

УДК 631.421

DOI: 10.19047/0136-1694-2020-102-164-182



**Ссылки для цитирования:**

Самсонова В.П., Мешалкина Ю.Л. Часто встречающиеся неточности и ошибки применения статистических методов в почвоведении // Бюллетень Почвенного института имени В.В. Докучаева. 2020. Вып. 102. С. 164-182. DOI: 10.19047/0136-1694-2020-102-164-182

**Cite this article as:**

Samsonova V.P., Meshalkina J.L., Common inaccuracies and errors in the application of statistical methods in soil science, Dokuchaev Soil Bulletin, 2020, V. 102, pp. 164-182, DOI: 10.19047/0136-1694-2020-102-164-182

## **Часто встречающиеся неточности и ошибки применения статистических методов в почвоведении**

© 2020 г. В. П. Самсонова<sup>1\*</sup>, Ю. Л. Мешалкина<sup>1,2\*\*</sup>

<sup>1</sup>МГУ им. М.В. Ломоносова, Россия,  
119991, Москва, Ленинские горы, 1,

\*e-mail: [ykbun@mail.ru](mailto:ykbun@mail.ru).

<sup>2</sup>РГАУ–МСХА имени К.А. Тимирязева, Россия,  
127550, Москва, ул. Тимирязевская, 49,

\*\*<https://orcid.org/0000-0003-1513-2439>, e-mail: [jlmesh@list.ru](mailto:jlmesh@list.ru).

Поступила в редакцию 31.03.2020, после доработки 01.04.2020,  
принята к публикации 22.06.2020

**Резюме:** Рассматриваются наиболее распространенные неточности и ошибки применения статистических методов, встречающиеся в отечественных публикациях по почвоведению. При обозначении случайных величин и параметров распределений греческими буквами нужно обозначать те, что относятся к генеральным совокупностям, а латинскими – к выборочным. Подробное описание эксперимента и того, к чему относятся повторности, позволяет делать корректные выводы из работы. Необходимо избегать мнимых повторностей, когда результаты в близко расположенных точках опробования рассматриваются как характеристики изменчивости почв на больших расстояниях. Расширение списка описательных статистик позволит использовать конкретное исследование в мета-анализе. Расчет доверительного интервала для среднего с использованием критерия Стьюдента при

разных уровнях значимости расширяет рамки возможных значений среднего, однако такой подход оправдан лишь в том случае, если показатель не слишком сильно отличается от нормального распределения. При проверке статистических гипотез необходимо обращать внимание не только на уровень значимости, но и на мощность критерия. Гипотеза о нормальности распределения может быть проверена при помощи различных критериев. Успех применения критерия зависит не только от истинности нулевой гипотезы (действительно нормального распределения), но и от других причин: от объема выборки и от альтернатив, относительно которых критерий проверяет гипотезу. Любое утверждение о виде связи между признаками на основании коэффициента корреляции (Пирсона или Спирмена) бессмысленно без указания числа повторностей, так как именно число повторностей определяет значимость отличия коэффициента корреляции от нуля. Предлагается, чтобы авторы и рецензенты статей обращали более пристальное внимание на такие ошибки.

**Ключевые слова:** обозначения статистик, мнимые повторности, уровень значимости, доверительный интервал, проверка гипотез, мощность критерия, коэффициент корреляции.

## Common inaccuracies and errors in the application of statistical methods in soil science

V. P. Samsonova<sup>1\*</sup>, J. L. Meshalkina<sup>1,2\*\*</sup>

<sup>1</sup>*Lomonosov Moscow State University,  
1 Leninskie Gori, Moscow 119234, Russian Federation,  
\*e-mail: [ykbun@mail.ru](mailto:ykbun@mail.ru).*

<sup>2</sup>*Russian State Agrarian University – Moscow Agricultural Academy  
named after K. A. Timiryazev,  
49 Timiryazevskaya Str., Moscow 127550, Russian Federation,  
\*\*<https://orcid.org/0000-0003-1513-2439>, e-mail: [jlmesh@list.ru](mailto:jlmesh@list.ru).*

