

Федеральная служба по надзору в сфере защиты  
прав потребителей и благополучия человека  
Федеральное казенное учреждение здравоохранения Иркутский ордена Трудового  
Красного Знамени научно-исследовательский противочумный институт Сибири и  
Дальнего Востока

**АЛГОРИТМ ДИФФЕРЕНЦИАЦИИ ТЕРРИТОРИЙ ПО РИСКУ ЗАБОЛЕВАНИЯ  
КЛЕЩЕВЫМ ВИРУСНЫМ ЭНЦЕФАЛИТОМ  
(на примере муниципальных образований Забайкальского края)**

**Учебно-методическое пособие**

**М.И. Толмачёва, А.Я. Никитин, А.О. Туранов, Ю.Н. Трушина**

**Иркутск, 2021**

**М.И. Толмачёва, А.Я. Никитин, А.О. Туранов, Ю.Н. Трушина**

**АЛГОРИТМ ДИФФЕРЕНЦИАЦИЯ ТЕРРИТОРИЙ ПО РИСКУ ЗАБОЛЕВАНИЯ  
КЛЕЩЕВЫМ ВИРУСНЫМ ЭНЦЕФАЛИТОМ  
(на примере муниципальных образований Забайкальского края)**

**Учебно-методическое пособие**

**ИРКУТСК – 2021**

УДК 616.988.833.1-036.8  
ББК 55.1  
А 45

Алгоритм дифференциации территорий по риску заболевания клещевым вирусным энцефалитом (на примере муниципальных образований Забайкальского края): учеб.-метод. пособие. / М.И. Толмачева, А.Я. Никитин, А.О. Туранов, Ю.Н. Трушина. – Иркутск: ИНЦХТ. - 20 с.

ISBN 978-5-98277-332-6

Учебно-методическое пособие предназначено для студентов, аспирантов высших учебных заведений, сотрудников Роспотребнадзора, проходящих профессиональную переподготовку или повышение квалификации по программам дополнительного профессионального образования в Иркутском научно-исследовательском противочумном институте.

Дифференциация территории по заболеваемости является основой разработки стратегии и тактики проведения профилактических мероприятий. Поэтому важно, чтобы специалисты Роспотребнадзора и других учреждений страны, могли научно-обоснованно общедоступными способами ранжировать территории отдельных административных районов в субъектах страны по риску заражения и проявлению природно-очаговых инфекций, правильно интерпретировать полученные результаты.

В пособии рассмотрен алгоритм проведения дифференциации территорий по уровню эпидемиологического риска с применением параметрических и непараметрических методов статистической обработки данных. В качестве примера взяты материалы о заболеваемости населения клещевым вирусным энцефалитом в административных районах Забайкальского края.

Статистические методы, изложенные в учебном пособии достаточно универсальны, просты в исполнении и находят широкое применение в научно-практических работах специалистов. Для освоения методов требуются знания математики в рамках программы средней школы. Пособие иллюстрировано примерами с демонстрацией возможности использования для их решения компьютерной программы Excel.

Печатается по решению учёного совета Иркутского научно-исследовательского противочумного института Роспотребнадзора (протокол № 2 от 20.02.2021)

© Толмачёва М.И., Никитин А.Я., Туранов А.О., Трушина Ю.Н.,  
2021  
© ИНЦХТ, 2021

## Содержание

1.	Введение	4
2.	Перечень сокращений, условных обозначений, используемых единиц и терминов	6
3.	Материалы, необходимые для проведения дифференциации территорий по эпидемиологическим рискам	7
4.	Характер представления данных о заболеваемости	7
5.	Оценка выпадающих значений	9
6.	Группировка административных районов субъекта РФ по эпидемиологическим рискам	10
6.1.	Группировка с применением параметрического метода расчета доверительного интервала	10
6.2.	Группировка с применением непараметрического (медианного) метода расчета доверительного интервала	15
7.	Список использованных источников	17
Приложение 1. Критические значения величины $t$ для исключения крайних вариантов («выпадающих» значений)		18
Приложение 2. Значения критерия Стьюдента ( $t_p$ ) при различных уровнях значимости и числе степеней свободы		19
Приложение 3. Границы непараметрического доверительного интервала (для медианы) при уровне значимости $P < 0,05$		20

## **ВВЕДЕНИЕ**

В настоящее время базовым подходом к профилактике природно-очаговых инфекций в РФ становится риск-ориентированный, основанный на дифференциации федеральных округов, субъектов и отдельных административных районов по инцидентности проявления болезней [Савилов и др., 2011; Ильин и др., 2013; Колпаков, Яковлев, 2015; Рудаков и др., 2019 а, б]. Подобный подход позволяет эффективно проводить мероприятия специфической и неспецифической профилактики, является экономически обоснованным, а при сравнении отдельных кластеров территорий с разной инцидентностью, может позволить выявить причины наблюдающихся различий и принять управленческие решения по улучшению эпидемиологической обстановки.

В ряде работ [Савилов и др., 2011; Колпаков, Яковлев, 2015; Ильин и др., 2013; Рудаков и др., 2019 б] продемонстрировано, что дифференциация территорий по уровню заболеваемости может проводиться с применением статистического метода, основанного на расчете доверительного интервала (ДИ) с применением параметрического (при нормальном распределении наблюдений) и непараметрического подходов.

Цель пособия – обучить алгоритму проведения дифференциации территорий по степени эпидемиологического риска. В качестве примера взяты данные о заболеваемости клещевым вирусным энцефалитом (КВЭ) в муниципальных образованиях (МО) Забайкальского края.

