

ОСНОВЫ СТАТИСТИКИ

Визуализации статистических
данных в MS Excel





Здравствуйтесь!

В этой лекции, вы ознакомитесь с достаточно интересным инструментом, позволяющим автоматически рассчитать элементы описательной статистики, а также рассмотрим процедуру вычисления доверительного интервала.

Наиболее развитыми средствами анализа данных являются статистические процедуры «Пакета анализа». Они обладают большими возможностями, чем статистические функции. С их помощью можно решать более сложные задачи обработки статистических данных и выполнять более тонкий анализ этих данных.

В «Пакет анализа» входят следующие статистические процедуры:

- генерация случайных чисел (Random number generation);
- выборка (Sampling);
- гистограмма (Histogram);
- описательная статистика (Descriptive statistics);
- ранги перцентиль (Rank and percentile);
- двухвыборочный z-тест для средних (z-Test: Two Sample for Means);
- двухвыборочный t-тест для средних с одинаковыми дисперсиями (t-Test: Two-Sample Assuming Equal Variances);
- двухвыборочный t-тест для средних с различными дисперсиями (t-Test: Two-Sample Assuming Unequal Variances);
- парный двухвыборочный t-тест для средних (t-Test: Paired Two Sample for Means);
- двухвыборочный F-тест для дисперсий (F-Test: Two Sample for Variances);
- ковариация (Covariance);
- корреляция (Correlation);
- регрессия (Regression);
- однофакторный дисперсионный анализ (ANOVA: Single Factor);
- двухфакторный дисперсионный анализ без повторений (ANOVA: Two Factor Without Replication);
- двухфакторный дисперсионный анализ с повторениями (ANOVA: Two Factor With Replication);
- скользящее среднее (Moving Average);
- экспоненциальное сглаживание (Exponential Smoothing);
- анализ Фурье (Fourier Analysis).

В данном случае мы имеем не статистические функции, а более удобный инструмент, это «Описательная статистика».

Пользователи MS Excel знают, что данная программа имеет очень широкий набор статистических функций, по уровню которых она вполне может потягаться со специализированными приложениями. Кроме того, у MS Excel имеется инструмент, с помощью которого производится обработка данных по целому ряду основных статистических показателей буквально в один клик.

Статистические процедуры Пакета анализа

Этот инструмент называется «Описательная статистика». С его помощью можно в очень короткие сроки использовав ресурсы программы, обработать массив данных и получить о нем информацию по целому ряду статистических критериев. Давайте взглянем, как работает данный инструмент, и остановимся на некоторых нюансах работы с ним.



Использование описательной статистики.

Под описательной статистикой понимают систематизацию эмпирических данных по целому ряду основных статистических критериев. Причем на основе полученного результата из этих итоговых показателей можно сформировать общие выводы об изучаемом массиве данных.

В MS Excel существует отдельный инструмент, входящий в «Пакет анализа», с помощью которого можно провести данный вид обработки данных. Он так и называется «Описательная статистика».

Среди критериев, которые высчитывает данный инструмент следующие показатели:

- Медиана;
- Мода;
- Дисперсия;
- Среднее;
- Стандартное отклонение;
- Стандартная ошибка;
- Асимметричность и др.

Как уже было сказано выше, инструмент «Описательная статистика» входит в более широкий набор функций, который принято называть «Пакет анализа». Но дело в том, что по умолчанию данная надстройка в MS Excel отключена. Поэтому, если вы до сих пор её не включили, то для использования возможностей описательной статистики, придется это сделать.

Переходим во вкладку «Файл». Далее производим перемещение в пункт «Параметры».

В активированном окне параметров перемещаемся в подраздел «Надстройки». В самой нижней части окна находится поле «Управление». Нужно в нем переставить переключатель в позицию «Надстройки MS Excel», если он находится в другом положении. Вслед за этим кликаем на кнопку «Перейти...».

Запускается окно стандартных надстроек MS Excel. Около наименования «Пакет анализа» ставим флажок. Затем кликаем на кнопку «ОК».

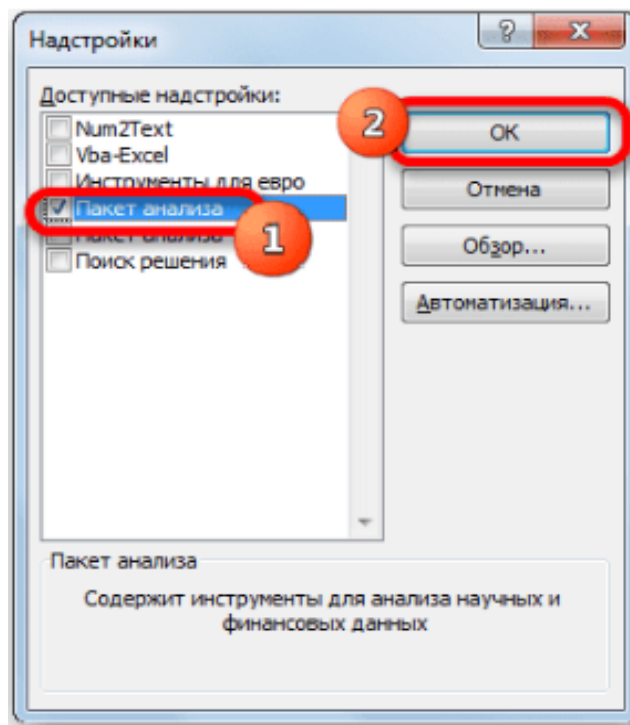


Рисунок 15.1. Окно настройки «Пакет анализ»



После вышеуказанных действий надстройка «Пакет анализа» будет активирована и станет доступной во вкладке «Данные» MS Excel. Теперь мы сможем использовать на практике инструменты описательной статистики.