*Received 31.03.2020, Revised 01.04.2020, Accepted 22.06.2020*

**Abstract:** The most common inaccuracies and errors in the application of statistical methods found in Russian publications on soil science are considered. When designating random variables and distribution parameters in Greek letters, it is necessary to designate those that refer to general populations, and Latin letters – to sampling ones. A detailed description of the experiment and what the replications relate to allows you to draw correct conclusions from the study. It is necessary to avoid pseudoreplication when

results at closely located sampling points are considered as characteristics of soil variability over large distances. Expanding the list of descriptive statistics will allow you to use a specific study in meta-analysis. Calculating the confidence interval for the average using the Student's test at different significance levels expands the scope of possible values of the average, but this approach is justified only if the indicator does not differ too much from the normal distribution. When testing statistical hypotheses, it is necessary to pay attention not only to the level of significance, but also to the power of the criterion. The normality distribution hypothesis can be tested using various criteria. The success of applying the criterion depends not only on the validity of the null hypothesis (a truly normal distribution), but also on other reasons: on the sample size and on the alternatives for which the criterion tests the hypothesis. Any statement about the type of relationship between features based on the correlation coefficient (Pearson or Spearman) is meaningless without specifying the number of replicates, since it is the number of replicates that determines the significance of the difference between the correlation coefficient and zero. It is proposed that authors and reviewers pay closer attention to such errors.

**Keywords:** statistics designations, pseudoreplication, significance level, confidence interval, hypothesis testing, criterion power, correlation coefficient.

## ВВЕДЕНИЕ

Практически любая научная публикация в российских журналах по почвоведению сопровождается упоминанием о том, что "...данные были статистически обработаны...". Более того, наличие статистической обработки часто воспринимается как некая индульгенция правильности исследования. Чаще всего, такая "статистическая обработка" сводится к расчету описательных статистик без дальнейшего обсуждения полученных результатов. Аналогичная ситуация наблюдалась ранее во многих зарубежных журналах, однако после публикаций о необходимости более жесткого редакционного контроля поступающих статей, например, статьи Ричарда Вебстера (Webster, 2001), положение изменилось к лучшему.

Цель настоящей работы – обсудить неточности, ошибки и неправильные интерпретаций результатов статистической обработки данных, часто встречающиеся в публикациях, связанных с

почвоведением. Задачи, решаемые статистикой в почвоведении, делятся на две большие группы: 1) оценивание неизвестных параметров и 2) проверка гипотез. В таком порядке мы их и будем рассматривать.

## ОЦЕНИВАНИЕ

### 1. Отсутствие единого стандарта для обозначения статистических характеристик

Разнобой в обозначениях одинаковых по смыслу статистик – не такая уж безобидная ситуация. Казалось бы, если в статье приведена расшифровка обозначений, то какие могут быть претензии к авторам. Однако неодинаковость обозначений приводит к путанице понятий. Так, среднее арифметическое может быть обозначено как  $\mu$ ,  $X$ ,  $M$ ,  $\bar{x}$ . Буквы греческого алфавита в статистике зарезервированы для обозначения параметров генеральной совокупности, то есть модельного представления об изучаемой величине ([ГОСТ Р 50779.10-2000](#)). Параметры генеральной совокупности – это по определению постоянные величины. Так что  $\mu$  – это *постоянная* величина, генеральное среднее и обозначать ею выборочное значение недопустимо. Все остальные показатели  $X$ ,  $M$ ,  $\bar{x}$  – оценки генерального значения. В силу того, что оценка той или иной статистической характеристики всегда определяется по выборке, это величина случайная, значения которой изменяются от выборки к выборке. Буква  $X$  зарезервирована для обозначения выборочной случайной величины. Обозначение среднего арифметического буквой  $M$  является устаревшим и в соответствующем ГОСТе не встречается ([ГОСТ Р 50779.10-2000](#)).

То же самое можно сказать и относительно характеристик варьирования, а именно: о дисперсии и стандартном отклонении. Обозначения  $\sigma^2$  и  $\sigma$  указывают на генеральную совокупность, и это величины постоянные, а их оценки должны обозначаться как  $s^2$  и  $s$ .

## 2. Описание методики эксперимента

Часто встречаются работы, в которых скупо описывается методика проведения эксперимента, что затрудняет обобщение результатов. Число повторностей, способ размещения точек опробования, характер отбираемого образца (индивидуальный или смешанный), его размеры и форма ([Кондрашкина, 1991](#)) могут сильно влиять на конечные выводы. Так, например, слова "...повторность определения трехкратная" могут означать:

- 1) трехкратное определение химического показателя из одного индивидуального образца;
- 2) трехкратное определение того же показателя из смешанного образца, составленного из трех индивидуальных проб;
- 3) однократное определение в каждом из трех индивидуальных образцов, отобранных в одном почвенном горизонте в одном разрезе;
- 4) однократное определение в каждом из трех образцов, отобранных в одном почвенном горизонте трех разрезов, размещенных на обследуемой территории.

Очевидно, что результаты определений в описанных случаях относятся к разным объектам исследования. В первом случае характеризуется неодинаковость значений, обусловленная аналитическими погрешностями, в третьем – характеризуется суммарная неодинаковость значений в пределах разреза и аналитическая погрешность, в последнем случае – суммарная аналитическая погрешность и пространственная неоднородность в пределах территории. Второй случай – более сложный, поскольку на аналитическую и пространственную погрешности накладывается еще и погрешность, обусловленная смешиванием. Если не указано, как конкретно проводился эксперимент, то непонятно, что же имели в виду авторы, сообщая о трехкратной повторности.

