. различия в типах почв этих двух лугов никак не влияют на высоту мятлика)***, а различия их средних значений вызваны случайными факторами.***

Для такой проверки существует показатель, который называется **t-критерий Стьюдента**. Однако прежде ещё несколько терминов, применяемых при статистических операциях сравнения и оценок взаимосвязей.

Выборки могу быть независимыми и связанными. Пример *независимых* выборок — когда, изучая два луга, на каждом случайно отбирали растения, т. е. одна выборка никак не повлияла на другую. *Связанные* выборки — это когда в качестве второй выборки выступает та же первая выборка, но после проведения определённых манипуляций над ней. Например, имеется случайная выборка пациентов с некоторым заболеванием, которое характеризуется каким-то количественным показателем — признаком. Это первая выборка. Затем этим людям дают определённое лекарство и смотрят, как меняется у них изучаемый признак, строят второй вариационный ряд (выступающий в роли второй выборки), определяют его показатели и сравнивают с первым. В данном случае также в первую очередь необходимо определить, имеется ли статистическая достоверность того, что изменение изучаемого признака связано именно с получаемым пациентами лекарством, а не вызвано иными, в том числе случайными, факторами.

***t*-критерий Стьюдента** можно применять к независимым выборкам. Ограничение заключается в том, что данный критерий применим только тогда, когда исходные данные имеют *нормальное распределение*. При несоблюдении этого условия вместо *t*-критерия должны использоваться аналогичные методы *непараметрической статистики*, среди которых наиболее известными являются ***U*-критерий Манна—Уитни** (для независимых выборок).

Перед выполнением статистических операций, имеющих целью подтвердить различия в нескольких выборках или доказать наличие связи между изменением некоторых показателей, всегда выдвигается так называемая ***нулевая гипотеза*** (H_0). В качестве нулевой гипотезы выступают гипотезы об отсутствии взаимосвязи или корреляции между исследуемыми переменными, об отсутствии различий (однородности) в распределениях (параметрах распределений) в двух и/или более выборках. В стандартном научном подходе проверки гипотез исследователь пытается показать несостоятельность нулевой гипотезы, несогласованность её с имеющимися опытными данными, другими словами, отвергнуть гипотезу. То есть исследователь проводит статистическое исследование, рассчитывает *t*-критерий Стьюдента или другие статистические показатели именно для того, чтобы опровергнуть нулевую гипотезу. Если ему это удаётся, тогда принимается вместо нулевой другая гипотеза, она называется ***альтернативная гипотеза*** (H_a) и исключает нулевую.

Итак, в качестве примера, возьмём две выборки с лугов с различными типами почв и определим для каждого растения значение признака, т. е. высоту:

Луг А	Луг Б
30, 45, 41, 38, 34, 36, 31, 30, 49, 50, 51, 46, 41, 37, 36, 34, 33, 49, 32, 46, 41, 44, 38, 50, 37, 39, 40, 46, 42,	46, 49, 52, 37, 56, 40, 47, 51, 58, 46, 46, 56, 53, 37, 44, 42, 40, 37, 54, 53, 51, 37, 56, 44, 42, 49

Для луга А среднее арифметическое $M = 40{,}2$, среднеквадратичное отклонение $\sigma = 6{,}4$, ошибка средней $m = 1{,}2$; для луга Б $M = 47{,}0$, $\sigma = 6{,}7$, $m = 1{,}3$.

Сейчас необходимо проверить, есть ли статистически доказанные различия в двух выборках или их нет, а различие средних значений вызвано случайными факторами.

t-критерий Стьюдента рассчитывается по следующей формуле:

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}},$$

где M_1 — средняя арифметическая первой сравниваемой совокупности (группы), M_2 — средняя арифметическая второй сравниваемой совокупности (группы), m_1 — средняя ошибка первой средней арифметической, m_2 — средняя ошибка второй средней арифметической.