Применение инструмента «Описательная статистика»

Теперь посмотрим, как инструмент описательная статистика можно применить на практике. Для этих целей используем готовую таблицу.

Переходим во вкладку «Данные» и выполняем клик по кнопке «Анализ данных», которая размещена на ленте в блоке инструментов «Анализ».

Открывается список инструментов, представленных в Пакете анализа. Ищем наименование «Описательная статистика», выделяем его и кликаем по кнопке «ОК».

После выполнения данных действий непосредственно запускается окно «Описательная статистика».

В поле «Входной интервал» указываем адрес диапазона, который будет подвергаться обработке этим инструментом. Причем указываем его вместе с шапкой таблицы. Для того, чтобы внести нужные нам координаты, устанавливаем курсор в указанное поле. Затем, зажав левую кнопку мыши, выделяем на листе соответствующую табличную область. Как видим, её координаты тут же отобразятся в поле. Так как мы захватили данные вместе с шапкой, то около параметра «Метки в первой строке» следует установить флажок. Тут же выбираем тип группирования, переставив переключатель в позицию «По столбцам» или «По строкам». В нашем случае подходит вариант «По столбцам», но в других случаях, возможно, придется выставить переключатель иначе.

Выше мы говорили исключительно о входных данных. Теперь переходим к разбору настроек параметров вывода, которые расположены в этом же окне формирования описательной статистики. Прежде всего, нам нужно определиться, куда именно будут выводиться обработанные данные:

- Выходной интервал;
- Новый рабочий лист;
- Новая рабочая книга.

В первом случае нужно указать конкретный диапазон на текущем листе или его верхнюю левую ячейку, куда будет выводиться обработанная информация. Во втором случае следует указать название конкретного листа данной книги, где будет отображаться результат обработки. Если листа с таким наименованием в данный момент нет, то он будет создан автоматически после того, как вы кликнете на кнопку «ОК». В третьем случае никаких дополнительных параметров указывать не нужно, так как данные будут выводиться в отдельном файле MS Excel (книге). Мы выбираем вывод результатов на новом рабочем листе под названием «Итоги».

Далее, если вы хотите, чтобы выводилась также итоговая статистика, то нужно установить флажок около соответствующего пункта. Также можно установить уровень надежности, поставив галочку около соответствующего значения. По умолчанию он будет равен 95%, но его можно изменить, внося другие числа в поле справа.

Кроме этого, можно установить галочки в пунктах «К-ый наименьший» и «К-ый наибольший», установив значения в соответствующих полях. Но в нашем случае этот параметр так же, как и предыдущий, не является обязательным, поэтому флажки мы не ставим.

После того, как все указанные данные внесены, кликаем на кнопку «ОК».

После выполнения этих действий таблица с описательной статистикой выводится на отдельном листе, который был нами назван «Итоги». Как видим, данные представлены сумбурно, поэтому их следует отредактировать, расширив соответствующие колонки для более удобного просмотра.



После того, как данные «причесаны» можно приступить к их непосредственному анализу. Как видим, при помощи инструмента описательной статистики были рассчитаны следующие показатели:

- Асимметричность;
- Интервал;
- Минимум;
- Стандартное отклонение;
- Дисперсия выборки;
- Максимум;
- Сумма;
- Эксцесс;
- Среднее;
- Стандартная ошибка;
- Медиана;
- Мода;
- Счет.

Года	Численность населения города Алматы (тыс.)	Экономически активное население (тыс.)	Численность населения города Алматы (тыс.)		
1					
2	1991	1 134,4	598,2	Среднее	1305,874444
3	1992	1 134,2	575,0	Стандартная ошибка	42,03213027
4	1993	1 125,8	581,1	Медиана	1209,49
5	1994	1 114,7	587,8	Мода	#Н/Д
6	1995	1 111,7	605,8	Стандартное отклонение	218,4053555
7	1996	1 117,7	670,8	Дисперсия выборки	47700,89933
8	1997	1 120,1	616,5	Эксцесс	-0,115538708
9	1998	1 129,0	639,1	Асимметричность	1,022665238
10	1999	1 130,6	651,1	Интервал	690,3
11	2000	1 128,8	604,9	Минимум	1111,7
12	2001	1 132,4	598,0	Максимум	1802
13	2002	1 149,6	563,3	Сумма	35258,61
14	2003	1 179,2	577,3	Счет	27
15	2004	1 209,5	602,0	Наибольший(1)	1802
16	2005	1 247,9	619,1	Наименьший(1)	1111,7
17	2006	1 270,3	645,4	Уровень надежности(95,0%)	86,39828114
18	2007	1 324,7	677,3		
19	2008	1 361,9	705,5		
20	2009	1 390,6	706,8		
21	2010	1 415,2	722,4		
22	2011	1 449,4	740,9		
23	2012	1 474,8	774,7		
24	2013	1 506,9	787,3		
25	2014	1 641,4	809,2		
26	2015	1 702,7	885,1		
27	2016	1751,3	916,3		
28	2017	1802,0	939,2		

Рисунок 15.2. Исходные данные и результат работы вкладки «Описательная статистика»

Если какие-то из вышеуказанных данных для конкретного вида анализа не нужны, то их можно удалить, чтобы они не мешали. Далее производится анализ с учетом статистических закономерностей.

Как видим, с помощью инструмента «Описательная статистика» можно сразу получить результат по целому ряду критериев, которые в ином случае рассчитывались с применением отдельно предназначенной для каждого расчета функцией, что заняло бы значительное время у пользователя. А так, все эти расчеты можно получить практически в один клик, используя соответствующий инструмент – «Пакет анализа».