## 3. Интерпретация доверительного интервала для аналитических ошибок

Если пробы анализируются в аккредитованной лаборатории, то результат может быть представлен в виде  $x \pm s_r$  или  $x \pm s_R$ , где  $x$  – значение свойства,  $s_r$  и  $s_R$  – аналитические ошибки в условиях повторяемости и воспроизводимости, соответственно ([ГОСТ](#)

[Р ИСО 5725-6-2002](#)). Это означает, что определение велось в одной повторности, а  $s_r$  и  $s_R$  характеризуют аналитическую ошибку метода для единичной лаборатории и группы лабораторий, соответственно.

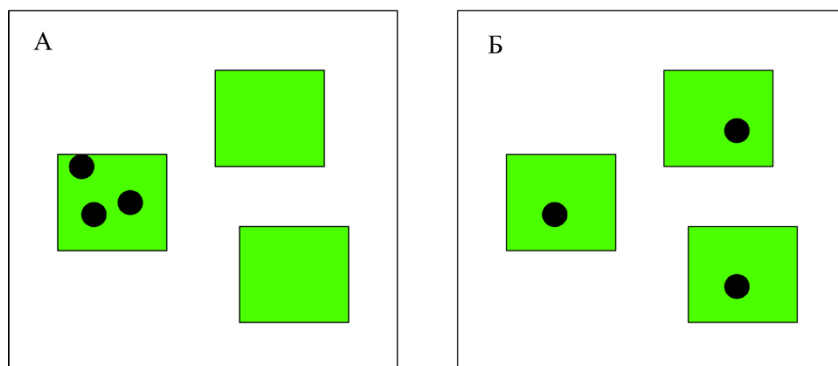
Отметим, что часто метод определения в некотором диапазоне концентраций характеризуется не абсолютной, а относительной ошибкой, например, для определений на хроматографе ([ПНД Ф 16.2.2:2.3.71-2011](#)). В этом случае результат выдается как  $X \pm \Delta$ , где  $\Delta = \delta \cdot X$ , а  $\delta$  – показатель точности, который при доверительной вероятности 95 % равен  $\delta = k s_R$ , где  $k$  – коэффициент охвата, равный 2 (округление значения 1.96, характерного для нормального распределения и доверительной вероятности 95 %).

Отсутствие ясного представления о том, чем характеризуется вариабельность первичных данных, может вносить большую путаницу при попытках их интерпретации и, более того, препятствовать их использованию при различного рода обобщениях, поскольку сравнение средних значений требует оценок дисперсий.

#### **4. Мнимые повторности**

Проблема мнимых повторностей (pseudoreplication), широко обсуждаемая в биологической литературе ([Козлов, 2003](#); [Hurlbert, 1984](#)), на самом деле является проблемой определения генеральной совокупности для проводимого исследования. За пределами чисто математических конструкций генеральную совокупность в почвоведении можно представить как объект, относительно которого делается вывод, основанный на том или ином способе опробования и определения его характеристик, физических, химических или других свойств. Этот объект должен иметь ясные границы, отделяющие его от похожих объектов. Способ опробования и метод определения показателей должны быть по крайней мере не слишком сильно отличающимися, например, размеры проб должны быть примерно одинаковыми, свойства должны определяться по единой методике, желательно, чтобы присутствовал контроль качества определений и тому подобное.

На рисунке 1 приведена схема, объясняющая понятие “мнимые повторности”.



**Рис. 1.** Иллюстрация для понимания термина “мнимые повторности”. Маленькие круги – элементы опробования (точки отбора проб, разрезы и т. п.), маленькие квадраты – неоднородности 1-го уровня, большие квадраты – неоднородность 2-го уровня. Относительно неоднородностей 2-го уровня на схеме А представлены мнимые повторности, на схеме Б – они отсутствуют.

**Fig. 1.** Illustration for understanding the term “pseudoreplication”. Small circles represent sampling elements (sampling points, soil profiles, etc.), small squares represent heterogeneities of the 1<sup>st</sup> level, large squares represent heterogeneities of the 2<sup>nd</sup> level. Concerning the heterogeneities of the 2<sup>nd</sup> level, pseudoreplication is presented in scheme A, and they are absent in scheme B.

