Обзор рекомендуемого софта для биомедицинских исследований

ФГБОУ ВО «Пермский государственный медицинский Университет имени академика Е.А. Вагнера» МЗ РФ
Кафедра детских инфекционных болезней

Образовательный проект «Школа молодого ученого ПГМУ»

Описательная статистика. Расчет количественных статистических показателей

- Среднее значение
- Медиана
- Стандартное отклонение
- Доверительный интервал

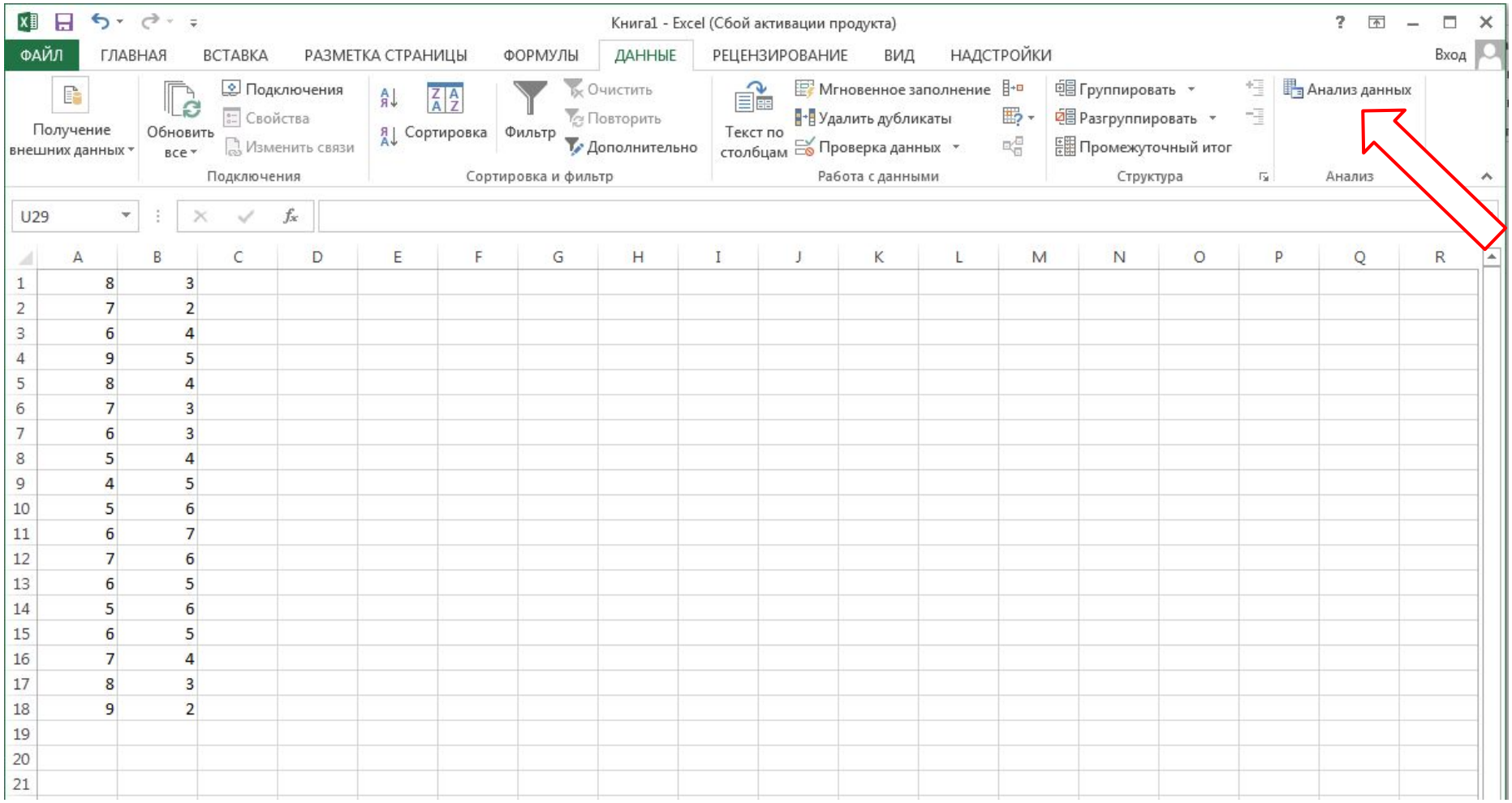
Наиболее распространенные статистические критерии, применяемые для решения задач медицинской статистики

Сравниваемые показатели	Параметрические тесты	Непараметрические тесты	Методы сравнения долей
Две независимые группы	Критерий Стьюдента (двухвыборочный), дисперсионный анализ	Критерий Манна-Уитни, критерий Уилкоксона, медианный критерий	Критерий χ^2, точный критерий Фишера
Более двух независимых групп	Дисперсионный анализ, критерий Стьюдента для множественных сравнений, критерий Даннета	Критерий Краскела-Уоллиса, медианный критерий	Критерий χ^2
Одна группа, связанные измерения	Критерий Стьюдента для связанных пар (парный тест), дисперсионный анализ	Т-критерий Уилкоксона, критерий знаков	Критерий Мак-Нимара
Одна группа, несколько связанных измерений	Дисперсионный анализ повторных измерений, критерий Шеффе для зависимых выборок	Критерий Фридмана	Критерий Кокрена
Два и более сопряженных ряда чисел (выявление корелятивной связи)	Коэффициент корреляции Пирсона	Коэффициент ранговой корреляции Спирмена	

Для начинающих

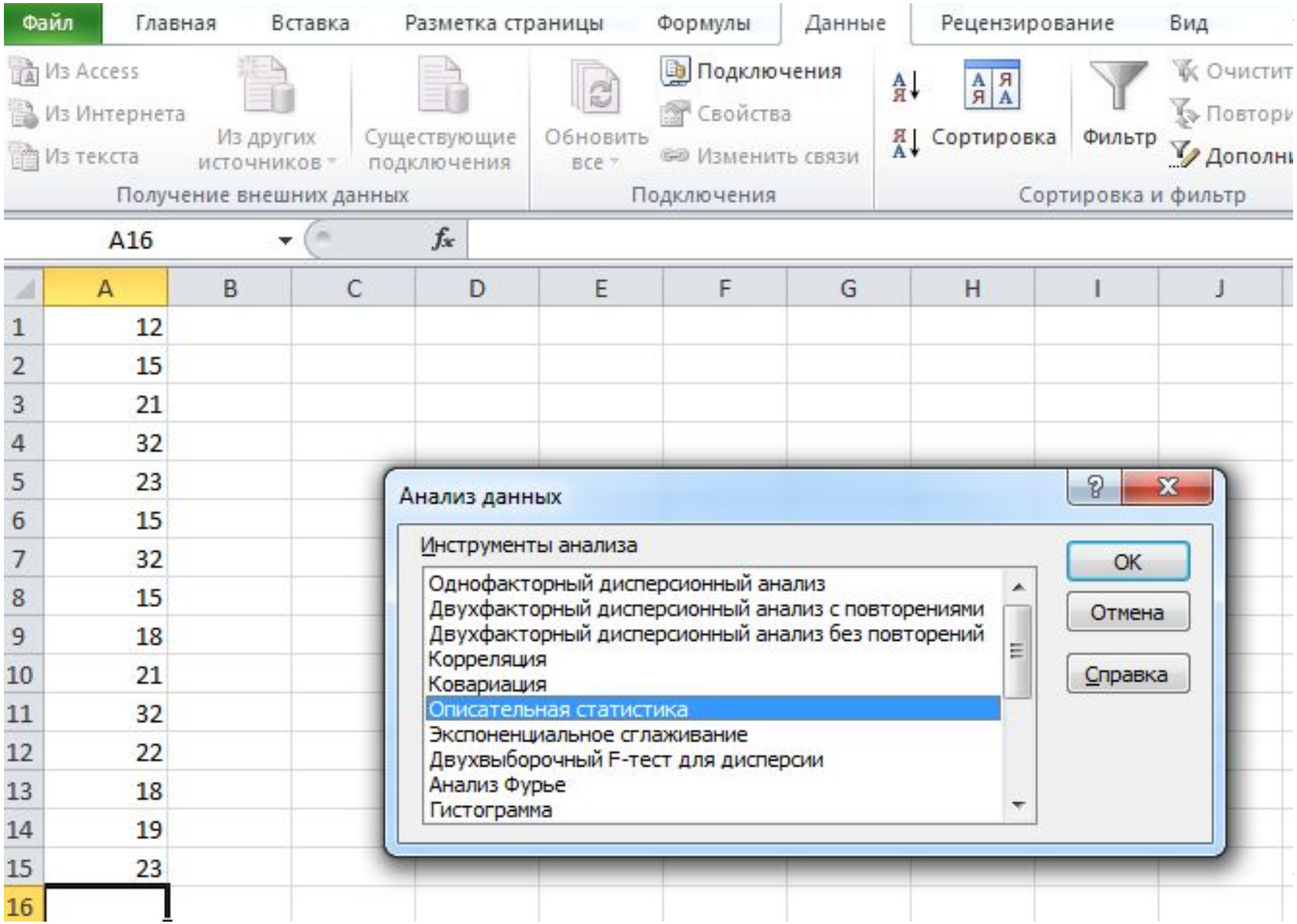
- MS Excel и его надстройка Пакет анализа
- Attestat
- Он-лайн калькуляторы
- Biostat.exe