Замечание. Следует обратить внимание на то, что расчет параметров в режиме Описательная статистика имеет ряд важных особенностей:

1. В качестве значений параметров: Стандартное отклонение, Дисперсия выборки, Эксцесс, Асимметричность – MS Excel генерирует оценки соответствующих параметров для генеральной совокупности, а не для выборки.
2. Для применения Описательной статистики предварительное ранжирование исходных данных не требуется: при вычислении показателей ранжирование выполняется автоматически.
3. Появление в ячейке Мода индикатора ошибки #Н/Д указывает на то, что в анализируемых данных нет одинаковых значений признака. В этом случае в качестве моды M_o выбирается то значение признака, которое соответствует максимальной ординате теоретической кривой распределения.
4. Индикатор ошибки #ДЕЛ/0! В ячейке Эксцесс и/или Асимметричность означает, что в результирующей таблице стандартное отклонение является нулевым или же заданный входной диапазон данных содержит менее четырех элементов данных
5. Стандартная ошибка – это разность между ожидаемыми и наблюдаемыми значениями исследуемого признака.

Стандартная ошибка или ошибка среднего находится из выражения: $m = S_n$.

Следующий не менее интересный инструмент: Средство – Ранг и перцентиль – настройки «Пакет Анализ» позволяет создать таблицу, содержащую порядковый и процентный ранги для каждого значения в заданном наборе данных, при этом значения упорядочиваются в порядке убывания. Ранг – номер (порядковое место) значения признака в наборе данных.

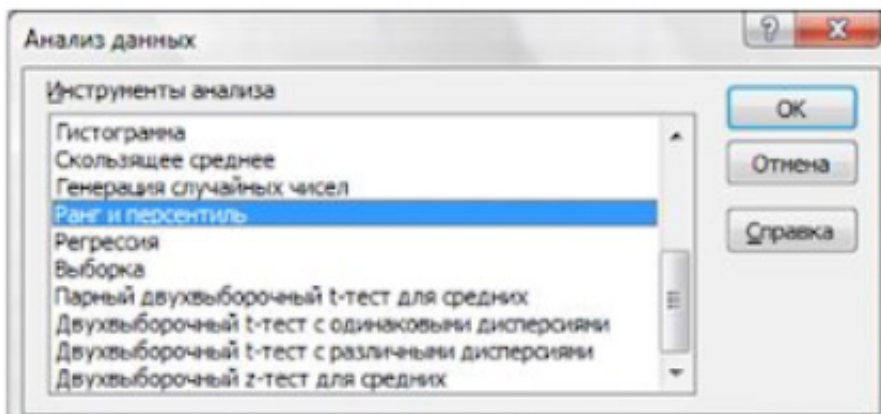


Рисунок 15.3 Окно запуска «Ранг и перцентиль»

Применим это средство для данных Численность населения (рис.15.2). Для этого выберем в Пакете анализа Ранг и перцентиль (рис.15.3).