## 2. Перечень сокращений, условных обозначений, используемых единиц и терминов

**Г<sub>0</sub>–Г<sub>4</sub>** – обозначение групп территорий (кластеров), выделенных по уровню заболеваемости

**ДИ** – доверительный интервал, рассчитывается по имеющимся выборочным показателям заболеваемости, и с 95 % доверительной вероятностью покрывает область истинных, но не известных нам значений параметра, характеризующих «нормальную» для данных условий инцидентность проявления болезни

**КВЭ** – клещевой вирусный энцефалит

**Me** – медиана

**МО** – муниципальное образование

**РФ** – Российская Федерация

**СМП** – среднемноголетний показатель

**<sup>0</sup>/<sub>0000</sub>** – величина инцидентности КВЭ (или другой нозоформы) в пересчете на 100 тысяч населения

**df** – число степеней свободы

**k** – порядковый номер статистики выборки для определения границ ДИ

**M** – значение среднего

**m** – значение стандартной ошибки величины среднеарифметической

**n** – объем выборки

**P** – уровень достоверности (значимости)

**t** – величина критерия Стьюдента

**τ** – критерий оценки принадлежности крайних значений к исследуемой выборке

### **3. Материалы, необходимые для дифференцирования территорий**

В качестве примера для статистического анализа и дифференциации (ранжирования) территорий по уровню заболеваемости, использованы данные об инцидентности КВЭ в МО Забайкальского края из работы А.О. Туранова и соавторов [Туранов и др., 2020].

Для выявления групп (кластеров) административных районов, значительно различающихся по СМП заболеваемости, использован расчет параметрического и непараметрического (основанного на медиане) 95 % ДИ [Закс, 1976; Савилов и др., 2011; Ильин и др., 2013]. Кроме того, применен непараметрический метод оценки принадлежности отдельных визуально выпадающих значений к исследуемой совокупности [Терентьев, Ростова, 1977].

### **4. Характер представления данных о заболеваемости**

Анализ данных следует начинать с ввода ранжированных (сортированных по убыванию) исходных значений в электронную таблицу Excel. Обычно по горизонтали (верхняя строка) вводятся наименования анализируемых показателей, по вертикали (столбцы) – их количественные значения (рис. 1). Возможно и иное представление исходных данных. Если введены не ранжированные данные, то эту операцию можно выполнить непосредственно в программе Excel: на вкладке «Главная» входим в окно «Сортировка и фильтр» и активируем функцию «Сортировка от Я до А». То есть, в нашем случае (рис. 1) ранжированные показатели заболеваемости КВЭ расположены от больших значений к меньшим.

Важно подчеркнуть, что для проведения дифференциации территорий по степени эпидемиологического риска могут быть использованы абсолютные [Ильин и др., 2013] или относительные показатели [Колпаков, Яковлев, 2015; Рудаков и др., 2019 б]. В случае нормального распределения данных в исследуемой выборке для ее анализа применяют параметрический статистический метод (раздел 6.1) [Закс, 1976; Савилов и др., 2011; Ильин и др., 2013; Колпаков, Яковлев, 2015], в противном случае – непараметрический (раздел 6.2) [Закс, 1976; Колпаков, Яковлев, 2015; Рудаков и др., 2019 б].

В качестве примера на рис. 1 приведены относительные показатели заболеваемости КВЭ в МО Забайкальского края.

Забайкальский край включает 32 административных района. За анализируемый период восемь МО являлись эндемичными (они объединены в группу, которая

обозначена, как  $\Gamma_0$ ). Все дальнейшие расчеты произведены по отношению к оставшимся 24 МО.

	A	B	C	D
1				
2			<b>Административный район</b>	<b>СМП</b>
3		1	Красночикийский	36
4		2	Петровск-Забайкальский	24,6
5		3	Шелопутинский	13,3
6		4	Улетовский	10,5
7		5	Газимуро-Заводский	9,2
8		6	Тунгокоченский	7,4
9		7	Балейский	6,7
10		8	Александрово-Заводский	5,1
11		9	Оловянинский	4,9
12		10	Могочинский	4,8
13		11	Сретенский	4,7
14		12	Нерчинский	4,4
15		13	Читинский	4,3
16		14	Шилкинский	3,4
17		15	Хилокский	2,8
18		16	Дульдургинский	2,5
19		17	Акшинский	2,2
20		18	г. Чита	2,2
21		19	Карымский	2,2
22		20	Калганский	2
23		21	Чернышевский	1,5
24		22	Агинский	1,4
25		23	Могойтуйский	1,2
26		24	Борзинский	1
27		25	Каларский	0
28		26	Краснокаменский	0
29		27	Кыринский	0
30		28	Приаргунский	0
31		29	Нерченско-Заводский	0
32		30	Тунгиро-Олекминский	0
33		31	Забайкальский	0
34		32	Ононский	0
35				

Рис. 1. Вариант ввода исходных данных в таблицу Excel на примере СМП заболеваемости населения КВЭ в Забайкальском крае (2003–2018 гг.).



## 5. Оценка выпадающих значений

Особенностью пространственного проявления КВЭ на территории Забайкальского края является наличие двух МО (Петровск-Забайкальский и Красночикойский) с инцидентностью болезни значительно более высокой, чем в остальных районах этого субъекта РФ (рис. 1). В этой связи, прежде, чем перейти к группировке территорий по уровню заболеваемости в отдельные кластеры на основе метода ДИ, необходимо оценить принадлежность двух «выпадающих» значений исследуемой совокупности. Для этого применим непараметрический метод, основанный на расчете показателя  $\tau$  [Терентьев, Ростова, 1977].