Предположим, что на некоторой территории присутствуют пространственные объекты двух уровней. Для характеристики объекта 1-го уровня можно заложить повторности по схеме А. Например, задача заключается в том, чтобы охарактеризовать почвенные свойства только вокруг конкретного разреза, и вокруг него были заложены разрезы (прикопки), то в этом случае нет мнимых повторностей и эксперимент вполне корректен (рис. 1). При изучении объектов 2-го уровня схема А будет представлять собой мнимые повторности, а схема Б – не будет. Например, если под характеристикой некоторого классификационного почвенного выдела, например, типа почвы, имеются в виду схожие почвенные профили, распространенные на площади в сотни или тысячи км<sup>2</sup>, то ясно, что сочетание условий почвообразования в конкретных

точках будет приводить к образованию несколько различных почвенных профилей, причем, чем дальше отстоят друг от друга эти профили, тем, скорее всего, различия будут больше, при условии, что, например, пробы отбирались одинаково и анализ проводился по одинаковой методике. Поэтому если вокруг одного почвенного разреза заложено 10 разрезов (прикопок) и утверждается, что свойства почвы измерены для некоего *типа почвы* с повторностью 10 и этот *тип почвы* обладает такой изменчивостью, то такое утверждение подпадает под определение мнимых повторностей ([Козлов, 2003](#)), поскольку в пределах больших расстояний свойства почвенного объекта, диагностируемого как определенный тип почвы, могут существенно различаться. Распространение результатов, полученных на малой территории, на весь ареал почвенного типа проблематично.

В общем случае, чем больше площадь, на которую распространяются результаты точечного опробования, тем больше вероятность отклонения от истинного значения для всех статистических характеристик. Большое количество примеров мнимых повторностей в экологических исследованиях приведено в книге ([Козлов, 2003](#)).

## **5. Необходимые описательные статистические характеристики**

В качестве статистических характеристик обычно в статьях приводятся среднее арифметическое и одна из характеристик варьирования: стандартное отклонение, ошибка среднего, коэффициент вариации или дисперсия. Добавление в публикациях информации о минимальных и максимальных значениях показателя, а также о медиане и квартилях (табл. 1) позволяет существенно увеличить информацию о результатах эксперимента и будет способствовать использованию их при обобщении данных в обзорных статьях с использованием приемов мета-анализа ([Gurevitch, 2001](#); [Oldfield et al., 2019](#)).



**Таблица 1.** Рекомендуемые статистики при описании результатов почвенных экспериментов и их обозначения

**Table 1.** Recommended statistic approaches for describing the results of soil experiments and their designation

Статистика	Обозначение
Объем выборки	$n$
Среднее	$\bar{x}$
Стандартное отклонение	$s$
Минимум	$min$
Нижний квартиль	$Q_{25}$
Медиана	$Med$ или $Q_{50}$
Верхний квартиль	$Q_{75}$
Максимум	$max$

#### 6. Доверительный интервал для среднего значения

Результаты анализов почвенных объектов часто публикуются в виде

$$\bar{x} \pm s, \quad (1)$$

Как было показано выше, в зависимости от плана эксперимента, величина  $s$  (стандартное отклонение) может характеризовать разные величины, но об этом практически всегда умалчивается. Более того, часто под  $s$  имеется в виду ошибка среднего

$$s_{\bar{x}} = \frac{s}{\sqrt{n}},$$

которая меньше стандартного отклонения в  $\sqrt{n}$  раз, где  $n$  – объем выборки.

Чем больше повторность  $n$ , тем меньше ошибка среднего и тем более точным кажется результат. Однако следует иметь в виду, что точность оценки относится к фиксированному способу опробования, зависящего от  $n$ . В случае нормального распределения вероятность нахождения среднего в интервале (1) равна примерно 67%, то есть за пределами интервала находится около трети возможной изменчивости среднего, а об этом забывается.

Расчет доверительного интервала для среднего

$$\bar{x} \pm t_{\alpha, \nu} S_x \quad (2)$$

с использованием  $t$ -критерия Стьюдента при разных уровнях значимости  $\alpha$  расширяет рамки возможных значений среднего, однако он оправдан лишь в том случае, если показатель имеет нормальное распределение (или не слишком сильно отличается от нормального). В противном случае при расчете можно получить абсурдные результаты, например, отрицательные значения нижней границы признака.

Важное значение при вычислении доверительного интервала имеет доверительная вероятность  $P$  или связанный с доверительной вероятностью уровень значимости  $\alpha = 1 - P$ . Чем меньше уровень значимости  $\alpha$ , тем шире доверительный интервал, тем меньше вероятность “пропустить” важные значения. Однако при малом числе повторностей доверительный интервал будет очень широким. Увеличение  $\alpha$  сужает интервал, но увеличивает возможность неправильного решения. [ГОСТ Р ИСО 5725-6-2002](#) рекомендует 95% доверительный интервал.