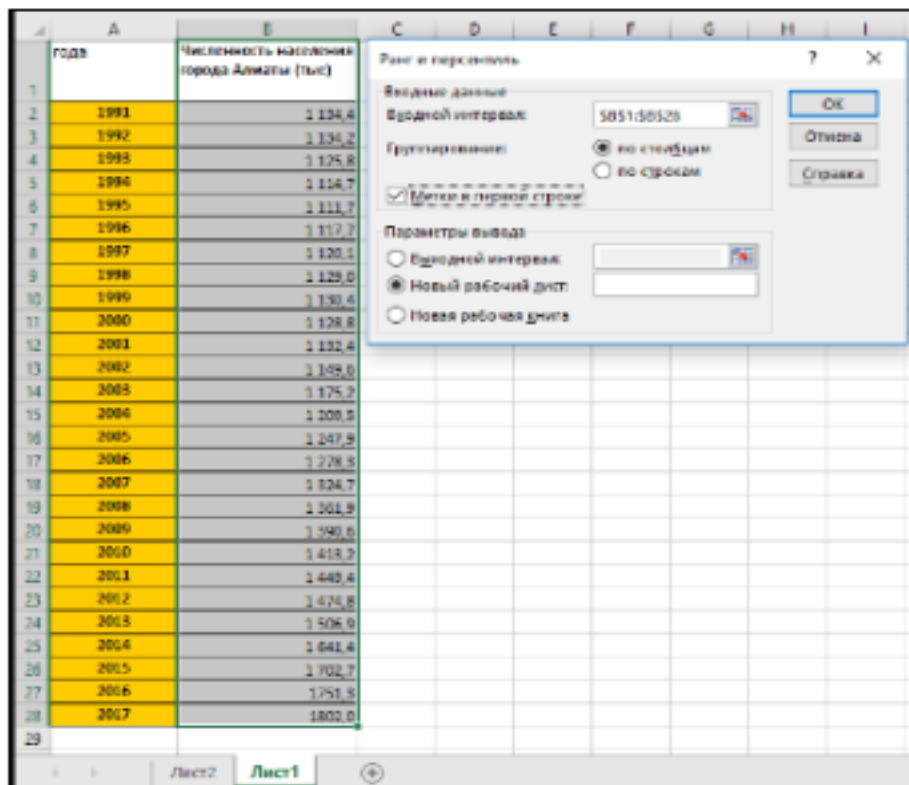


Рисунок 15.4 Средство анализа данных Ранг и перцентиль

год	Численность населения города Алматы (тыс)	Численность населения города Алматы (тыс)		
		Ранг	Процент	Процент
1991	1 134,4	27	1 802,0	1 100,00%
1992	1 134,2	26	1 751,3	2 96,10%
1993	1 125,8	25	1 702,7	3 62,40%
1994	1 114,7	24	1 641,4	4 88,40%
1995	1 111,7	23	1 506,9	5 84,60%
1996	1 117,7	22	1 474,8	6 80,70%
1997	1 120,1	21	1 449,4	7 76,90%
1998	1 129,0	20	1 413,2	8 73,00%
1999	1 130,4	19	1 390,6	9 69,20%
2000	1 128,8	18	1 361,9	10 65,30%
2001	1 132,4	17	1 324,7	11 61,50%
2002	1 135,0	16	1 278,3	12 57,60%
2003	1 175,2	15	1 247,9	13 53,80%
2004	1 209,5	14	1 209,5	14 50,00%
2005	1 247,9	13	1 175,2	15 46,10%
2006	1 228,3	12	1 149,6	16 42,30%
2007	1 324,7	1	1 134,4	17 38,40%
2008	1 361,9	2	1 134,2	18 34,60%
2009	1 390,6	11	1 132,1	19 30,70%
2010	1 413,2	9	1 130,4	20 26,90%
2011	1 449,4	8	1 129,0	21 23,00%
2012	1 474,8	10	1 128,8	22 19,20%
2013	1 506,9	8	1 125,8	23 15,30%
2014	1 641,4	7	1 120,1	24 11,50%
2015	1 702,7	6	1 117,7	25 7,60%
2016	1 751,3	4	1 114,7	26 3,80%
2017	1802,0	5	1 111,7	27 0,00%

Рисунок 15.5 Результаты вычислений Ранг и перцентиль.



На рисунках 15.4 и 15.5 показаны диалоговое окно Ранг и перцентиль и исходные данные. Результат применения этого средства представлен на рисунке 15.5. Итоговая таблица содержит порядковый номер выборочного значения, столбец выборочных значений, отсортированных в порядке убывания, столбец рангов и столбец процентных рангов этих значений, причем наибольшему значению присваивается ранг 1 и процентный ранг 100%, а наименьшему – наибольший ранг и процентный ранг, равный 0%.

Вычисление доверительного интервала в Microsoft Excel

Одним из методов решения статистических задач является вычисление доверительного интервала. Он используется, как более предпочтительная альтернатива точечной оценке при небольшом объеме выборки. Нужно отметить, что сам процесс вычисления доверительного интервала довольно сложный. Но инструменты программы MS Excel позволяют несколько упростить его. Давайте узнаем, как это выполняется на практике.

Одним из методов решения статистических задач является вычисление доверительного интервала. Он используется, как более предпочтительная альтернатива точечной оценке при небольшом объеме выборки. Нужно отметить, что сам процесс вычисления доверительного интервала довольно сложный. Но инструменты программы MS Excel позволяют несколько упростить его.

В MS Excel существуют два основных варианта произвести вычисления с помощью данного метода: когда дисперсия известна, и когда она неизвестна. В первом случае для вычислений применяется функция ДОВЕРИТ.НОРМ, а во втором – ДОВЕРИТ.СТЮДЕНТ.

Способ 1: функция ДОВЕРИТ.НОРМ.

Функция ДОВЕРИТ, относится к статистической группе функций. Задачей этого оператора является расчет доверительного интервала с нормальным распределением для средней генеральной совокупности. Его синтаксис выглядит следующим образом:

=ДОВЕРИТ(альфа;стандартное_откл;размер)

Аргументы функции:

«Альфа» — аргумент, указывающий на уровень значимости, который применяется для расчета доверительного уровня. Доверительный уровень равняется следующему выражению: $(1 - \text{«Альфа»}) * 100$

«Стандартное отклонение» — это аргумент, суть которого понятна из наименования. Это стандартное отклонение предлагаемой выборки.

«Размер» — аргумент, определяющий величину выборки.

Все аргументы данного оператора являются обязательными.

Граница доверительного интервала определяется при помощи формулы следующего вида:

$X + (-) \text{ДОВЕРИТ.НОРМ}$

где X — это среднее выборочное значение, которое расположено посередине выбранного диапазона.

Теперь давайте рассмотрим, как рассчитать доверительный интервал на конкретном примере. Была проведена выборка за 27 лет экономически активного населения г. Алматы в динамике с 1991 по 2017 г., вследствие которых были получены различные результаты, занесенные в таблицу. Это и есть наша совокупность. Стандартное отклонение равно 111,98. Нам нужно рассчитать доверительный интервал при уровне доверия 97%.

Для этого в диапазоне набираем данные «экономически активного населения г. Алматы» по годам. Затем этому диапазону ячеек присваиваем имя «Выборка». Выделяем ячейку, куда будет выводиться результат обработки данных. Кликаем по кнопке «Вставить функцию». Появляется Мастер функций. Переходим в категорию «Статистические» и выделяем наименование



«ДОВЕРИТ.НОРМ». После этого кликаем по кнопке «ОК».