Пакет анализа

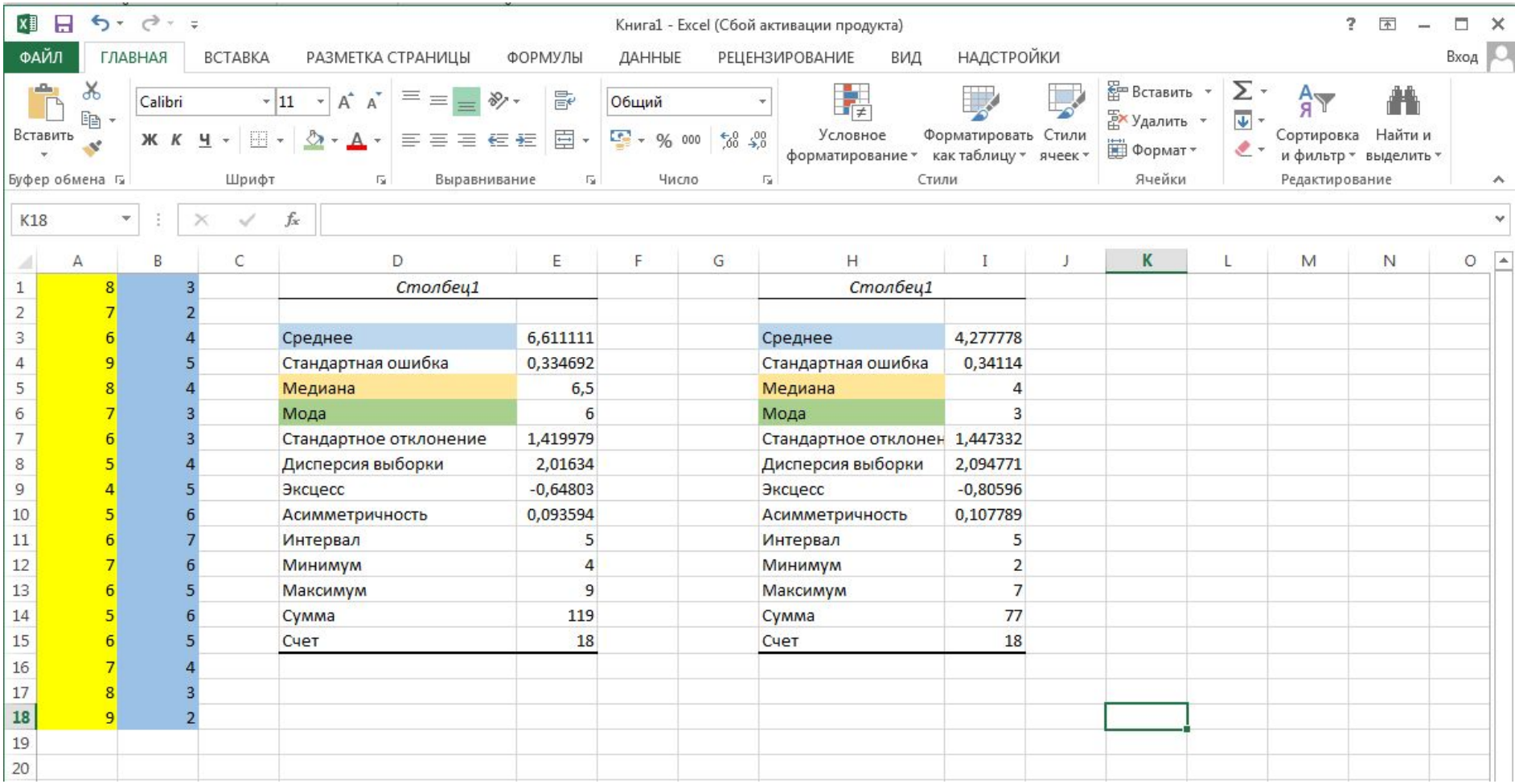


	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	8	3																
2	7	2																
3	6	4																
4	9	5																
5	8	4																
6	7	3																
7	6	3																
8	5	4																
9	4	5																
10	5	6																
11	6	7																
12	7	6																
13	6	5																
14	5	6																
15	6	5																
16	7	4																
17	8	3																
18	9	2																
19																		
20																		
21																		

Пакет анализа



Пример описательной статистики в Пакете анализа



Книга1 - Excel (Сбой активации продукта)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	8	3		*Столбец1*				*Столбец1*							
2	7	2													
3	6	4		Среднее	6,611111			Среднее	4,277778						
4	9	5		Стандартная ошибка	0,334692			Стандартная ошибка	0,34114						
5	8	4		Медиана	6,5			Медиана	4						
6	7	3		Мода	6			Мода	3						
7	6	3		Стандартное отклонение	1,419979			Стандартное отклонен	1,447332						
8	5	4		Дисперсия выборки	2,01634			Дисперсия выборки	2,094771						
9	4	5		Эксцесс	-0,64803			Эксцесс	-0,80596						
10	5	6		Асимметричность	0,093594			Асимметричность	0,107789						
11	6	7		Интервал	5			Интервал	5						
12	7	6		Минимум	4			Минимум	2						
13	6	5		Максимум	9			Максимум	7						
14	5	6		Сумма	119			Сумма	77						
15	6	5		Счет	18			Счет	18						
16	7	4													
17	8	3													
18	9	2													
19															
20															

Пакет анализа

Статистические средства надстройки Пакет анализа

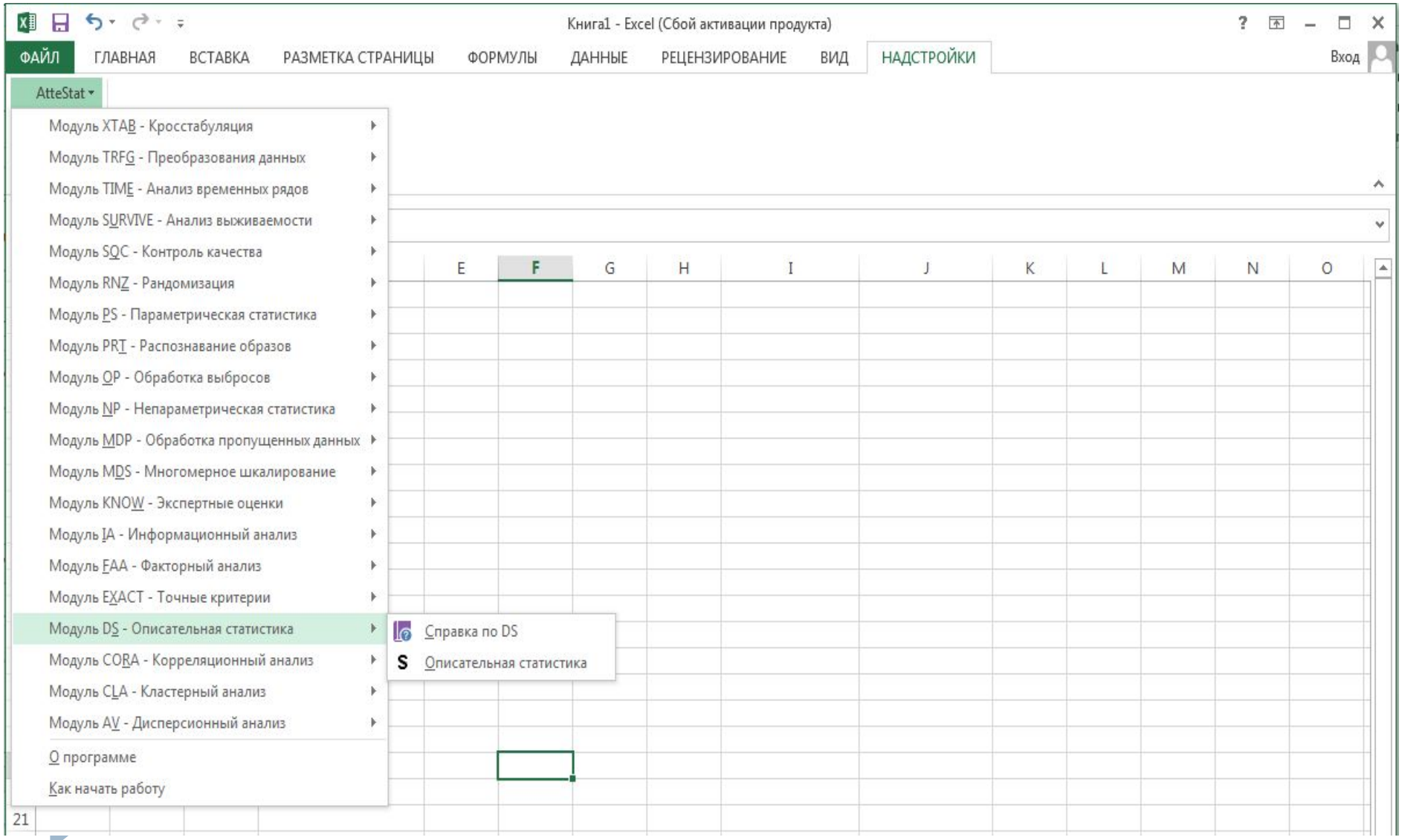
Средство	*Описание*
Однофакторный дисперсионный анализ	Используется для проверки гипотезы о равенстве математических ожиданий двух или более выборок
Двухфакторный дисперсионный анализ без повторений	Двухфакторный дисперсионный анализ на основе одной выборки
Двухфакторный дисперсионный анализ с повторениями	Двухфакторный дисперсионный анализ на основе нескольких выборок
Корреляция	Вычисляет корреляционную матрицу
Ковариация	Вычисляет матрицу ковариаций
Описательная статистика	Создает отчет, содержащий статистические характеристики представленной выборки
Экспоненциальное сглаживание	Реализует метод экспоненциального сглаживания данных
Двухвыборочный F-тест для дисперсий	Применяется для сравнения дисперсий двух генеральных совокупностей
Анализ Фурье	Реализует метод быстрого преобразования Фурье (БПФ) для анализа данных
Гистограмма	Используется для анализа распределения выборочных данных и построения гистограмм
Скользящее среднее	Используется для сглаживания данных
Генерация случайных чисел	Генерирует случайные числа, имеющие заданное распределение
Ранг и персентиль	Используется для вычисления рангов и квантилей