Помимо самого t-критерия, необходимо рассчитать показатель, который называется **число степеней свободы f**. В данном случае он определяется как сумма объёмов двух выборок минус два, т. е.

$$f = (n_1 + n_2) - 2.$$

По данным о числе степеней свободы и требуемой точности (величины допустимой ошибки, синонимы: уровень значимости, доверительный уровень) находим по таблице критическое значение t-критерия. **Если расчётное значение t-критерия выше критического, то нулевая гипотеза отвергается** и считается статистически доказанным с вероятностью более 95 %, т. е. $p < 0{,}05$, тот факт, что различия в средних значениях высот мятлика на двух лугах обусловлены не случайными факторами, а именно влиянием типа почвы. Таким образом, считается статистически доказанной гипотеза о влиянии типов почв на высоту мятлика (разумеется, пока только лишь для изученных типов почв). **Если же расчётное значение t-критерия окажется меньше критического, то нулевая гипотеза подтверждается**, значит различия сравниваемых величин статистически не значимы, а следовательно, различия в средних значениях случайны и гипотеза о влиянии типов почв на высоту мятлика (пока что!) не получила статистического доказательства. В этом случае необходимо попытаться увеличить объём выборки, постараться исключить влияние наиболее сильных случайных факторов (закладывать площади так, чтобы они различались только по типу почв и были более или менее сходны по другим параметрам — уровню грунтовых вод, экспозиции склонов, подстилающим, отсутствию антропогенного воздействия и т. д. (табл. 1).

Таблица 1 — Критические значения t-критерия

f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$
1	12,706	63,65	18	2,101	2,878	35	2,030	2,724	66—67	1,997	2,652
2	4,303	9,925	19	2,093	2,861	36	2,028	2,719	68—69	1,995	2,650
3	3,182	5,841	20	2,086	2,845	37	2,026	2,715	70—71	1,994	2,648
4	2,776	4,604	21	2,080	2,831	38—39	2,024	2,712	72—73	1,993	2,646
5	2,571	4,032	22	2,074	2,819	40—41	2,021	2,704	74—75	1,993	2,644
6	2,447	3,707	23	2,069	2,807	42—43	2,018	2,698	76—77	1,992	2,642
7	2,365	3,499	24	2,064	2,797	44—45	2,015	2,692	78—79	1,991	2,640

Окончание таблицы 1

f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$	f	$p < 0{,}05$	$p < 0{,}01$
8	2,306	3,355	25	2,060	2,787	46—47	2,013	2,687	80—89	1,990	2,639
9	2,262	3,250	26	2,056	2,779	48—49	2,011	2,682	90—99	1,987	2,632
10	2,228	3,169	27	2,052	2,771	50—51	2,009	2,678	100—109	1,984	2,626
11	2,201	3,106	28	2,048	2,763	52—53	2,007	2,674	110—119	1,982	2,621
12	2,179	3,055	29	2,045	2,756	54—55	2,005	2,670	120—139	1,980	2,617
13	2,160	3,012	30	2,042	2,750	56—57	2,003	2,667	140—159	1,977	2,611
14	2,145	2,977	31	2,040	2,744	58—59	2,002	2,663	160—179	1,975	2,608
15	2,131	2,947	32	2,037	2,738	60—61	2,000	2,660	180—199	1,973	2,605
16	2,120	2,921	33	2,035	2,733	62—63	1,999	2,657	200	1,972	2,601
17	2,110	2,898	34	2,032	2,728	64—65	1,998	2,655	∞	1,960	2,590

Примечание: f — число степеней свободы.

Любые статистические операции используют случайные величины. Поэтому остаётся вероятность ошибки для любого рассчитанного показателя. Эта ошибка называется ***уровень значимости*** и обозначается p. Мы не можем, используя статистический аппарат, её полностью исключить, однако мы можем снизить вероятность этой ошибки до приемлемого уровня вероятности. В большинстве исследований, в том числе экологических, таким приемлемым уровнем является 5 %. Запись «$p < 0{,}05$» означает, что вероятность ошибки рассчитанного значения менее 5 %, «$p < 0{,}01$» — менее 1 %, «$p < 0{,}001$» — менее 0,1 %.