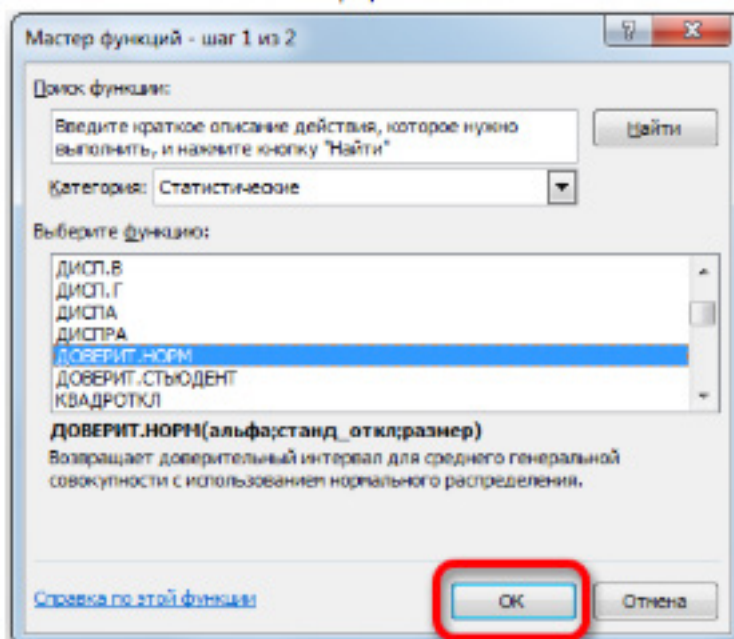


Рисунок 15.7. Окно выбора статистической функции

Открывается окно аргументов. Его поля закономерно соответствуют наименованиям аргументов. Устанавливаем курсор в первое поле – «Альфа». Тут нам следует указать уровень значимости. Как мы помним, уровень доверия у нас равен 97%. В то же время мы говорили, что он рассчитывается таким путем: $(1 - \text{«Альфа»}) * 100$.

Значит, чтобы посчитать уровень значимости, то есть, определить значение «Альфа» следует применить формулу такого вида: $(1 - \text{уровень доверия}) / 100$. То есть, подставив значение, получаем: $(1 - 97) / 100$. Путем нехитрых расчетов узнаем, что аргумент «Альфа» равен 0,03. Вводим данное значение в поле. Как известно, по условию стандартное отклонение равно 111,98. Поэтому в поле «Стандартное отклонение» просто записываем это число. В поле «Размер» нужно ввести количество наблюдаемых лет. Как мы помним, их 27. Но чтобы автоматизировать формулу и не редактировать её каждый раз при проведении нового испытания, давайте зададим данное значение не обычным числом, а при помощи функции СЧЁТ. Итак, устанавливаем курсор в поле «Размер», а затем кликаем по треугольнику, который размещен слева от строки формул. Появляется список недавно применяемых функций. Если функции СЧЁТ применялся вами недавно, то он должен быть в этом списке. В таком случае, нужно просто кликнуть по его наименованию. В обратном же случае, если вы его не обнаружите, то переходите по пункту «Другие функции...».

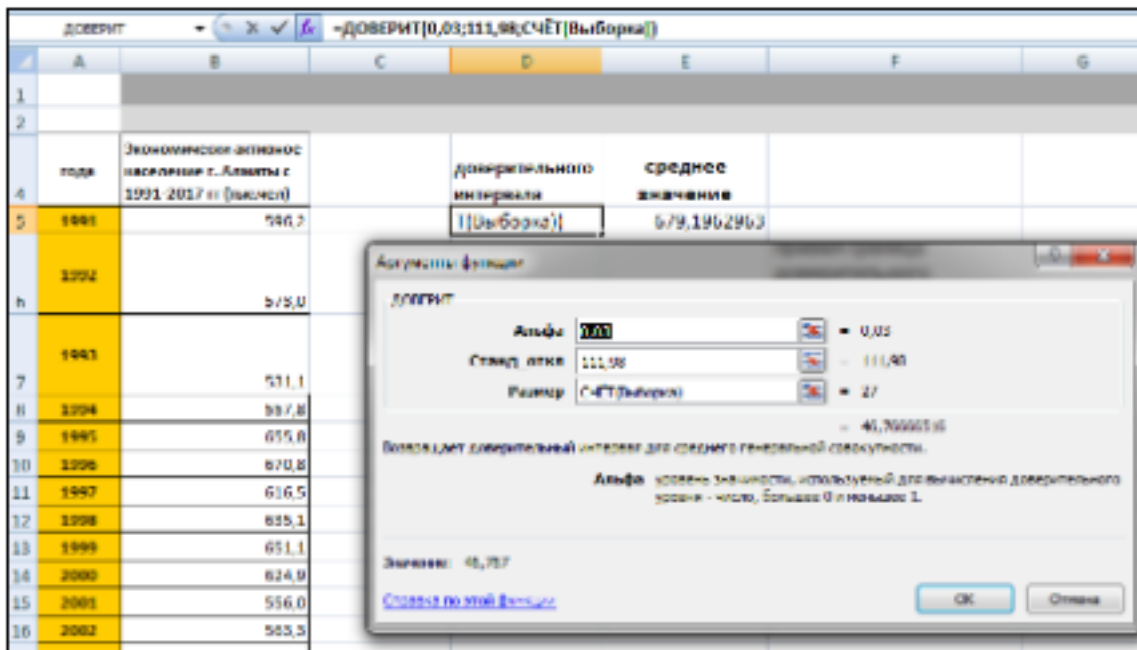


Рисунок 15.8. Диалоговое окно функции ДОВЕРИТ

Появляется уже знакомый нам Мастер функций. Опять перемещаемся в группу «Статистические». Выделяем там наименование «СЧЁТ». Кликаем по кнопке «ОК».