Пакет анализа

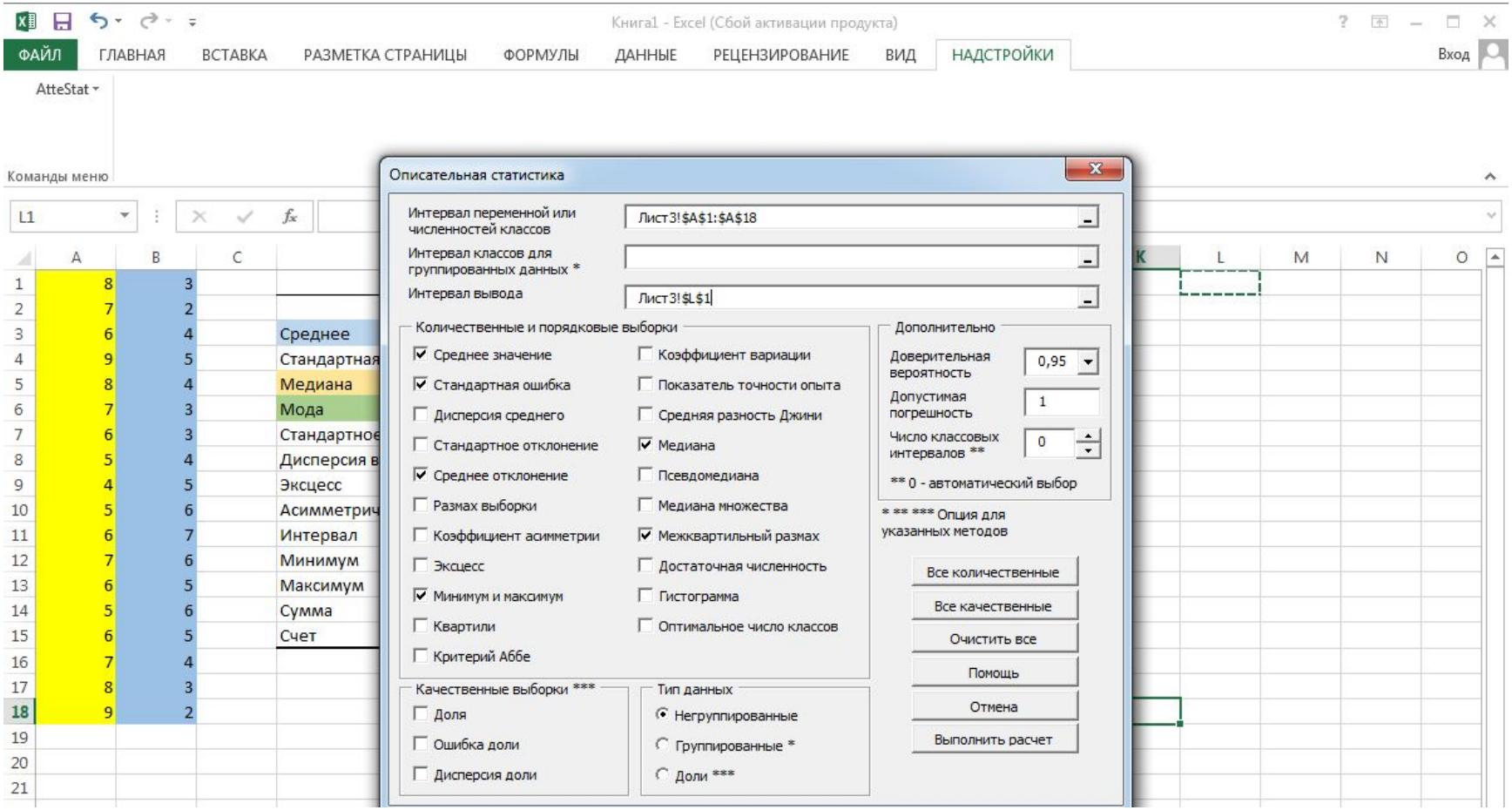
Средство	*Описание*
Регрессия	Испо[illegible]ьзуется для [illegible]ения линейной регрессии
Выборка	Со[illegible]рку, рас[illegible]тривая входной [illegible]ую совокупность
Парный двухвыбор[illegible] тест [illegible]	[illegible]стве [illegible]мерной выборки
Дву[illegible] один[illegible]	[illegible]равенстве мате- [illegible]рок. Предпо- [illegible]сий генеральных сово-
Двухвыборочный [illegible] разными дисперс[illegible]ми	Используется для проверки гипотезы о равенстве математических ожиданий для двух выборок. Не требует предположения о равенстве дисперсий генеральных совокупностей
Двухвыборочный z-тест для средних	Используется для проверки гипотезы о различии между математическими ожиданиям двух генеральных совокупностей