## ПРОВЕРКА ГИПОТЕЗ

### 7. Мощность критерия

Статистическая гипотеза – это некоторые утверждения о генеральной совокупности, проверяемые на основании выборки из этой совокупности. Статистические гипотезы всегда парные:  $H_0$  – нулевая гипотеза всегда имеет альтернативу, например:

$H_0$ : рН = 5.0 (среднее значение рН почвы на конкретном поле равно 5.0),

$H_1$ : рН = 6.0 (простая альтернатива),

$H_1$ : рН > 6.0 (сложная альтернатива, без указания конкретных значений).

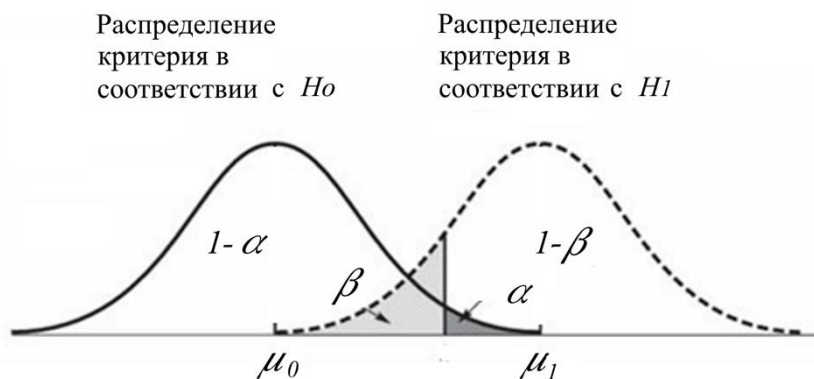
Гипотезы проверяются при помощи статистических критериев, по которым можно оценить вероятность правильности выдвигаемых гипотез. Статистические критерии *не доказывают* правильность утверждений, они лишь оценивают вероятность событий при тех или иных условиях, например, вероятность того,

что экспериментальные данные соответствуют генеральной совокупности (в нашем случае – почвам со средним значением  $pH = 5.0$ ).

Процедура проверки гипотез заключается в расчете значения критерия и сравнения его с критическим значением, соответствующим определенному уровню доверительной вероятности  $P$  (или уровню значимости  $\alpha$ ). Уровень значимости  $\alpha$  оценивает так называемую ошибку первого рода, т. е. вероятность отвергнуть правильную нулевую гипотезу, когда она верна. В докомпьютерные времена, когда вычисленное значение критерия нужно было сравнивать с теоретическим значением, эти теоретические значения печатались в специальных таблицах для фиксированных значений  $\alpha$  ( $0.10$ ;  $0.05$ ;  $0.01$ ;  $0.001$ ). В настоящее время можно рассчитать уровень значимости в программе Excel, не говоря уже о специальных статистических программах. Однако теперь возникает вопрос: как поступать, если в результате вычислений получено значение  $\alpha = 0.049$  или  $0.051$ ? Отвергнуть или принять нулевую гипотезу? Решение остается за исследователем, но значение  $\alpha$  необходимо указывать в публикации.

Поскольку есть альтернативная гипотеза (гипотезы), то принимая нулевую гипотезу, мы рискуем совершить ошибку второго рода – гипотезу принимаем, а на самом деле она неверна. Вероятность такой ошибки –  $\beta$ .

Вероятность  $1 - \beta$  называется мощностью критерия и позволяет оценить способность критерия различать альтернативы. Из рисунка 2 видно, что уменьшая ошибку первого рода ( $\alpha$ ), мы будем реже отвергать нулевую гипотезу, но в то же время ошибка второго рода будет расти, то есть риск принять нулевую гипотезу, когда она в самом деле неверна, будет возрастать. Из этого же рисунка следует, что, если ограничиться рассмотрением лишь ошибки первого рода, то мы полностью закрываем глаза на возможную альтернативу, которая может быть весьма существенной.



**Рис. 2.** Иллюстрация ошибок первого и второго рода при проверке гипотезы  $H_0 : \mu_0 = \mu_1$  против альтернативной гипотезы  $H_1 : \mu_0 < \mu_1$ .  
**Fig. 2.** Illustration of errors of the first and the second kind when testing the hypothesis  $H_0 : \mu_0 = \mu_1$  against the alternative hypothesis  $H_1 : \mu_0 < \mu_1$ .