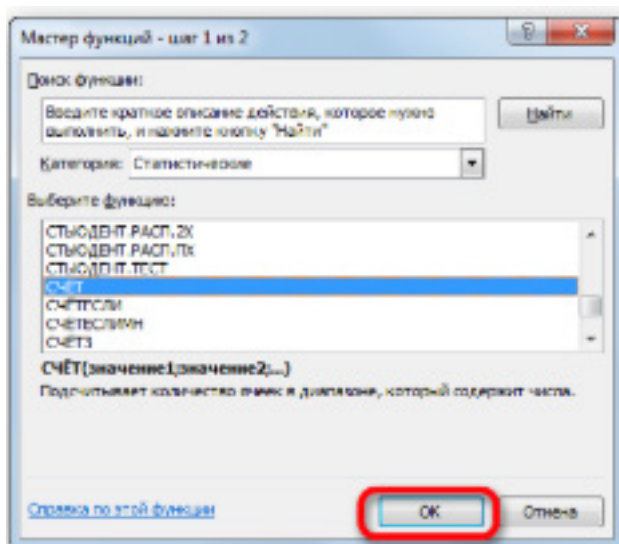


Рис. 15.9. Диалоговое окно Мастера функций

Появляется окно аргументов вышеуказанной функции. Данная функция предназначена для того, чтобы вычислять количество ячеек в указанном диапазоне, которые содержат числовые значения. Синтаксис её следующий: =СЧЁТ(значение1;значение2;...).

Группа аргументов «Значения» представляет собой ссылку на диапазон, в котором нужно рассчитать количество заполненных числовыми данными ячеек. Всего может насчитываться до 255 подобных аргументов, но в нашем случае понадобится лишь один.

Устанавливаем курсор в поле «Значение1» и, зажав левую кнопку мыши, выделяем на листе



диапазон, который содержит нашу совокупность. Затем его адрес будет отображен в поле. Кликаем по кнопке «ОК».

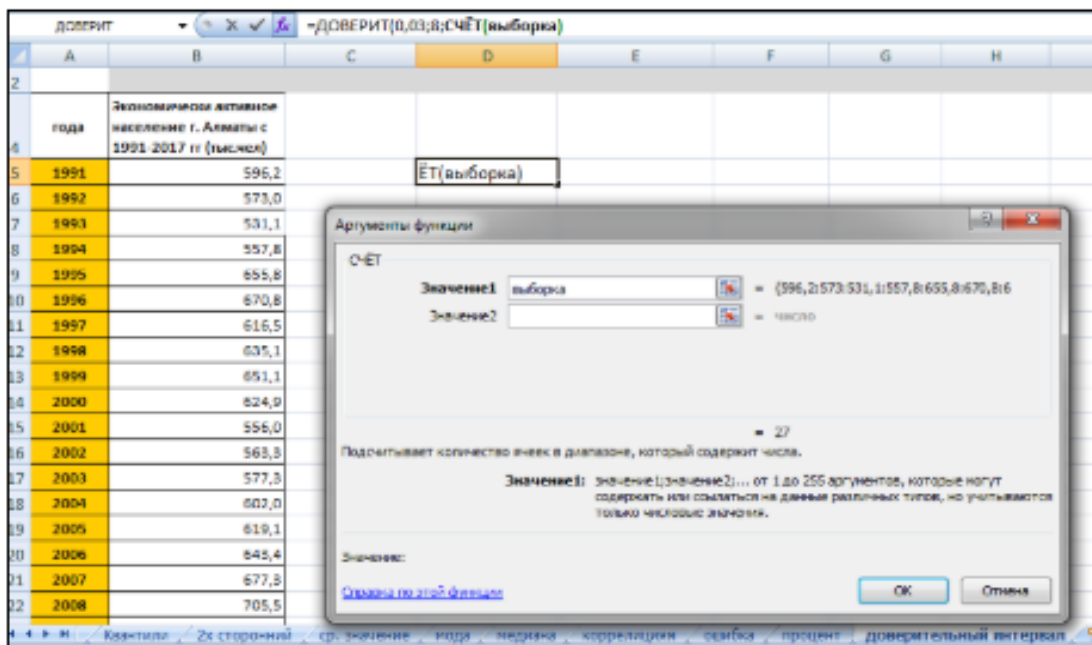


Рисунок 15.10 Диалоговое окно функции СЧЕТ

После этого приложение произведет вычисление и выведет результат в ту ячейку, где она находится сама. В нашем конкретном случае формула получилась такого вида:

$=\text{ДОВЕРИТ.НОРМ}(0,03;111,98;\text{СЧЁТ}(B5:B31))$

Общий результат вычислений составил 46,76.

Но это ещё не все. Как мы помним, граница доверительного интервала вычисляется путем сложения и вычитания от среднего выборочного значения результата вычисления ДОВЕРИТ.НОРМ. Таким способом рассчитывается соответственно правая и левая граница доверительного интервала. Само среднее выборочное значение можно рассчитать при помощи функции СРЗНАЧ. Данная функция предназначена для расчета среднего арифметического значения выбранного диапазона чисел. Она имеет следующий довольно простой синтаксис: $=\text{СРЗНАЧ}(\text{число1};\text{число2};\dots)$

Аргумент «Число» может быть как отдельным числовым значением, так и ссылкой на ячейки или даже целые диапазоны, которые их содержат.

Итак, выделяем ячейку, в которую будет выводиться расчет среднего значения, и кликаем по кнопке «Вставить функцию». Открывается Мастер функций. Снова переходим в категорию «Статистические» и выбираем из списка наименование «СРЗНАЧ». Как всегда, кликаем по кнопке «ОК». Запускается окно аргументов. Устанавливаем курсор в поле «Число1» и с зажатой левой кнопкой мыши выделяем весь диапазон значений. После того, как координаты отобразились в поле, кликаем по кнопке «ОК».