Критерий  $\tau$  рассчитывается по формуле (1), если сомнительным является одно минимальное по величине значение. В случае, если сомнительна принадлежность к выборке максимального по величине значения, используется формула (2).

Если «сомнительными» являются сразу два крайних значения (как в нашем примере) и оба находятся в начале ранжированного ряда величин, применяется формула (3), а если в конце – (4).

(1) для минимальной варианты:  $\tau_{1(1)} = \frac{x_2 - x_1}{x_n - x_1}$ , где  $x_1$  – первое значение в ранжированной выборке;  $x_2$  – второе значение;  $x_n$  – последнее значение.

(2) для максимальной варианты:  $\tau_{n(1)} = \frac{x_n - x_{n-1}}{x_n - x_1}$ .

(3) для двух «выпадающих» значений ( $x_1$  и  $x_2$ ) в начале ранжированного ряда:  
$$\tau_{1(2)} = \frac{x_2 - x_1}{x_n - x_1}$$

(4) для двух «выпадающих» значений ( $x_n$  и  $x_{n-1}$ ) в конце ранжированного ряда:  
$$\tau_{n(2)} = \frac{x_n - x_{n-1}}{x_n - x_1}$$

Если сомнительными являются одновременно и минимальное, и максимальное значения, то  $\tau$ , рассчитывают по формулам (5) и (6):

(5)  $\tau_{1(2)} = \frac{x_2 - x_1}{x_{n-1} - x_1}$  (оценка принадлежности к выборке  $x_1$ )

(6)  $\tau_{n(2)} = \frac{x_n - x_{n-1}}{x_n - x_2}$  (оценка принадлежности к выборке  $x_n$ )

После расчета величины  $\tau$ , полученное значение сравнивается с критическим по Приложению 1. Если расчетное значение  $\tau$  больше или равно значению, приведенному в

Приложении 1, то оцениваемое «выпадающее» значение должно быть исключено из дальнейшего анализа (с уровнем значимости  $P < 0,05$  или  $P < 0,001$  см. Приложение 1).

В нашем примере мы используем критерий  $\tau$  для оценки принадлежности к выборке двух крайних значений (рис. 1), которые находятся в одном и том же конце (максимальные значения) ранжированного ряда (административные районы с инцидентностью болезни  $36,0 \text{ }^0/_{0000}$  и  $24,6 \text{ }^0/_{0000}$ ). Рассматриваем только эндемичные по КВЭ территории. То есть, проверка будет осуществляться для  $n = 25$  (ближайшая величина  $n$  в Приложении 1 к фактическому значению наблюдений в анализируемой нами выборке). Расчет производится по формуле:

$$\tau_{n(2)} = \frac{x_n - x_{n-2}}{x_n - x_1} \text{ (для оценки принадлежности к выборке } x_n \text{ и } x_{n-1}) =$$
$$= (36,0 - 13,3) / (36,0 - 1,0) = 0,649$$

Полученное значение  $\tau$  сверяем с таблицей критических значений показателя (Приложение 1).

Критическое значение  $\tau$  при  $P < 0,05$  и  $n = 25$  равно  $0,347$ ; а при  $P < 0,01$  и  $n = 25 = 0,434$  (Приложение 1). Полученное значение  $\tau$  ( $0,649$ ) выше, чем критические по Приложению 1. Следовательно, Красночикийский и Петровск-Забайкальский МО по инцидентности КВЭ с высоким уровнем достоверности ( $P < 0,01$ ) не принадлежат к группе, в которую входят все остальные эндемичные МО. Таким образом, статистически обосновано, что два анализируемых МО формируют самостоятельный кластер ( $\Gamma_4$ ) с очень высокой инцидентностью КВЭ, который необходимо изучать отдельно от всех остальных групп административных районов.

## **6. Группировка административных районов субъекта России по эпидемиологическим рискам**

Продолжим рассмотрение примера с дифференцированием по уровню заболеваемости КВЭ МО Забайкальского края на основе материалов за 2003–2018 гг. Единицами наблюдения в нашем случае являются территории МО.

Все МО Забайкальского края по СМП заболеваемости КВЭ разделены на пять групп: (Г<sub>0</sub>-Г<sub>4</sub>).

Г<sub>0</sub> кластер, включающий неэндемичные по КВЭ МО. В него вошли 8 районов (рис. 1).

В группу Г<sub>4</sub> – с самой высокой заболеваемостью КВЭ (раздел 5), вошли два района, которые характеризуются «выпадающими» по величине высокими уровнями заболеваемости КВЭ: Красночикойский и Петровск-Забайкальский.

Для группировки по эпидемиологическим рискам остальных 22 МО применили вариант дифференциации (ранжирования) территорий, основанный на оценке величины доверительного интервала (ДИ) (алгоритмы расчета параметрического ДИ приведены в разделе 6.1, а непараметрического в 6.2).

В результате расчета величины ДИ формируется еще три группы (кластера) МО по уровню проявления КВЭ: Г<sub>1</sub> – кластер МО с СМП заболеваемости ниже величины минимального значения ДИ; Г<sub>2</sub> – группа МО с СМП заболеваемости, который находится в границах ДИ; Г<sub>3</sub> – группа МО с СМП заболеваемости выше верхней границы ДИ.