Например, мы предполагаем, что новый препарат против сорняков увеличивает урожайность пшеницы. Нулевая гипотеза будет свидетельствовать о том, что урожайность не изменяется. Простая альтернатива – урожайность увеличивается на 10%. Если мы заинтересованы в том, чтобы препарат был оценен как обеспечивающий повышение урожайности, мы можем увеличивать  $\alpha$ , то есть отвергать нулевую гипотезу об отсутствии влияния препарата при более высокой ошибке первого рода. Однако если препарат не влияет, то мы не будем отвергать нулевую гипотезу, хотя она на самом деле неверна (ошибка второго рода). Последствия ошибки первого рода будут выражаться в том, что нулевую гипотезу о неизменности урожая в результате воздействия отвергли, посчитали препарат действующим, а на самом деле он ни на что не влияет, однако на его производство затрачены большие деньги. Последствия ошибки второго рода – препарат посчитали неэффективным, а он на самом деле дает заметную прибавку урожайности. Решение о том, какая из этих ошибок более важна, лежит за пределами статистики.

Приведенный пример показывает, что статистика дает лишь

оценку шансов того или иного исхода, но не доказывает их однозначно. Поэтому фраза о том, что "...результаты достоверны, поскольку обработаны статистически", на самом деле ничего не означает.

## 8. Нормальность распределения

Нормальность (подчинение распределения случайной величины нормальному закону) генеральной совокупности, из которой извлекается выборка, характеризующая объект, для большинства статистических методов является необходимым условием их применимости.

Простым индикатором того, что выборочные данные получены из нормальной совокупности, служит коэффициент вариации. Если выборочный коэффициент вариации (для свойств, принимающих положительные значения)

$$V = \frac{s}{\bar{x}} * 100\%$$

превышает 70%, то, скорее всего, распределение отличается от нормального. В более "строгих" руководствах ([Кобзарь, 2012](#)) предлагается считать распределение отличающимся от нормального, если коэффициент вариации больше 35%. Для проверки гипотезы о типе распределения требуется достаточно большое число повторностей. Так, в соответствии с [ГОСТ Р ИСО 5725-6-2002](#), если число повторностей меньше 15, то вообще никакие проверки не проводятся и распределение считается нормальным, хотя в соответствии с предыдущей рекомендацией, для выборок, имеющих коэффициент вариации более 70%, все-таки стоило бы хотя бы провести выбраковку наиболее сильно "отскакивающих" значений.

Гипотеза о нормальности распределения может быть проверена при помощи различных критериев, некоторые из них приведены в таблице 2 ([Кобзарь, 2012](#); [Лемешко, Рогожников, 2012](#)). В почвоведении применяется лишь небольшая часть.

Успех применения критерия зависит не только от истинности нулевой гипотезы (действительно нормального распределения), но и от других причин – от объема выборки и от альтерна-

тив, относительно которых он проверяет гипотезу. Так в статье ([Лемешко, Рогожников, 2012](#)) показано, что критерий Шапиро-Уилка наилучшим образом работает против логарифмически нормального закона, а критерий Шпигельхальтера – против логистического. В книге ([Кобзарь, 2012](#)) приведена ранжированная таблица критериев по отношению к асимметрии и эксцессу альтернативных распределений, против которых они являются наилучшими.

**Таблица 2.** Критерии для проверки гипотезы о нормальности распределения

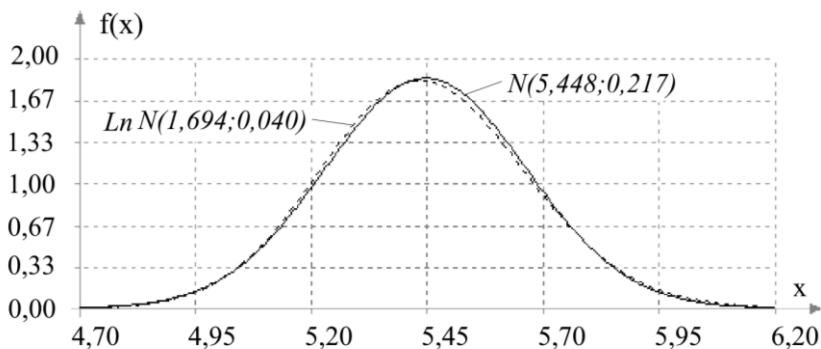
**Table 2.** Criteria for testing the normal distribution hypothesis

1	<b>Критерий Шапиро-Уилка *</b>	11	<b>Критерий асимметрии и эксцесса</b>
2	Критерий Дарбина	12	Критерий Д'Агостино
3	Критерий Васичека	13	Критерий Дэвида-Хартли-Пирсона
4	<b>Критерий хи-квадрат</b>	14	Критерий Андерсона-Дарлинга
5	Критерий Филлибена	15	<b>Критерий Колмогорова-Смирнова</b>
6	Критерий Мартинса-Иглевича	16	Критерий Лина-Мудхолкара
7	Критерий Шпигельхальтера	17	Критерий Саркади
8	Критерий Смирнова-Крамера-фон Мизеса	18	Критерий Локка-Спурье
9	Критерий Оя	19	Критерий Хегази-Грина
10	Критерий Муроты-Такеучи		