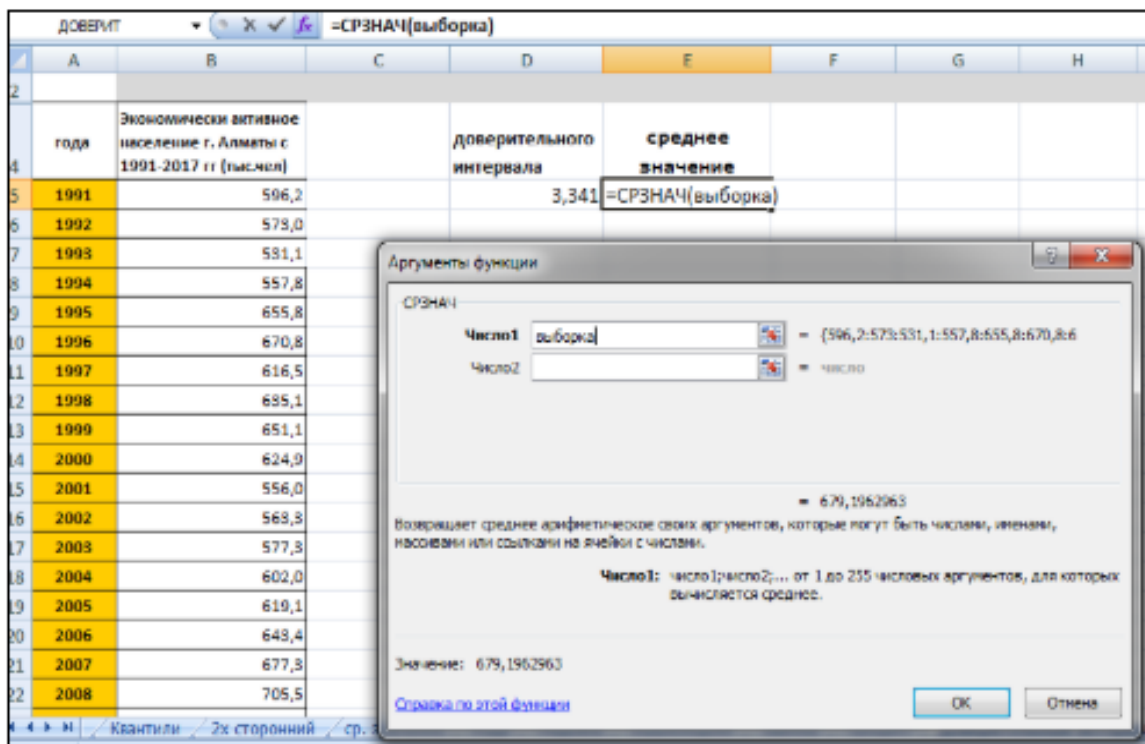


Рисунок 15.11 Диалоговое окно функции СРЗНАЧ

После этого СРЗНАЧ выводит результат расчета в элемент листа. Она равна 679,19 тыс. человек. Производим расчет правой границы доверительного интервала. Для этого выделяем отдельную ячейку, ставим знак «=» и складываем содержимое элементов листа, в которых расположены результаты вычислений функций СРЗНАЧ и ДОВЕРИТ.НОРМ. Для того, чтобы выполнить расчет, кликаем на клавишу Enter. В нашем случае получилась следующая формула: =E5+D5. Результат вычисления: 725,963.

Таким же образом производим вычисление левой границы доверительного интервала, только на этот раз от результата вычисления СРЗНАЧ отнимаем результат вычисления функции ДОВЕРИТ.НОРМ. Получается формула для нашего примера следующего типа: =E5-D5. Результат вычисления: 632,430.

Мы попытались подробно описать все действия по вычислению доверительного интервала, поэтому детально расписали каждую формулу. Но можно все действия соединить в одной формуле. Вычисление правой границы доверительного интервала можно записать так:

=СРЗНАЧ(Выборка)+ДОВЕРИТ.НОРМ(0,03;111,98;СЧЁТ(Выборка))

Аналогичное вычисление левой границы будет выглядеть так:

=СРЗНАЧ(Выборка)-ДОВЕРИТ.НОРМ(0,03;111,98;СЧЁТ(Выборка))

Способ 2: функция ДОВЕРИТ.СТЮДЕНТ

Кроме того, в MS Excel есть ещё одна функция, которая связана с вычислением доверительного интервала – ДОВЕРИТ.СТЮДЕНТ. Данная функция выполняет вычисление доверительного интервала генеральной совокупности с использованием распределения Стьюдента. Ее очень удобно использовать в том случае, когда дисперсия и, соответственно, стандартное отклонение неизвестны. Синтаксис функции такой:

=ДОВЕРИТ.СТЮДЕНТ(альфа;стандартное_откл;размер)

Как видим, наименования функций и в этом случае остались неизменными. Посмотрим, как

рассчитать границы доверительного интервала с неизвестным стандартным отклонением на примере всё той же задачи, что мы рассматривали в предыдущем способе. Уровень доверия, как и в прошлый раз, возьмем 97%. Рассмотрим тот же самый пример. Выделяем ячейку, в которую будет производиться расчет. Кликаем по кнопке «Вставить функцию». В открывшемся Мастере функций переходим в категорию «Статистические». Выбираем наименование «ДОВЕРИТ.СТЮДЕНТ». Кликаем по кнопке «ОК». Производится запуск окна аргументов указанной функции. В поле «Альфа», учитывая, что уровень доверия составляет 97%, записываем число 0,03. Второй раз на принципах расчета данного параметра останавливаться не будем. После этого устанавливаем курсор в поле «Стандартное отклонение». На этот раз данный показатель нам неизвестен и его требуется рассчитать. Делается это при помощи специальной функции – СТАНДОТКЛОН.В. Чтобы вызвать окно данного оператора, кликаем по треугольнику слева от строки формул. Если в открывшемся списке не находим нужного наименования, то переходим по пункту «Другие функции...».