19 статистических процедур и около 50 функций

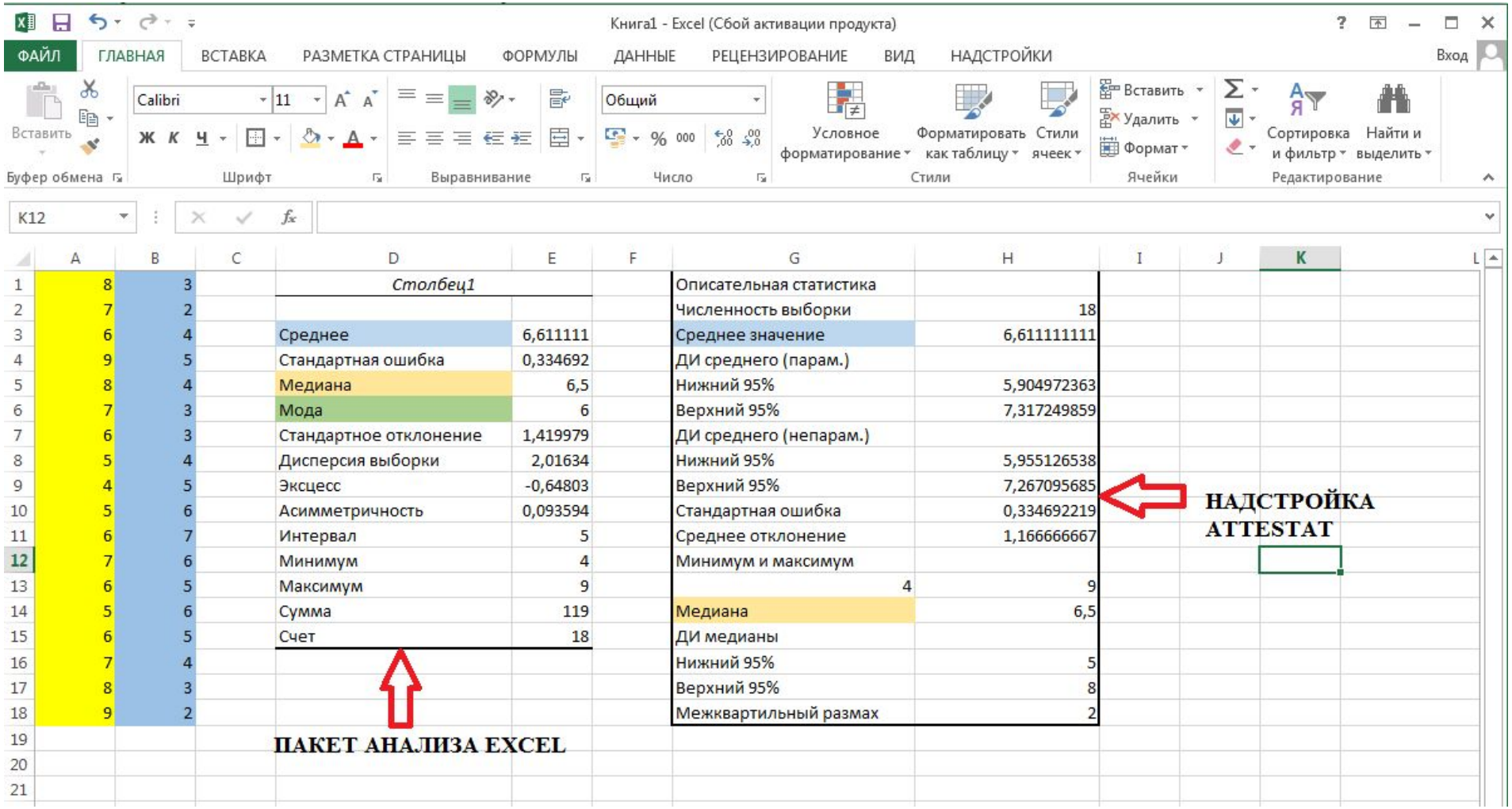
Attestat –надстройка Excel

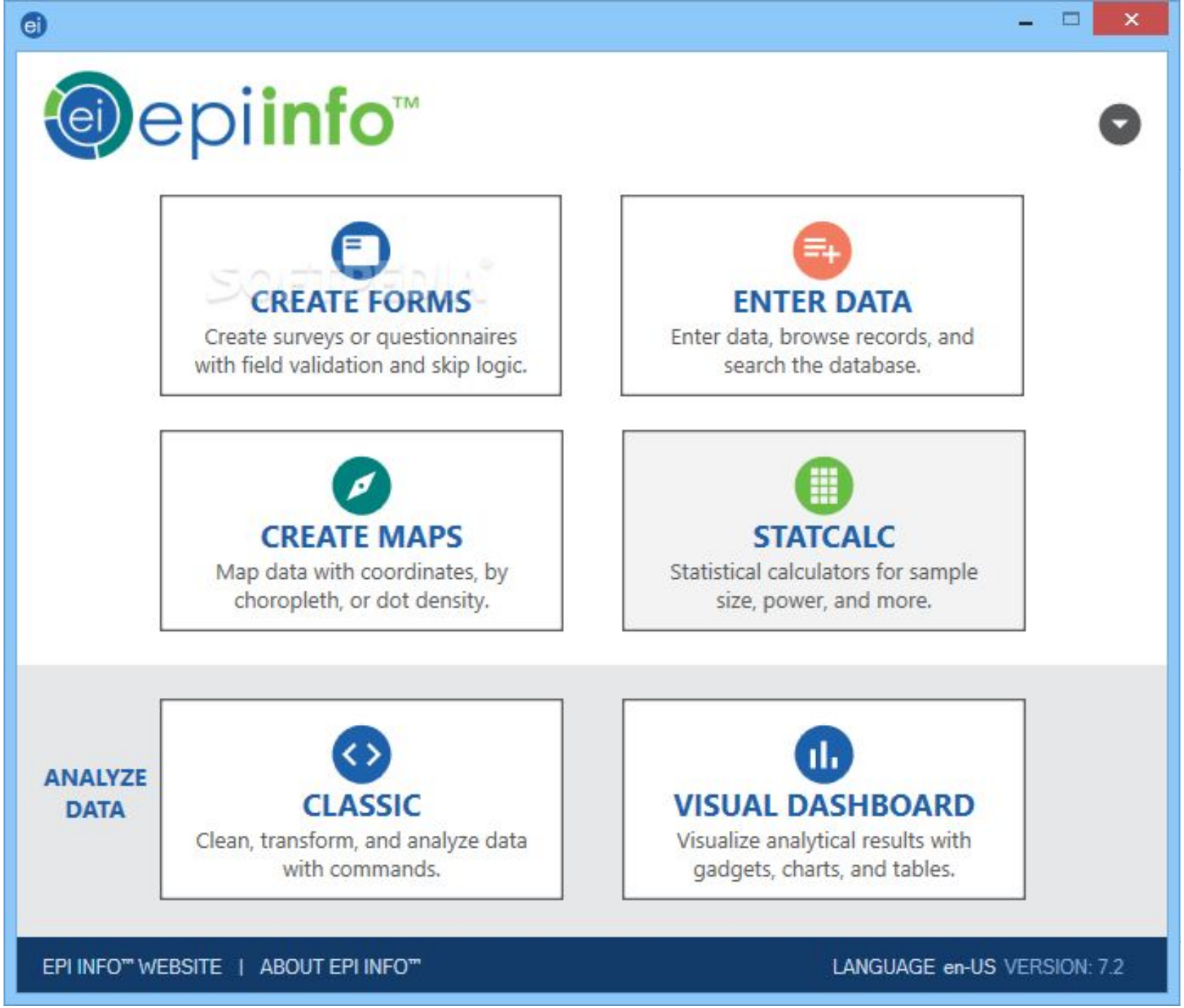


Пример описательной статистики в надстройке Excel - Attestat

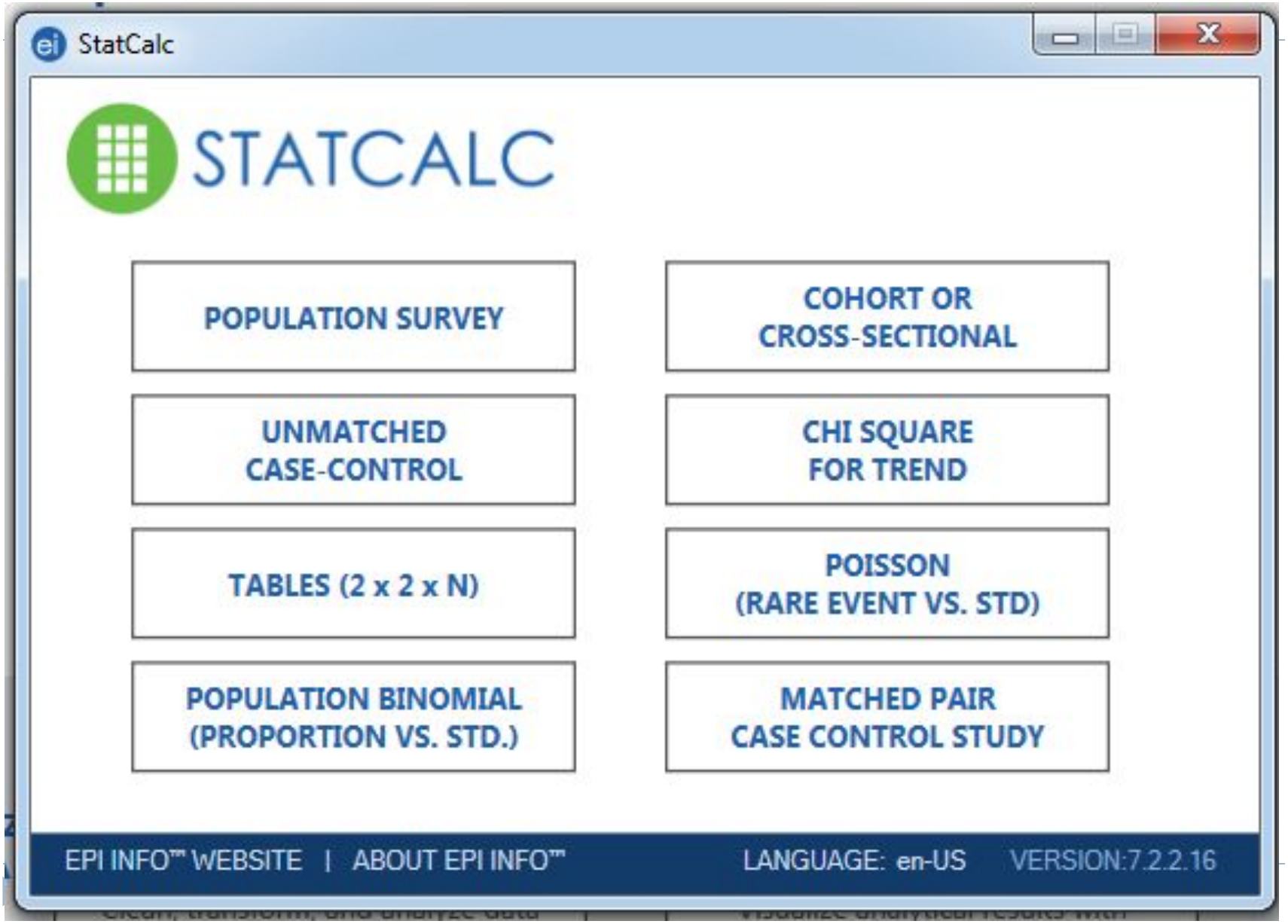


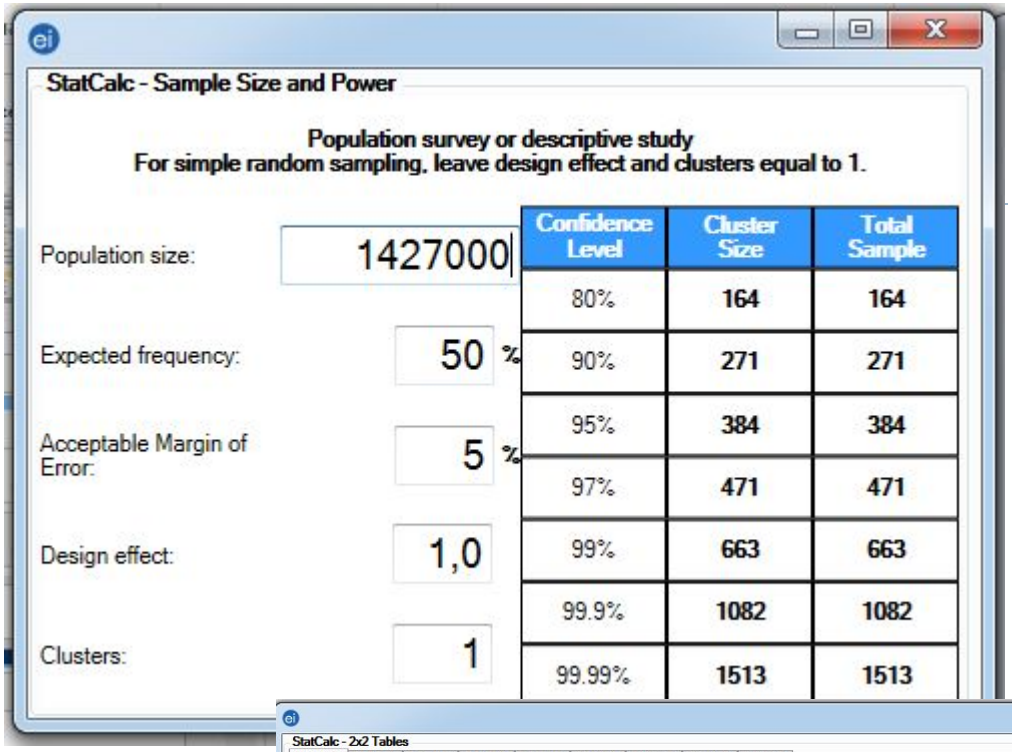
Пакет анализа VS Attestat




epiinfo™
CREATE FORMS
Create surveys or questionnaires with field validation and skip logic.
ENTER DATA
Enter data, browse records, and search the database.
CREATE MAPS
Map data with coordinates, by choropleth, or dot density.
STATCALC
Statistical calculators for sample size, power, and more.
ANALYZE DATA
CLASSIC
Clean, transform, and analyze data with commands.
VISUAL DASHBOARD
Visualize analytical results with gadgets, charts, and tables.
EPI INFO™ WEBSITE | ABOUT EPI INFO™
LANGUAGE en-US VERSION: 7.2

EpiInfo


StatCalc
STATCALC
POPULATION SURVEY
COHORT OR CROSS-SECTIONAL
UNMATCHED CASE-CONTROL
CHI SQUARE FOR TREND
TABLES (2 x 2 x N)
POISSON (RARE EVENT VS. STD)
POPULATION BINOMIAL (PROPORTION VS. STD.)
MATCHED PAIR CASE CONTROL STUDY
EPI INFO™ WEBSITE | ABOUT EPI INFO™
LANGUAGE: en-US
VERSION:7.2.2.16



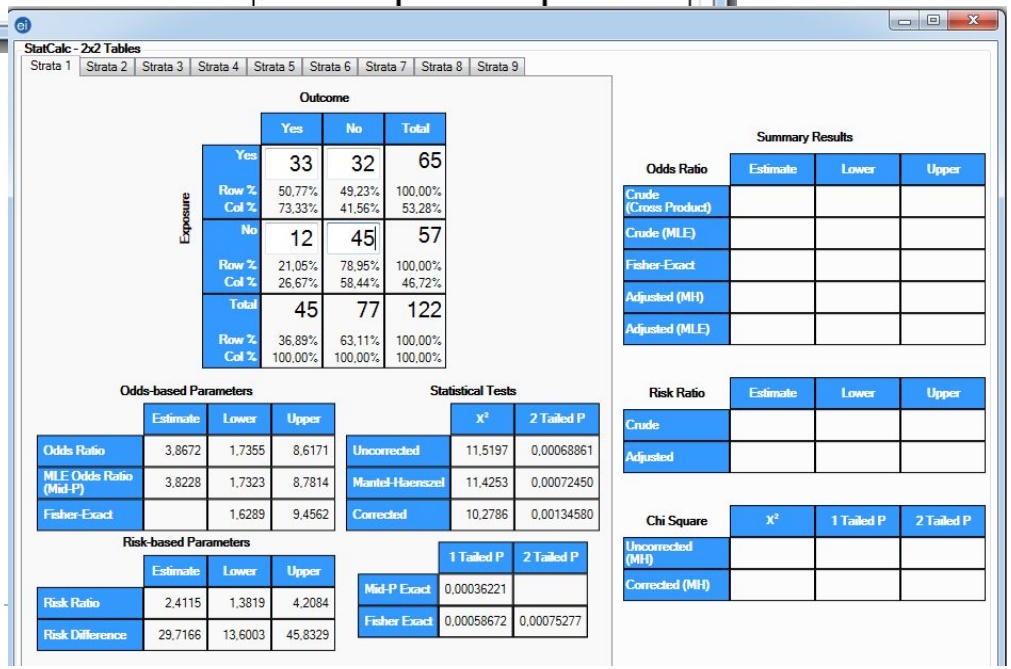