То есть даже если критическое значение, определённое по таблице для случая $p < 0{,}05$ (для случаев с меньшей вероятностью ошибки и критические значения существенно выше), оказалось меньше рассчитанного, что говорит о статистически достоверных различиях выборок, всё равно существует менее 5 % вероятности, что данный вывод ошибочен. Повторимся, этот уровень вероятности ошибки является приемлемым при естественно-научных исследованиях.

Итак, рассчитаем t-критерий для примера, приведённого выше: две выборки с лугов с различными типами почв.

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}} = \frac{|47{,}0 - 40{,}2|}{\sqrt{1{,}2^2 + 1{,}3^2}} = 3{,}84.$$

Вычисляем число степеней свободы, учитывая, что в первой выборке (луг А) её объём $n = 29$, для луга Б $n = 26$:

$$f = (n_1 + n_2) - 2 = (29 + 26) - 2 = 53.$$

По таблице 1 определяем критическое значение t-критерия при числе степеней свободы 53 и $p < 0{,}05$ — это 2,007. Поскольку полученное значение 3,84 больше, чем критическое, то нулевая гипотеза отвергается и принимается, что две эти выборки действительно имеют статистически достоверные различия.

Рассчитаем t-критерий для следующего примера. Сравниваются два водоёма по плотности личинок комара. В качестве выборок выступают совокупности проб из этих водоёмов, в качестве признака — количество личинок в каждой пробе.

Водоём А	Водоём Б
8, 4, 6,10, 8, 7, 12, 3	5, 6, 8, 7, 9, 8, 4, 4

Здесь для водоёма А $M = 7{,}25$, $\sigma = 2{,}96$, $m = 1{,}05$; для водоёма Б $M = 6{,}36$, $\sigma = 1{,}92$, $m = 0{,}68$. Количество степеней свободы $(8 + 8) - 2 = 14$.

Рассчитаем t-критерий:

$$t = \frac{|7{,}25 - 6{,}36|}{\sqrt{1{,}05^2 + 0{,}68^2}} = 0{,}71.$$

Поскольку полученное значение 0,71 меньше критического значения для числа степеней свободы 14 (по таблице оно равно 2,145), то нулевая гипотеза принимается и считается, что статистически достоверных различий по исследуемому показателю между водоёмами нет.

Мы видим, что с увеличением числа степеней свободы снижается величина критических значений t-критерия, поэтому большой объём выборки в большей степени позволяет делать статистически доказанные выводы.

Также помним, что расчёт t-критерия Стьюдента относится к *методам параметрической статистики*, т. е. исходные данные должны иметь *нормальное распределение* — кривая, отражающая распределение значений признака по частотам, имеет приблизительно колоколообразный вид и чем дальше от центра распределения, тем меньше частоты значений признака, максимальные частоты концентрируются около центра.

Однако данные могут и не иметь нормального распределения. В этом случае для сравнения двух независимых выборок по уровню какого-либо признака используют *методы непараметрической статистики*, например ***U*-критерий Манна-Уитни.**

U-критерий не требует наличия нормального распределения сравниваемых совокупностей. Он подходит для сравнения малых выборок: в каждой из выборок должно быть не менее трёх значений признака. Допускается, чтобы в одной выборке было два значения, но во второй тогда должно быть не менее пяти.

Методика расчёта следующая (мы будем её демонстрировать на последнем примере с водоёмами и пробами воды, в которых подсчитывалось количество личинок комара).

Сначала из обеих сравниваемых выборок составляется **единый ранжированный ряд** путём расставления единиц наблюдения по степени возрастания признака и присвоения меньшему значению меньшего ранга. В таблице в первой строчке все значения обеих выборок расположены в едином ряду в порядке возрастания. Значения, относящиеся к первой выборке, показаны полужирным шрифтом, ко второй — обычным (табл. 2).