Следует отметить, что если в нашем примере выделено пять групп по уровню эпидемиологического риска проявления КВЭ, то в других субъектах или за другой период времени, число кластеров может быть иным. Например, отсутствовать группы Г<sub>0</sub> и/или Г<sub>4</sub>.

### **6.1. Группировка с применением параметрического метода расчета доверительного интервала**

Все дальнейшие расчеты, в соответствии с приводимым ниже алгоритмом, проведены в программе Excel, имеющейся во всех компьютерах, и не требуют для освоения специальных знаний. Эта программа широко используется специалистами Роспотребнадзора. Большинство наших расчетов проведено во вкладке «Анализ данных», которая, в свою очередь, расположена во вкладке «Данные» на панели Excel (рис. 2). Если модуль «Анализ данных» отсутствует во вкладке «Данные», так как ранее на этом компьютере не использовался, то его необходимо установить. Для этого после входа в пункт меню «Настройки» диалогового окна «Сервис» достаточно активировать блок команд «Пакет анализа».

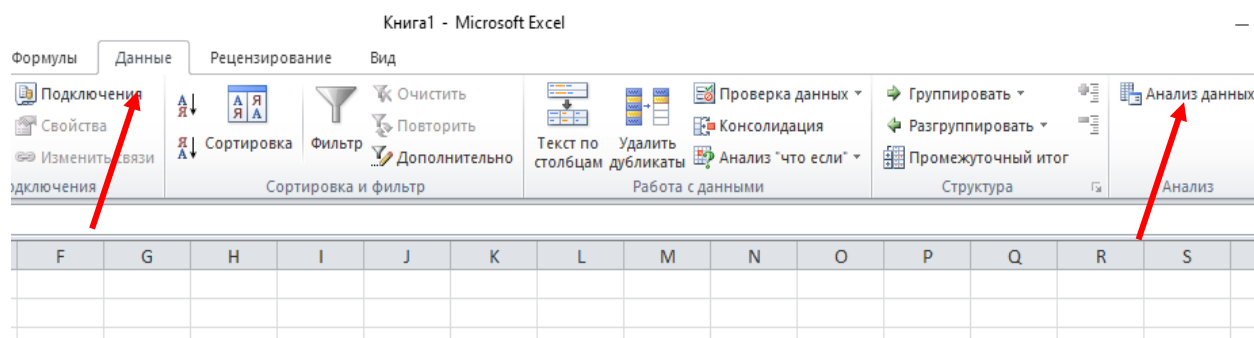


Рис. 2. Расположение вкладок «Данные» и «Анализ данных» (на них направлены стрелки) на панели Excel

Рассмотрим пример дифференциации территорий по уровню эпидемиологического риска с применением параметрического метода расчета ДИ. Для этого необходимо курсором выделить вставку «Пакет анализа», и дать команду «ввод». В активированном контекстном меню выбираем пункт «Описательная статистика» (рис. 3).

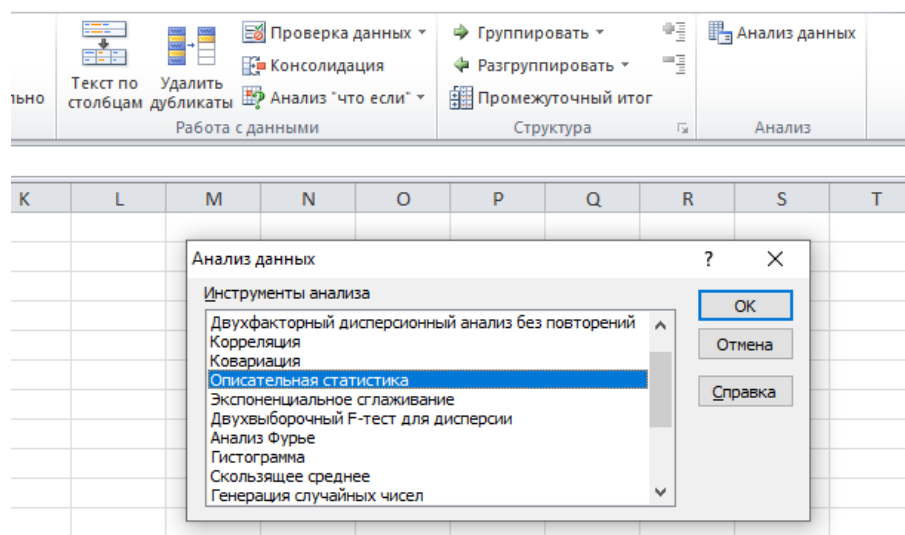


Рис. 3. Выбор функции описательная статистика для обработки значений

Далее действуем в соответствии с подсказками вкладок настроек опции «Описательная статистика» (рис. 4). То есть, в окно «Входной интервал» вводим данные о СМП заболеваемости КВЭ в 22 административных районах Забайкальского края.

Во вкладке настроек «Описательной статистики» выделяем курсором в меню «Параметры вывода» иконку «Выходной интервал». Обязательно отмечаем галочкой пункт «Итоговая статистика» (рис. 4). Затем активируем курсором одну из ячеек вблизи нашего массива данных. Нажимаем «ОК» и получаем итоги расчета в виде таблицы (рис. 5).