StatCalc - Sample Size and Power

Population survey or descriptive study
For simple random sampling, leave design effect and clusters equal to 1.

Population size:	1427000
Expected frequency:	50 %
Acceptable Margin of Error:	5 %
Design effect:	1,0
Clusters:	1

Confidence Level	Cluster Size	Total Sample
80%	164	164
90%	271	271
95%	384	384
97%	471	471
99%	663	663
99.9%	1082	1082
99.99%	1513	1513

POPULATION SURVEY



StatCalc - 2x2 Tables

Strata 1 | Strata 2 | Strata 3 | Strata 4 | Strata 5 | Strata 6 | Strata 7 | Strata 8 | Strata 9

Outcome

Exposure	Yes	No	Total
Yes	33	32	65
Row %	50,77%	49,23%	100,00%
Col %	73,33%	41,56%	53,28%
No	12	45	57
Row %	21,05%	78,95%	100,00%
Col %	26,67%	58,44%	46,72%
Total	45	77	122
Row %	36,89%	63,11%	100,00%
Col %	100,00%	100,00%	100,00%

Odds-based Parameters

	Estimate	Lower	Upper
Odds Ratio	3,8672	1,7355	8,6171
MLE Odds Ratio (Mid-P)	3,8228	1,7323	8,7814
Fisher-Exact		1,6289	9,4562

Statistical Tests

	X^2	2 Tailed P
Uncorrected	11,5197	0,00068861
Mantel-Haenszel	11,4253	0,00072450
Corrected	10,2786	0,00134580

Risk-based Parameters

	Estimate	Lower	Upper
Risk Ratio	2,4115	1,3819	4,2084
Risk Difference	29,7166	13,6003	45,8329