**Примечание.** \*Жирным выделены наиболее часто используемые критерии.

В работе ([Лемешко, Рогожников, 2012](#)) также показано, что для ряда надежных экспериментальных данных практически все применяемые критерии не отвергают гипотезу о нормальном распределении ошибок, однако, "...с другой стороны, очевидно, что для описания ошибок измерений в данных экспериментах можно подобрать и другие параметрические модели законов, не менее хорошо согласующиеся с результатами измерений", иллюстраци-

ей чего могут быть два разных статистических распределения, приведенных на рисунке 3. При такой близости кривых, ни один из критериев не сможет их различить.



**Рис. 3.** Иллюстрация близости нормального (N) и логарифмически нормального (Ln N) распределений (Лемешко, Рогожников, 2012).

**Fig. 3.** Illustration of the proximity of normal (N) and logarithmically normal (Ln N) distributions (Lemeshko, Rogozhnikov, 2012).

Для использования в качестве теоретической модели распределения, подчиняющегося тому или иному закону, должны быть веские основания. В противном случае можно обнаружить, что выборочные распределения могут быть аппроксимированы разными законами.

### 9. Коэффициенты корреляции

Коэффициент корреляции Пирсона – один из наиболее используемых индикаторов для оценки силы взаимосвязи двух признаков. Однако часто забывается, что 1) коэффициент корреляции Пирсона оценивает степень линейной взаимосвязи нормально распределенных совокупностей и 2) оценка значимости “силы” связи зависит не только от значения коэффициента корреляции, но и от числа повторностей ( $n$ ), при которых этот коэффициент был рассчитан. Без указания числа повторностей утверждение, например, что “между признаками существует тесная связь”, если ко-

эффицент корреляции равен  $r = 0.8$ , бессмысленно, поскольку при  $n < 7$  и  $\alpha = 0.05$  такой коэффициент незначимо отличен от нуля. В современных статистических программах значимые коэффициенты корреляции при выбранном уровне значимости выделяются цветом, что упрощает интерпретацию результатов.

Похожие ошибки возникают и при сравнении коэффициентов корреляции. Утверждение, что некое воздействие привело к росту коэффициента корреляции между признаками от 0.5 до 0.6 без специальной проверки также бессмысленно, поскольку решение вопроса о различиях зависит и от объемов выборки, и от уровня значимости. Так, например, при объеме выборок  $n = 10$  коэффициенты корреляции  $r_1 = 0.8$  и  $r_2 = 0.2$  и уровне значимости  $\alpha = 0.05$  не различаются (минимальна повторность для различий составляет  $n = 14!$ ).

Непараметрический коэффициент корреляции Спирмена позволяет избежать ошибок, возникающих из-за присутствия в выборках “отскоков”, обуславливающих неоправданно высокие коэффициенты корреляции Пирсона, однако он также зависит от повторностей. Таблицы для оценки значимости коэффициентов корреляции можно найти в книге Е.А. Дмитриева ([Дмитриев, 1995](#)) или в интернете. Отметим, что, если в выборке присутствуют “отскоки”, то значения коэффициентов Спирмена будут иметь более низкие значения по сравнению с коэффициентами Пирсона. Для выборок из нормальных распределений коэффициенты будут практически одинаковы. Подробно различные аспекты применения и интерпретации коэффициентов корреляции обсуждаются в книге Ю.Н. Благовещенского ([Благовещенский, 2009](#)).

## ЗАКЛЮЧЕНИЕ

Перечисленные ошибки и неточности отнюдь не исчерпывают всех возможных огрехов в применении статистики в почвоведении. Так, в стороне остались вопросы расчетов необходимых повторностей для обоснования тех или иных выводов, вопросы планирования эксперимента для того, чтобы избежать мнимых повторностей, вопросы множественного сравнения средних значений и обоснованности применения непараметрических оценок и



другие. Каждая из этих тем требует отдельного подробного обсуждения. Однако, если на обсуждаемые вопросы будут обращать внимание авторы статей и рецензенты, качество публикаций в отечественных журналах должно улучшиться, а интерпретации результатов будут более обоснованы.