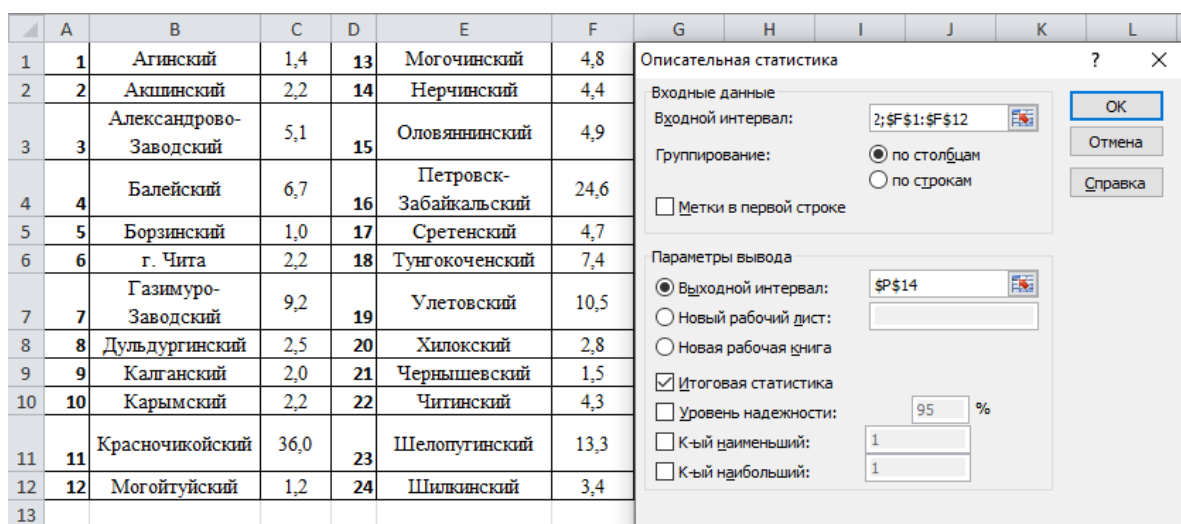


Рис. 4. Блок настроек меню опции «Описательная статистика»

	A	B	C	D	E	F	G	H	I
1	1	Агинский	1,4	13	Могочинский	4,8		Столбец1	
2	2	Акшинский	2,2	14	Нерчинский	4,4		Среднее	6,595833
3	3	Александрово-Заводский	5,1	15	Оловянинский	4,9		Стандартная ошибка	1,65519
4	4	Балейский	6,7	16	Петровск-Забайкальский	24,6		Медиана	4,35
5	5	Борзинский	1,0	17	Сретенский	4,7		Мода	2,2
6	6	г. Чита	2,2	18	Тунгокоченский	7,4		Стандартное отклонение	8,108743
7	7	Газимуро-Заводский	9,2	19	Улетовский	10,5		Дисперсия выборки	65,75172
8	8	Дульдургинский	2,5	20	Хилокский	2,8		Эксцесс	7,802283
9	9	Калганский	2,0	21	Чернышевский	1,5		Асимметричность	2,711683
10	10	Карымский	2,2	22	Читинский	4,3		Интервал	35
11	11	Красночикойский	36,0	23	Шелопутинский	13,3		Минимум	1
12	12	Могойтуйский	1,2	24	Шилкинский	3,4		Максимум	36
13								Сумма	158,3
14								Счет	24

Рис. 5. Вывод итогов расчетов опции «Описательная статистика» на экран монитора

Для расчета границ ДИ из таблицы рис. 5 нам понадобятся только три строки: среднее, стандартная ошибка и счет.

Расчет верхней и нижней границ ДИ проводится по формулам:

$$M - t \times m < \text{ДИ} < M + t \times m,$$

где  $M$  – значение среднего (рис. 5:  $M$  округленно равно 6,60);  $t$  – величина критерия Стьюдента по Приложению 2;  $m$  – значение стандартной ошибки (рис. 5:  $m$  округленно равно 1,655). Обычно при округлении значение  $m$  дается на один знак больше, чем округленное значение  $M$ .

Таким образом, нижняя граница ДИ равна:  $M - t \times m$ ; а верхняя:  $M + t \times m$ .

В медицине и биологии часто используют ДИ равный 95 % вероятности проявления события. Это означает, что при повторных исследованиях (испытаниях) лишь в 5 % случаев

(или в 0,05 в долях от единицы), новое полученное значение уровня заболеваемости окажется за пределами границ ДИ.

Значение  $t$  находим по таблицам критерия Стьюдента (Приложение 2). При этом  $t$  берется на пересечении выбранного значения  $P$  (берем 95 % = 0,05 в таблице Приложения 2) и величины  $df$  (числа степеней свободы) равного в общем случае  $n-1$ . То есть, в нашем примере при расчете ДИ для 22 эндемичных по КВЭ МО Забайкальского края:

$df = 22-1 = 21$ ;  $t = 2,08$ ; нижняя граница ДИ:  $6,6 - (1,66*2,08) = 3,2$ ; верхняя граница ДИ:  $6,6 + (1,66*2,08) = 10,0$ .

Соответственно, верхний порог заболеваемости КВЭ в приводимом примере составит  $10,0 \text{ ‰}$ , а нижний –  $3,2 \text{ ‰}$ . Значения между  $3,2 \text{ ‰}$  и  $10,0 \text{ ‰}$  входят в границы 95 % ДИ и формируют кластер  $\Gamma_2$  (раздел 6).

Расчет ДИ можно провести в программе Excel через выбор функции «ДОВЕРИТ.СТЬЮДЕНТ», где необходимо предварительно рассчитать стандартное отклонение по генеральной совокупности (рис. 6).