Таблица 2

Значение признака	**3**	4	4	**4**	5	6	**6**	**7**	7	8	8	**8**	**8**	9	**10**	**12**
Ранг	**1**	2	3	**4**	5	6	**7**	**8**	9	10	11	**12**	**13**	14	**15**	**16**
Скорректированный ранг	**1**	3	3	**3**	5	6,5	**6,5**	**8,5**	8,5	11,5	11,5	**11,5**	**11,5**	11,5	**15**	**16**

Во второй строчке таблицы расставлены ранги значений по возрастанию — от самого меньшего значения, имеющего самый низкий ранг, — к большему. Видно, что в ряду признаков некоторые значения признака равны (три значения равны 3, два значения равны 6, четыре значения равны 8 и т. д.). Поэтому ряд, в котором показаны ранги, нужно скорректировать. В случае равных значений признака у нескольких единиц каждой из них присваивается среднее арифметическое последовательных значений рангов. То есть последовательно идут три значения, равные 4 с рангами 2, 3, 4. Из этих значений рангов мы вычисляем среднее (2 + 3 + 4) / 3 = 3, и это получившееся число присваиваем рангу всех трёх одинаковых значений признака. Если в ряду значений признака идут два значения, равных 7, с рангами 8 и 9 (как в таблице), то находим среднее их рангов — 8,5 — и присваиваем его рангам обоих этих значений. В результате получается скорректированный ряд.

Подсчитаем отдельно сумму рангов для первой выборки и для второй выборки. Для водоёма А она равна 73, для водоёма Б — 63.

Находим значение U-критерия Манна-Уитни по формуле:

$$U = n_1 \cdot n_2 + \frac{n_x \cdot (n_x + 1)}{2} - T_x,$$

где n_1 — объём первой выборки; n_2 — объём второй выборки; n_x — объём выборки, которая соответствует бóльшему значению суммы рангов; T_x — бóльшая из двух сумм рангов.

Полученное значение U-критерия сравниваем по таблице с критическим значением U при заданной численности сопоставляемых выборок (табл. 2):

— если полученное значение U **меньше табличного** или **равно ему**, то признаётся статистическая значимость различий между уровнями признака в рассматриваемых выборках (принимается альтернативная гипотеза). Достоверность различий тем выше, чем меньше значение U;

— если же полученное значение U **больше табличного**, принимается нулевая гипотеза, т. е. отсутствие статистически достоверных отличий.

В нашем случае $U = 27$. Обе выборки имели одинаковый объём — 8, ищем в таблице 3 соответствующее этим объёмам значение. Оно равно 13, т. е. рассчитанное значение намного больше табличного, что говорит об отсутствии статистически доказанных различий между выборками.

Таблица 3 — Критические значения U-критерия Манна-Уитни

n_1	n_2													
	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	12	15	17	20	23	26	28	30	34	37	39	42	45	48
10	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	16	19	23	26	30	33	37	40	44	48	51	55	58	62
12	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	22	26	30	36	40	45	50	55	59	64	67	74	78	83
15	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	26	31	37	42	48	53	59	64	70	75	81	86	92	98
17	28	34	39	45	51	57	63	67	75	81	87	93	99	105
18	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	34	41	48	55	62	69	76	83	90	98	105	112	119	127

t-критерий Стьюдента и U-критерий Манна-Уитни применяется для сравнения ***независимых*** выборок. Если выборки связанные (получены при повторных измерениях одного параметра у одинаковых объектов, например, плотность вредителей на определённых растениях до обрызгивания инсектицидом и плотность вредителей на тех же растениях через три дня после обрызгивания), то применяется парный t-критерий Стьюдента (параметрический) или T-критерий Уилкоксона (непараметрический).

Парный t-критерий рассчитывается по формуле:

$$t = \frac{M_d}{\sigma_d / \sqrt{n}},$$

где M_d — средняя арифметическая разностей показателей одних и тех же объектов, измеренных до и после, σ_d — среднеквадратичное отклонение разностей показателей, n — число наблюдений (объём выборки).