## СПИСОК ЛИТЕРАТУРЫ

1. *Благовецкий Ю.Н.* Тайны корреляционных связей в статистике. М.: [ИНФРА-М](#), 2009. 158 с.
2. ГОСТ Р ИСО 5725-6-2002. Точность (правильность и прецизионность) методов и результатов измерений, М.: Стандартинформ, 2009. 58 с.
3. ГОСТ Р 50779.10-2000. Статистические методы. Вероятность и основы статистики. Термины и определения, М.: Стандартинформ, 2005. 46 с.
4. *Дмитриев Е.А.* Математическая статистика в почвоведении. М.: МГУ, 1995. 320 с.
5. *Кобзарь А.И.* Прикладная математическая статистика. Для инженеров и научных работников. М.: ФИЗМАТЛИТ, 2012. 816 с.
6. *Козлов М.В.* Мнимые повторности (pseudoreplication) в экологических исследованиях: проблема, не замеченная российскими учеными // Журнал общей биологии. 2003. Т. 64. № 4. С. 292–307.
7. *Лемешко Б.Ю., Rogozhnikov А.П.* О нормальности погрешностей измерений в классических экспериментах и мощности критериев, применяемых для проверки отклонения от нормального закона // Метрология. 2012. № 5. С. 3–26.
8. *Кондрашкина М.И.* Влияние размера и формы образца на информацию о почвенном объекте: Автореферат дис. ... канд. биологических наук: 03.00.27. М., 1991. 24 с.
9. ПНД Ф 16.2.2:2.3.71-2011. Количественный химический анализ почв. Методика измерений массовых долей металлов в осадках сточных вод донных отложениях, образцах растительного происхождения спектральными методами.
10. *Hurlbert S.H.* Pseudoreplication and the design of ecological field experiments // Ecological Monographs. Vol. 54 (2). 1984. P. 187–211.
11. *Gurevitch J., Hedges L.V.* Meta-analysis. Combining the results of independent experiments // Design and analysis of ecological experiments / Schneider SM, Gurevitch J (eds). Oxford University Press. Oxford. USA. 2001. P. 347–369.

12. Oldfield E.E., Bradford M.A., Wood S.A. Global meta-analysis of the relationship between soil organic matter and crop yields // *Soil*. Vol. 5. Iss. 1. 2019. P. 15–32.
13. Webster R. Statistics to support soil research and their presentation // *European Journal of Soil Science*. 2001. Vol. 52. P. 331–340.

## REFERENCES

1. Blagoveshchenskii Yu.N., *Tainy korrelyatsionnykh svyazei v statistike* (Secrets of correlations in statistics), Moscow: INFRA-M, 2009, 158 p.
2. GOST R ISO 5725-6-2002 (State standard), Moscow: Standartinform, 2009, 58 p.
3. GOST R 50779.10-2000 (State standard), Moscow: Standartinform, 2005, 46 p.
4. Dmitriev E.A., *Matematicheskaya statistika v pochvovedenii* (Mathematical statistics in soil science), Moscow: MGU, 1995, 320 p.
5. Kobzar' A.I., *Prikladnaya matematicheskaya statistika* (Applied mathematical statistics), Moscow: Fizmatlit, 2006, 816 p.
6. Kozlov M.V., Mnimye povtornosti (pseudoreplication) v ekologicheskikh issledovaniyakh: problema, ne zamechennaya rossiiskimi uchenymi (Pseudoreplication in environmental research: a problem not noticed by Russian scientists), *Zhurnal obshchei biologii*, 2003, Vol. 64, No. 4, pp. 292–307.
7. Lemeshko B.Yu., Rogozhnikov A.P., O normal'nosti pogreshnosti izmerenii v klassicheskikh eksperimentakh i moshchnosti kriteriev, primenyaemykh dlya proverki otkloneniya ot normal'nogo zakona (On the normality of measurement errors in classical experiments and the power of criteria used to check deviations from the normal law), *Metrologiya*, 2012, No. 5, pp. 3–26.
8. Kondrashkina M.I., Vliyanie razmera i formy obraztsa na informatsiyu o pochvennom ob'ekte: *Avtoref. dis. ... kand. biol. nauk* (Effect of sample size and shape on soil object information, Extended abstract of cand. bio. sci. thesis), 03.00.27, Moscow: MGU, 1991, 24 p.
9. PND F 16.2.2:2.3.71-2011 (Federal Environmental Regulations), Moscow, 2011, 45 p.
10. Hurlbert S.H., Pseudoreplication and the design of ecological field experiments, *Ecological Monographs*, Vol. 54, No. 2, 1984, pp. 187–211.
11. Gurevitch J., Hedges L.V., Meta-analysis. Combining the results of independent experiments, *Design and analysis of ecological experiments*, Schneider S.M., Gurevitch J. (eds), Oxford University Press, 2001, pp. 347–369.

12. Oldfield E.E., Bradford M.A., Wood S.A., Global meta-analysis of the relationship between soil organic matter and crop yields, *Soil*, Vol. 5, Iss. 1, 2019, pp. 15–32.

13. Webster R., Statistics to support soil research and their presentation, *European Journal of Soil Science*, 2001, Vol. 52, pp. 331–340.