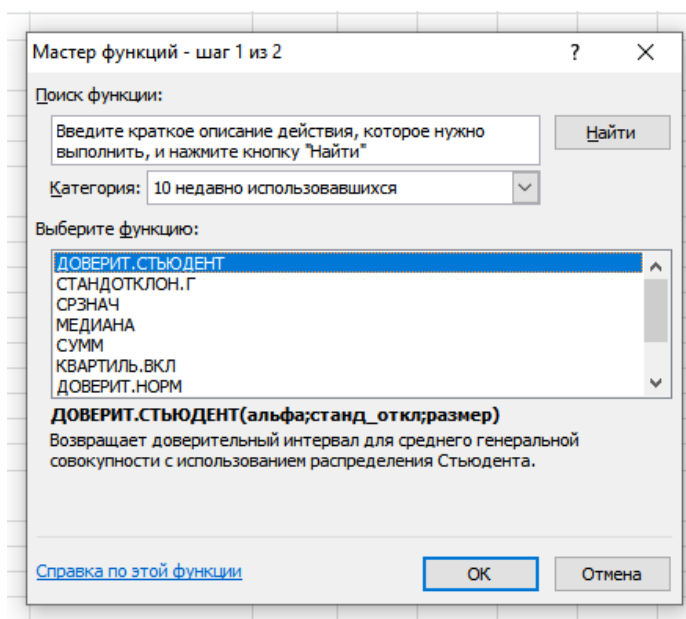


Рис. 6. Окно «мастер функций» для выбора функции расчета стандартного отклонения и ДИ

После выделения кластера с очень высокой инцидентностью КВЭ (раздел 5) и неэндемичных районов ( $\Gamma_0$ ), СМП заболеваемости в остальных 22 эндемичных административных районах Забайкальского края с применением метода расчета 95 % ДИ разбиты на три группы: с низкой ( $\Gamma_1$ ), средней ( $\Gamma_2$ ) и высокой ( $\Gamma_3$ ) заболеваемостью.

Кластер с низкой заболеваемостью КВЭ ( $\Gamma_1$  – равны или ниже нижней границы ДИ) включает десять МО. По отдельным МО этого кластера заболеваемость варьирует от  $1,0 \text{ ‰}$  (Борзинский район) до  $2,8 \text{ ‰}$  (Хилокский район).

Для кластера со средним уровнем заболеваемости КВЭ ( $\Gamma_2$  – десять районов) наблюдается вариация СМП от 3,4 ‰ (Шилкинский район) до 9,2 ‰ (Газимуро-Заводский район). Вся эта группа укладывается в границы ДИ.

И, наконец, в группу с высокой заболеваемостью КВЭ ( $\Gamma_3$  – включает МО со значениями равными или выше верхней границы ДИ) входят три МО при вариации инцидентности между территориями от 10,5 ‰ (Улетовский район) до 13,3 ‰ (Шелопугинский район).

## 6.2. Группировка с применением непараметрического (медианного) метода расчета доверительного интервала

В качестве инструмента градации шкалы с применением непараметрического подхода [Закс, 1976; Колпаков, Яковлев, 2015; Рудаков и др., 2019 б] применим расчет количества элементов выборки, находящихся между доверительными границами медианы (**Me**). Алгоритм применения метода медианы, начнем с ранжирования (упорядочивания) всех значений выборки от меньшего к большему (пример на рис. 7).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Новосибирская область															
2	Заболеваемость населения КВЭ: числослучаев на 100 тысяч населения															
3	2016	10	4,2													
4	2008	2	4,8													
5	2014	8	5,6													
6	2007	1	5,8													
7	2012	6	6,1													
8	2015	9	6,2													
9	2010	4	6,4													
10	2009	3	6,5													
11	2011	5	7,4													
12	2013	7	7,9													

Рис. 7. Ранжирование значений выборки. Стрелка указывает на иконку «Сортировка и фильтр», которая содержит команды упорядочивания значений

Если выборка содержит нечетное число наблюдений, то медианой будет являться значение, находящееся в середине исследуемой выборки. Если число четное, то величина медианы рассчитывается путем усреднения двух соседних центральных значений.

Еще проще рассчитать **Me** с применением программы Excel. Для этого необходимо выбрать из перечня статистических функций вкладку «МЕДИАНА» и нажать ОК (рис. 8).