	1 Tailed P	2 Tailed P
Mid-P Exact	0,00036221	
Fisher Exact	0,00058672	0,00075277

Summary Results

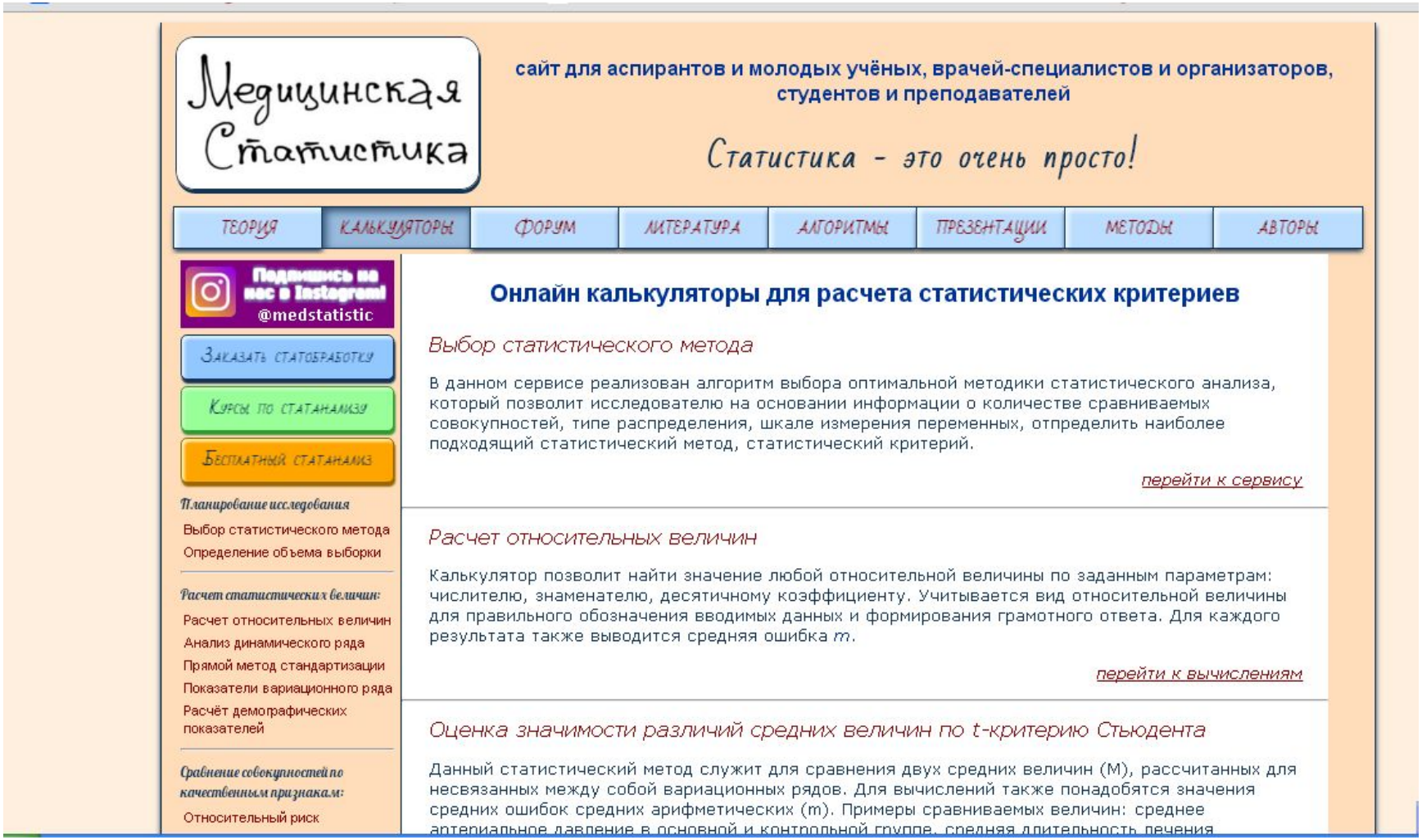
Odds Ratio	Estimate	Lower	Upper
Crude (Cross Product)			
Crude (MLE)			
Fisher-Exact			
Adjusted (MH)			
Adjusted (MLE)			

Risk Ratio	Estimate	Lower	Upper
Crude			
Adjusted			

Chi Square	X^2	1 Tailed P	2 Tailed P
Uncorrected (MH)			
Corrected (MH)			

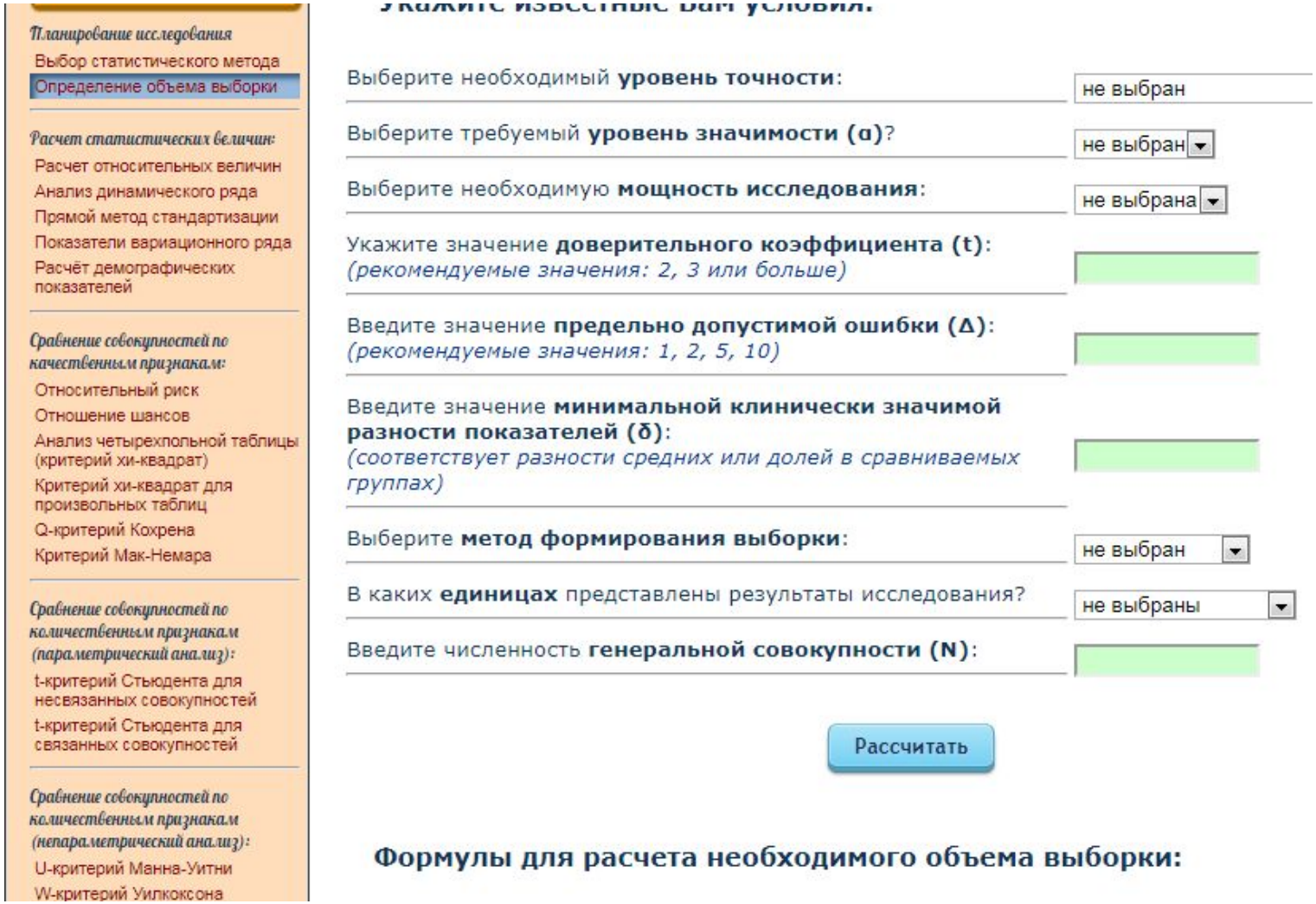
TABLES (2 x 2 x N)