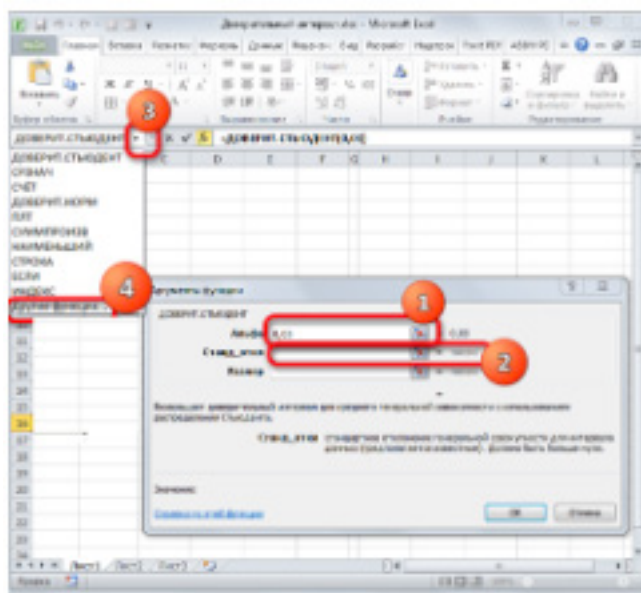


Рисунок 15.12 Диалоговое окно функции «ДОВЕРИТ.СТЮДЕНТ»

Запускается Мастер функций. Перемещаемся в категорию «Статистические» и отмечаем в ней наименование «СТАНДОТКЛОН.В». Затем кликаем по кнопке «ОК». Открывается окно аргументов. Задачей функции СТАНДОТКЛОН.В является в определение стандартного отклонения при выборке. Ее синтаксис выглядит так:

=СТАНДОТКЛОН.В(число1; число2;...).

Нетрудно догадаться, что аргумент «Число» — это адрес элемента выборки. Если выборка размещена единым массивом, то можно, использовав только один аргумент, дать ссылку на данный диапазон. Устанавливаем курсор в поле «Число1» и, как всегда, зажав левую кнопку мыши, выделяем совокупность. После того, как координаты попали в поле, не спешим кликать на кнопку «ОК», так как результат получится некорректным. Прежде нам нужно вернуться к окну аргументов функции ДОВЕРИТ.СТЮДЕНТ, чтобы внести последний аргумент. Для этого кликаем по соответствующему наименованию в строке формул. Снова открывается окно аргументов уже знакомой функции. Устанавливаем курсор в поле «Размер». Опять кликаем на уже знакомый нам треугольник для перехода к выбору функций. Как вы поняли, нам нужно наименование «СЧЁТ». Так как мы использовали данную функцию при вычислениях в предыдущем способе, в данном списке она присутствует, так что просто кликаем по ней. Если же вы её не обнаружите, то



действуйте по алгоритму, описанному в первом способе. Попав в окно аргументов СЧЁТ, ставим курсор в поле «Число1» и с зажатой кнопкой мыши выделяем совокупность. Затем кликаем по кнопке «ОК».

год	Экономически активное население г. Алматы с 1991-2017 гг (тысяч)	доверительный интервал	среднее значение
1991	596,2	46,765	675,1962963
1992	573,0		
1993	591,1		
1994	557,8		
1995	655,8		
1996	670,8		
1997	655,5		
1998	639,1		
1999	651,1		

Рисунок 15.13 Результаты расчета

После этого программа производит расчет и выводит значение доверительного интервала. Для определения границ нам опять нужно будет рассчитать среднее значение выборки. Но, учитывая то, что алгоритм расчета при помощи формулы СРЗНАЧ тот же, что и в предыдущем способе, и даже результат не изменился, не будем на этом подробно останавливаться второй раз. Сложив результаты вычисления СРЗНАЧ и ДОВЕРИТ. СТЮДЕНТ, получаем правую границу доверительного интервала. Отняв от результатов расчета функции СРЗНАЧ результат расчета ДОВЕРИТ. СТЮДЕНТ, имеем левую границу доверительного интервала.

Если расчет записать одной формулой, то вычисление правой границы в нашем случае будет выглядеть так:

$$= \text{СРЗНАЧ}(\text{Выборка}) + \text{ДОВЕРИТ.СТЮДЕНТ}(0,03;\text{СТАНДОТКЛОН.В}(\text{Выборка});\text{СЧЁТ}(\text{Выборка}))$$

Соответственно, формула расчета левой границы будет выглядеть так:

$$= \text{СРЗНАЧ}(\text{Выборка}) - \text{ДОВЕРИТ.СТЮДЕНТ}(0,03;\text{СТАНДОТКЛОН.В}(\text{Выборка});\text{СЧЁТ}(\text{Выборка}))$$

Как видим, инструменты программы MS Excel позволяют существенно облегчить вычисление доверительного интервала и его границ. Для этих целей используются отдельные операторы для выборок, у которых дисперсия известна и неизвестна.

Подведем небольшие итоги, сегодня мы рассмотрели одну из возможностей Надстройки Анализа данных, которые выводит данные по категории описательная статистика. Кроме того, рассмотрели достаточно непростую процедуру вычисления доверительного интервала с пошаговым выполнением операций.

Выполнив домашнее задание, вы закрепите данный материал и сможете эффективно использовать его в повседневной и профессиональной деятельности.