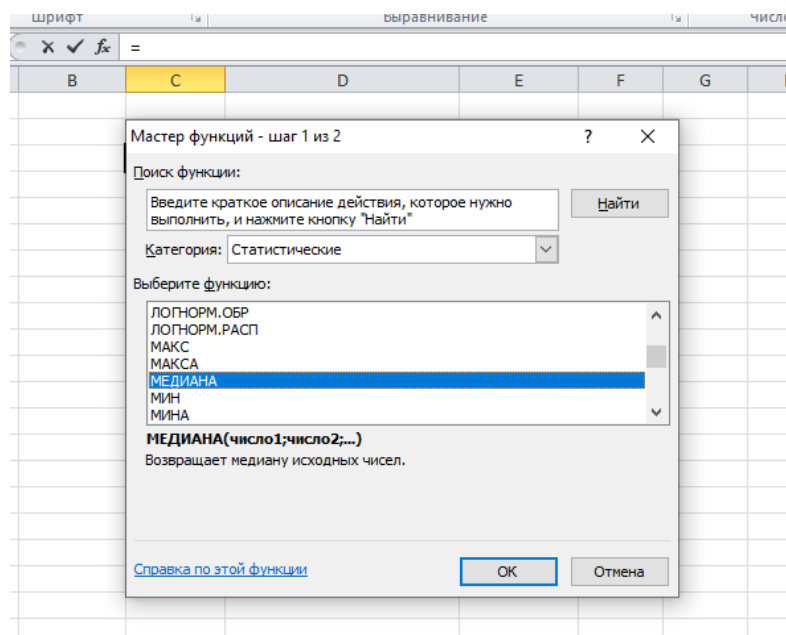


Рис. 8. Расчет значения медианы

**Me**, в нашем примере, составляет 3,85 (значения совпали по обоим методам вычисления).

При расчете пороговых значений ДИ медианы для выборок, содержащих меньше 100 наблюдений, можно воспользоваться таблицей Приложения 3. В ней даны значения номеров левой (нижней) и правой (верхней) границы ДИ. Зная эти номера, можно найти конкретные значения возможной вариации признака по исходной последовательности значений заболеваемости (рис. 1).

В нашем примере  $n = 22$ . Для этого случая нижней границей 95 % ДИ для **Me** будет значение с порядковым номером 6, а верхней – 17 (Приложение 3). Этим порядковым номерам соответствуют (рис. 1) значения уровня заболеваемости КВЭ для случае нижнего порога – 2,0 ‰, верхнего – 5,1 ‰.

Число групп, полученных в ходе дифференцирования территорий по степени эпидемиологического риска методом расчета непараметрического ДИ, результате будет равно 5. Так как нами изначально выделена группа  $\Gamma_0$  – неэндемичные территории и  $\Gamma_4$  – МО с очень высокими уровнями показателя заболеваемости КВЭ (разделы 4 и 5).

Группа со средним уровнем заболеваемости ( $\Gamma_2$  – включает двенадцать районов), формируется из МО, СМП инцидентности КВЭ которых находится между нижней и верхней границами 95 % ДИ, найденного по методу медианы. Соответственно, заболеваемость в них варьирует от 2,2 ‰ (Карымский район) до 5,1 ‰ (Александрово-Заводский район).

Кластер с низким уровнем СМП КВЭ ( $\Gamma_1$  – включает пять МО) и объединяет территории со значениями заболеваемости меньше нижней границы **Me**. По отдельным



МО этого кластера заболеваемость варьирует от 1,0 ‰ (Борзинский район) до 2,0 ‰ (Калганский район).

Значения, превышающие верхнюю доверительную границу **Me**, образуют группу Г<sub>3</sub> (включает пять МО) с высокой заболеваемостью КВЭ. Инцидентность КВЭ в них варьирует от 6,7 ‰ (Балейский район) до 13,3 ‰ (Шелопугинский район).

## 7. Список использованных источников

1. Закс Л. Статистическое оценивание. Пер. с нем. В.Н. Варыгина. Под ред. Ю.П. Адлера, В.Г. Горского. М., «Статистика», 1976. – 598 с.
2. Ильин В.П. и др. Алгоритм классификации муниципальных образований субъекта Российской Федерации, основанный на оценке доверительного интервала заболеваемости клещевым вирусным энцефалитом (на примере Забайкальского края). Бюллетень ВСНЦ СО РАМН. 2013; 6: 115–121.
3. Колпаков С.Л., Яковлев А.А. О методологии оценки эпидемиологической ситуации. Эпидемиология и инфекционные болезни. 2015; (20); 4: 34–39.
4. Рудакова С.А. и др. (а). Интенсивность и тенденции развития эпидемического процесса иксодовых клещевых боррелиозов в Российской Федерации в 2002–2018 гг. и прогноз на 2019 г. Проблемы особо опасных инфекций. 2019; (24); 2: 22–29.
5. Рудаков Н.В. и др. (б). Дифференциация эндемичных территорий по уровням заболеваемости клещевыми трансмиссивными инфекциями как основа выбора стратегии и тактики профилактики. Здоровье населения и среда обитания. 2019; 12: 56–61.
6. Савилов Е.Д. и др. Эпидемиологический анализ: Методы статистической обработки материала. Новосибирск: Наука-Центр, 2011. – 156 с.
7. Терентьев П.В., Ростова Н.С. Практикум по биометрии. Учебное пособие. Л.: Изд-во Ленинградского университета, 1977. – 152 с.
9. Туранов А.О. и др. Дифференциация территории Забайкальского края по заболеваемости клещевым вирусным энцефалитом. Проблемы особо опасных инфекций. 2020; 2:108–114. DOI: 10.21055/0370-1069-2020-2-108-114

Критические значения величины  $\tau$  для исключения крайних вариантов («выпадающих» значений)

n (число наблюдений)	$\tau$ для оценки возможности исключить «выпадающее» значение ( $\tau_{(1)}$ )		$\tau$ для оценки возможности исключить два «выпадающих» значения ( $\tau_{(2)}$ )	
	P = 0,05	P = 0,01	P = 0,05	P = 0,01
3	0,941	0,988	1,000	1,000
4	0,765	0,889	0,967	0,992
5	0,642	0,780	0,845	0,929
6	0,560	0,698	0,736	0,836
7	0,507	0,637	0,661	0,778
8	0,468	0,590	0,607	0,710
9	0,437	0,555	0,565	0,667
10	0,412	0,527	0,531	0,632
11	0,392	0,592	0,504	0,603
12	0,376	0,482	0,481	0,579
15	0,338	0,438	0,430	0,522
20	0,300	0,391	0,372	0,464
25	0,281	0,367	0,347	0,434
30	0,260	0,341	0,322	0,402