Он-лайн калькулятор

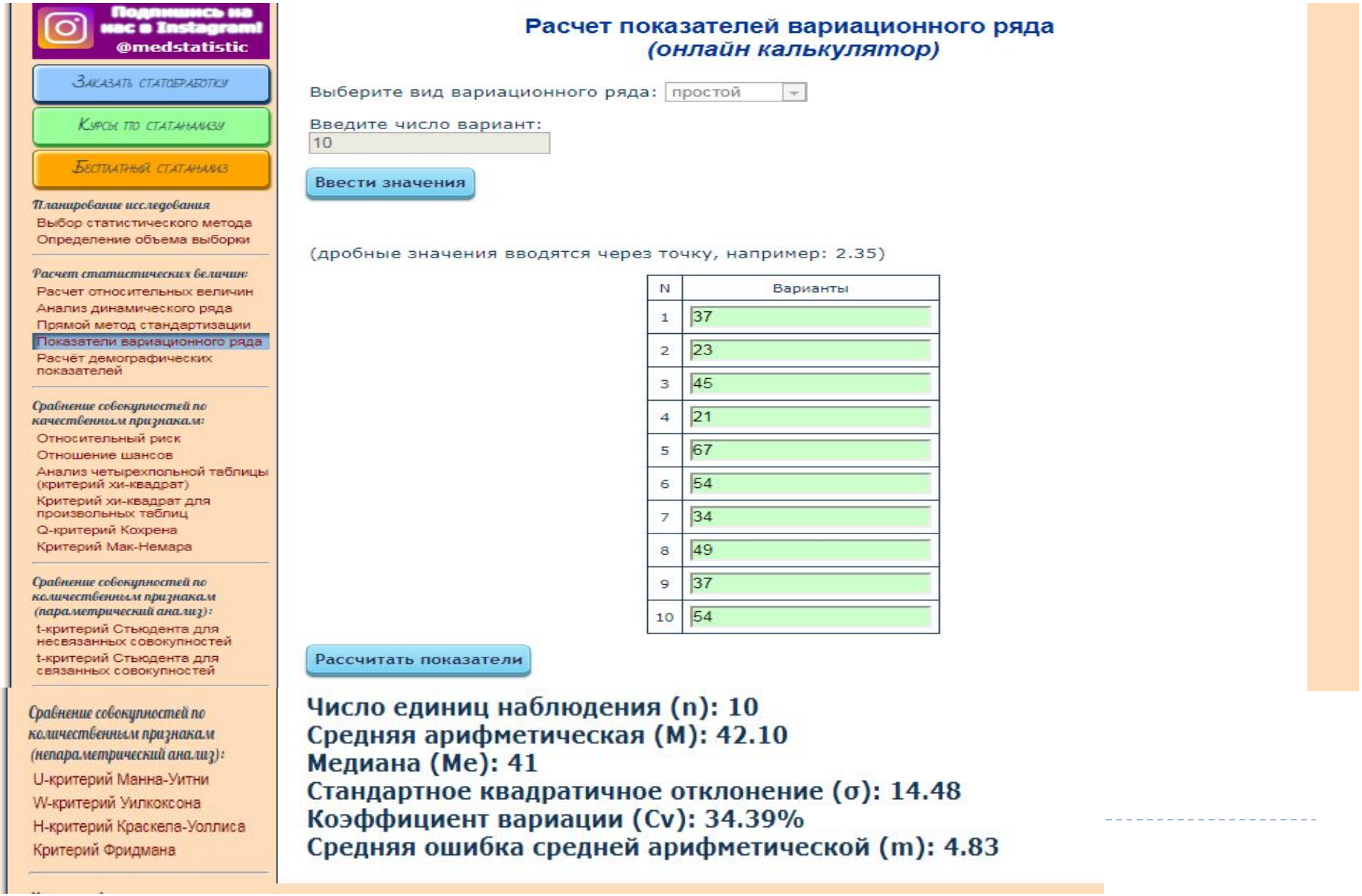


http://www.medstatistic.ru/calculators/averagestudent.html

Он-лайн калькулятор Медицинская статистика



Он-лайн калькулятор Медицинская статистика



Он-лайн калькулятор Медицинская статистика

Планирование исследования
- Выбор статистического метода
- Определение объема выборки

Расчет статистических величин:
- Расчет относительных величин
- Анализ динамического ряда
- Прямой метод стандартизации
- Показатели вариационного ряда
- Расчёт демографических показателей

Сравнение совокупностей по качественным признакам:
- Относительный риск
- Отношение шансов
- Анализ четырехпольной таблицы (критерий хи-квадрат)
- Критерий хи-квадрат для произвольных таблиц
- Q-критерий Кохрена
- Критерий Мак-Немара

Сравнение совокупностей по количественным признакам (параметрический анализ):
- t-критерий Стьюдента для несвязанных совокупностей
- t-критерий Стьюдента для связанных совокупностей

Сравнение совокупностей по количественным признакам (непараметрический анализ):
- U-критерий Манна-Уитни
- W-критерий Уилкоксона
- H-критерий Краскела-Уоллиса
- Критерий Фридмана

Изучение связи между признаками:

Введите число значений результативного признака (столбцов): 2

Обозначить признаки

Значения факторного признака:

Фактор1
Фактор2

Значения результативного признака:

Результат1
Результат2

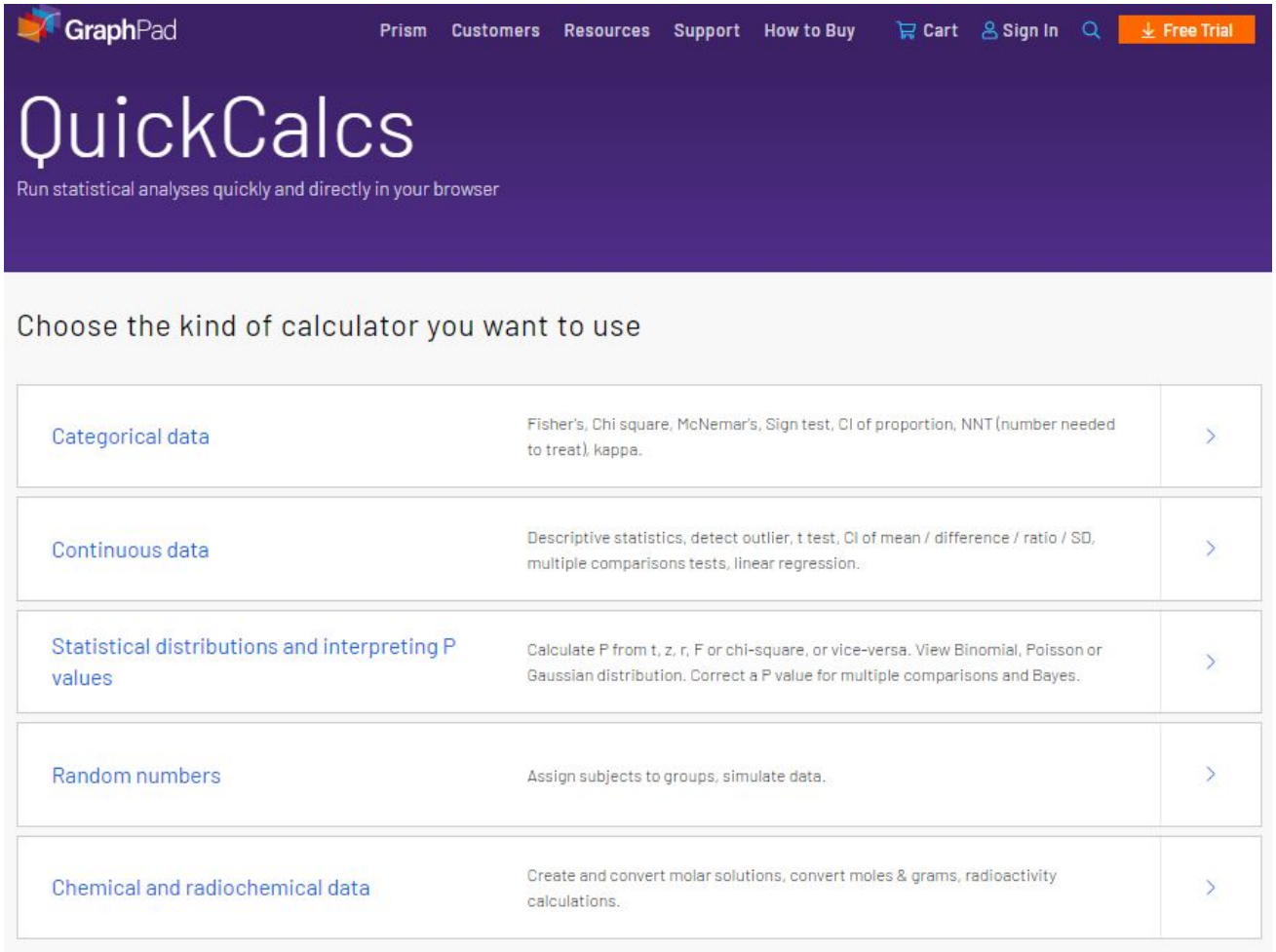
Ввести данные

Факторный признак	Результативный признак		Сумма
	Результат1	*Результат2*	
Фактор1	34	45	**79**
Фактор2	23	67	**90**
Всего	**57**	**112**	**169**

Рассчитать

Число степеней свободы равно 1
Значение критерия χ^2 составляет 5.753
Критическое значение χ^2 при уровне значимости p=0.05 составляет 3.841
Связь между факторным и результативным признаками статистически значима при уровне значимости p<0.05
Уровень значимости p=0.017

Он-лайн калькулятор



https://www.graphpad.com/quickcalcs/

raphPad Prism Customers Resources Support How to Buy

QuickCalcs

1. Select category 2. Choose calculator **3. Enter data** 4. View results

Descriptive statistics and CI of mean

Enter raw data and this calculator will calculate the mean, SD, SEM and confidence interval of the mean. Enter mean, N and SD or SEM, and it will calculate the confidence interval of the mean.

1. Choose data entry format

- (•) Enter up to 50 rows.
- ( ) Enter or paste up to 10000 rows.
- ( ) Enter mean, SEM and N.
- ( ) Enter mean, SD and N.

Caution: Changing format will erase your data.