Число степеней свободы оценивается по формуле $f = n - 1$. Далее аналогично t-критерию для независимых выборок полученное значение парного t-критерия сравнивается с критическим значением по той же таблице 1 для соответствующего числа степеней свободы. Если вычисленное значение больше или равно табличному, то нулевая гипотеза отвергается и принимается альтернативная.

T-критерий Уилкоксона не требует наличия нормального распределения. Он показывает, достоверно ли меняются показатели в сторону увеличения или уменьшения в связанной выборке после воздействия какого-либо фактора. Объём выборки должен быть не менее 5 и не более 50. Методика расчёта довольно проста. Сначала необходимо вычислить разность значений для каждой пары измерений — «до» и «после». Определить, в какую сторону направлена большая часть изменений показателя — в сторону возрастания или в сторону уменьшения. На основании этого выделить типичные изменения (в ту сторону, куда направлено большинство изменений) и нетипичные (в противоположную). Затем для каждой вычисленной разности взять её модуль и ранжировать по модулю от меньших значений к большим. Отдельно посчитать сумму рангов, соответствующих нетипичным изменениям. Полученное значение сравнить с критическим по таблице 4 для соответствующего числа наблюдений. Если полученное значение меньше критического или равное ему, то нулевая гипотеза отвергается, если больше, то признаётся статистическая значимость изменений показателя в типичную сторону (принимается альтернативная гипотеза).

Таблица 4 — Критические значения критерия Уилкоксона

T	$p<0{,}05$	$p<0{,}01$	*T*	$p<0{,}05$	$p<0{,}01$	*T*	$p<0{,}05$	$p<0{,}01$	*T*	$p<0{,}05$	$p<0{,}01$	*T*	$p<0{,}05$	$p<0{,}01$
5	0		**14**	25	15	**23**	83	62	**32**	175	140	**41**	302	252
6	2		**15**	30	19	**24**	91	69	**33**	187	151	**42**	319	266
7	3	0	**16**	35	23	**25**	100	76	**34**	200	162	**43**	336	281
8	5	1	**17**	41	27	**26**	110	84	**35**	213	173	**44**	353	296
9	8	3	**18**	47	32	**27**	119	92	**36**	227	185	**45**	371	312
10	10	5	**19**	53	37	**28**	130	101	**37**	241	198	**46**	389	328
11	13	7	**20**	60	43	**29**	150	110	**38**	256	211	**47**	407	345
12	17	9	**21**	67	49	**30**	151	120	**39**	271	224	**48**	426	362
13	21	12	**22**	75	55	**31**	163	130	**40**	286	238	**49**	446	379

Например, на 8 метеостанциях региона измерили количество осадков за 2015 и 2016 гг. Необходимо определить, имеется ли статистически достоверное увеличение количества осадков в 2016 г. по сравнению с 2015 (табл. 5).

Таблица 5 — Пример для применения критерия Уилкоксона

Метеостанция	Осадки в 2015 г., мм	Осадки в 2016 г., мм	Разность	Модуль разности	Ранг разности
1	634	642	8	8	6,5
2	602	612	10	10	8
3	***643***	***635***	***–8***	***8***	***6,5***
4	***652***	***648***	***–4***	***4***	***2***
5	597	603	6	6	3
6	625	632	7	7	4,5
7	***636***	***629***	***–7***	***7***	***4,5***
8	612	613	1	1	1

Для каждой метеостанции вычислили разность осадков в 2016 и 2015 гг. Определили типичные изменения (в сторону увеличения, так как их 5) и нетипичные (в сторону уменьшения, их только 3). Нетипичные изменения выделили полужирным курсивом. Для каждой разности определили и записали её модуль, затем в последней колонке проранжировали значения разности рангов.

Сумма рангов, соответствующих нетипичным изменениям, равна:

$$T = 6{,}5 + 2 + 4{,}5 = 13.$$

По таблице смотрим критическое значение для числа наблюдений, равного восьми. Оно равно 5. Поскольку $13 > 5$, то принимается нулевая гипотеза об отсутствии статистически достоверного изменения количества осадков.

Окончание в следующем номере.