Значения критерия Стьюдента ( $t_p$ ) при различном числе степеней свободы ( $df$ ) и уровне значимости ( $P$ )

Число степеней свободы ( $df$ )	Уровни значимости ( $P$ )			
	0,1	0,05	0,01	0,001
1	6,31	12,71	63,66	636,62
2	2,92	4,30	9,93	31,60
3	2,35	3,18	5,84	12,92
4	2,13	2,78	4,60	8,61
5	2,02	2,57	4,03	6,87
6	1,94	2,45	3,71	5,96
7	1,90	2,37	3,50	5,41
8	1,86	2,31	3,36	5,04
9	1,83	2,26	3,25	4,78
10	1,81	2,23	3,17	4,59
11	1,80	2,20	3,11	4,44
12	1,78	2,18	3,06	4,32
13	1,77	2,16	3,01	4,22
14	1,76	2,15	2,98	4,14
15	1,75	2,13	2,95	4,07
16	1,75	2,12	2,92	4,02
17	1,74	2,11	2,90	3,97
18	1,73	2,10	2,88	3,92
19	1,73	2,09	2,86	3,88
20	1,73	2,09	2,85	3,85
21	1,72	2,08	2,83	3,82
22	1,72	2,07	2,82	3,79
23	1,71	2,07	2,81	3,77
24	1,71	2,06	2,80	3,75
25	1,71	2,06	2,79	3,73
26	1,71	2,06	2,78	3,71
27	1,70	2,05	2,77	3,69
28	1,70	2,05	2,76	3,67
29	1,70	2,05	2,76	3,66
30	1,70	2,04	2,75	3,65
35	1,69	2,03	2,72	3,59
40	1,68	2,02	2,70	3,55
60	1,67	2,00	2,66	3,46
100	1,66	1,98	2,63	3,39
> 100	1,65	1,96	2,58	3,29

Границы непараметрического доверительного интервала (для медианы – **Me**)  
при уровне значимости **P < 0,05**)

<b>n</b> (число наблюдений)	Порядковые номера наблюдений для определения границ ДИ		<b>n</b> (число наблюдений)	Порядковые номера наблюдений для определения границ ДИ		<b>n</b>	Порядковые номера наблюдений для определения границ ДИ	
	Нижняя	Верхняя		Нижняя	Верхняя		Нижняя	Верхняя
<b>5</b>	0	6	<b>37</b>	13	25	<b>69</b>	26	44
<b>6</b>	1	6	<b>38</b>	13	26	<b>70</b>	27	44
<b>7</b>	1	7	<b>39</b>	13	27	<b>71</b>	27	45
<b>8</b>	1	8	<b>40</b>	14	27	<b>72</b>	28	45
<b>9</b>	2	8	<b>41</b>	14	28	<b>73</b>	28	46
<b>10</b>	2	9	<b>42</b>	15	28	<b>74</b>	29	46
<b>11</b>	2	10	<b>43</b>	15	29	<b>75</b>	29	47
<b>12</b>	3	10	<b>44</b>	16	29	<b>76</b>	29	48
<b>13</b>	3	11	<b>45</b>	16	30	<b>77</b>	30	48
<b>14</b>	3	12	<b>46</b>	16	31	<b>78</b>	30	49
<b>15</b>	4	12	<b>47</b>	17	31	<b>79</b>	31	49
<b>16</b>	4	13	<b>48</b>	17	32	<b>80</b>	31	50
<b>17</b>	5	13	<b>49</b>	18	32	<b>81</b>	32	50
<b>18</b>	5	14	<b>50</b>	18	33	<b>82</b>	32	51
<b>19</b>	5	15	<b>51</b>	19	33	<b>83</b>	33	51
<b>20</b>	6	15	<b>52</b>	19	34	<b>84</b>	33	52
<b>21</b>	6	16	<b>53</b>	19	35	<b>85</b>	33	53
<b>22</b>	6	17	<b>54</b>	20	35	<b>86</b>	34	53
<b>23</b>	7	17	<b>55</b>	20	36	<b>87</b>	34	54
<b>24</b>	7	18	<b>56</b>	21	36	<b>88</b>	35	54
<b>25</b>	8	18	<b>57</b>	21	37	<b>89</b>	35	55
<b>26</b>	8	19	<b>58</b>	22	37	<b>90</b>	36	55
<b>27</b>	8	20	<b>59</b>	22	38	<b>91</b>	36	56
<b>28</b>	9	20	<b>60</b>	22	39	<b>92</b>	37	56
<b>29</b>	9	21	<b>61</b>	23	39	<b>93</b>	37	57
<b>30</b>	10	21	<b>62</b>	23	40	<b>94</b>	38	57
<b>31</b>	10	22	<b>63</b>	24	40	<b>95</b>	38	58
<b>32</b>	10	23	<b>64</b>	24	41	<b>96</b>	38	59
<b>33</b>	11	23	<b>65</b>	25	41	<b>97</b>	39	59
<b>34</b>	11	24	<b>66</b>	25	42	<b>98</b>	39	60
<b>35</b>	12	24	<b>67</b>	26	42	<b>99</b>	40	60
<b>36</b>	12	25	<b>68</b>	26	43	<b>100</b>	40	61