2. Enter data

1:	22
2:	43
3:	21
4:	45
5:	32
6:	56
7:	32
8:	12
9:	34
10:	32
11:	65
12:	54
13:	67
14:	54
15:	36
16:	47
17:	

3. View the results

Calculate now

Clear the form

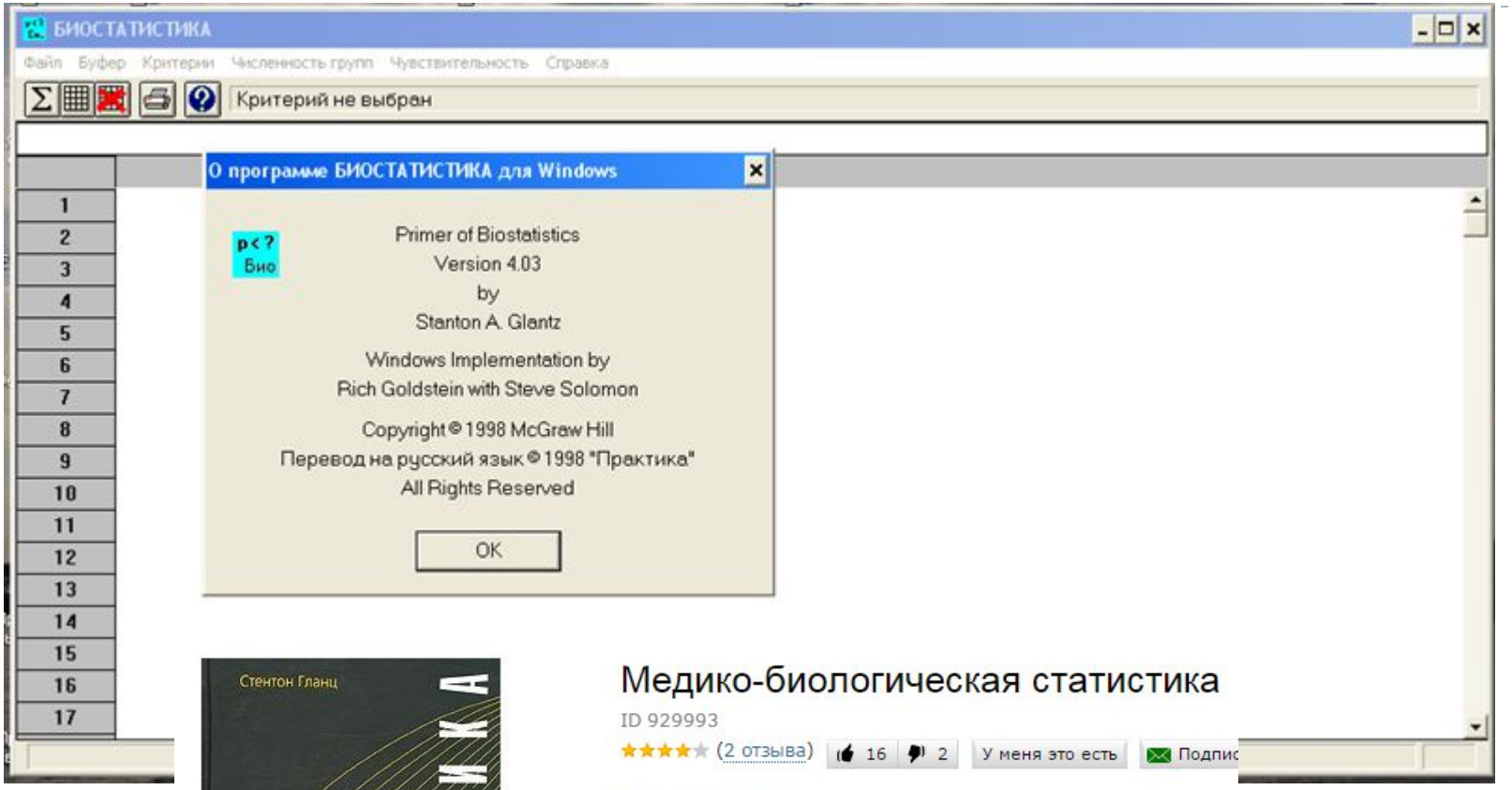
QuickCalcs

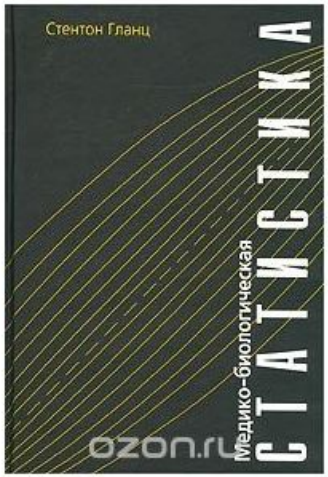
1. Select category 2. Choose calculator

Descriptive statistics results

Parameter	Value
Mean	40.75
SD	15.89
SEM	3.97
N	16
90% CI	33.78 to 47.72
95% CI	32.28 to 49.22
99% CI	29.04 to 52.46
Minimum	12
Median	39.5
Maximum	67

Biostat.exe





Медико-биологическая статистика

ID 929993

★★★★☆ (2 отзыва) 16 2 У меня это есть Подпис

Primer of Biostatistics

Автор: Стентон Гланц

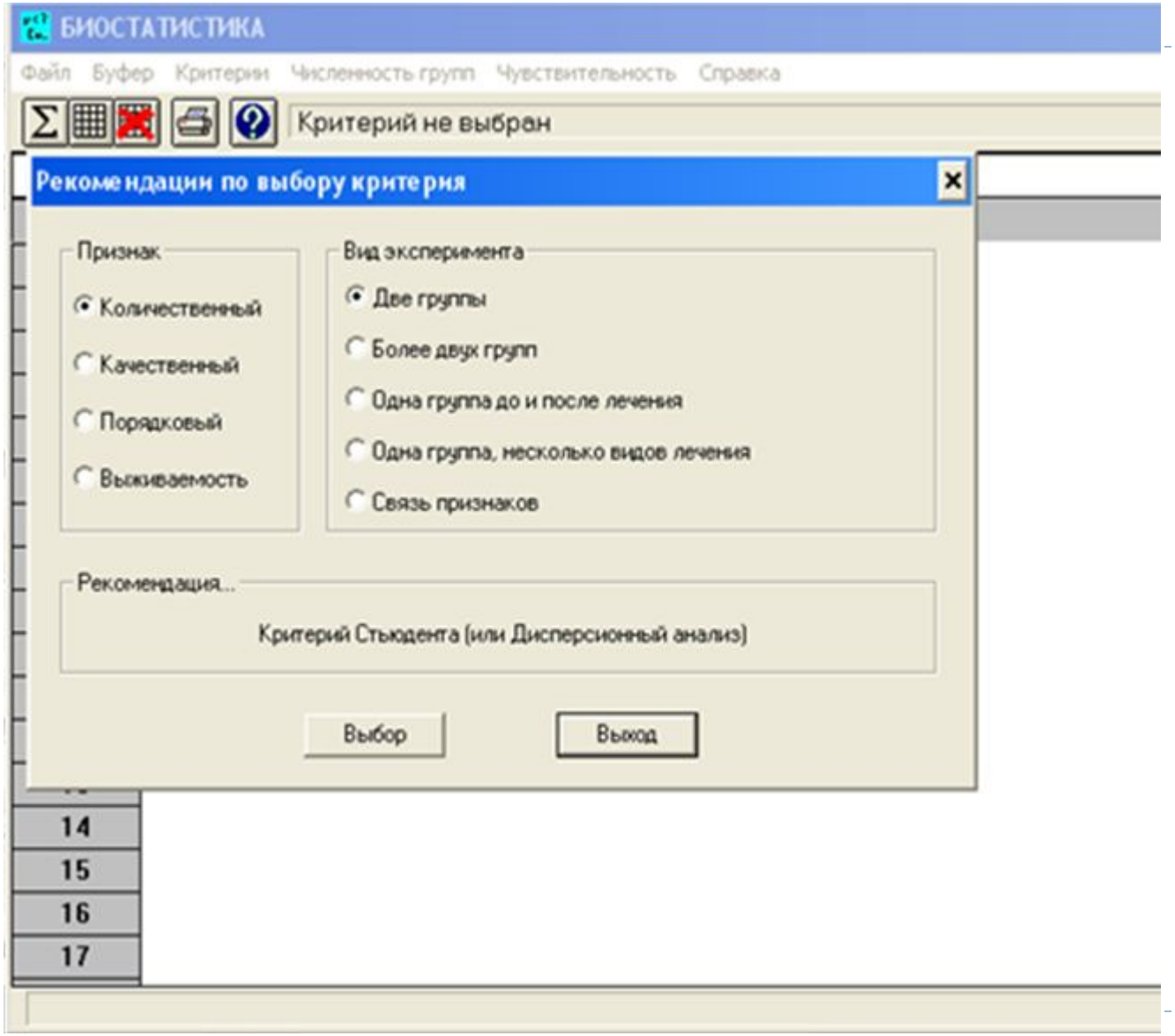
Переводчики: Юлий Данилов, Н. Бузикашвили, Д. Самойлов

Языки: Русский

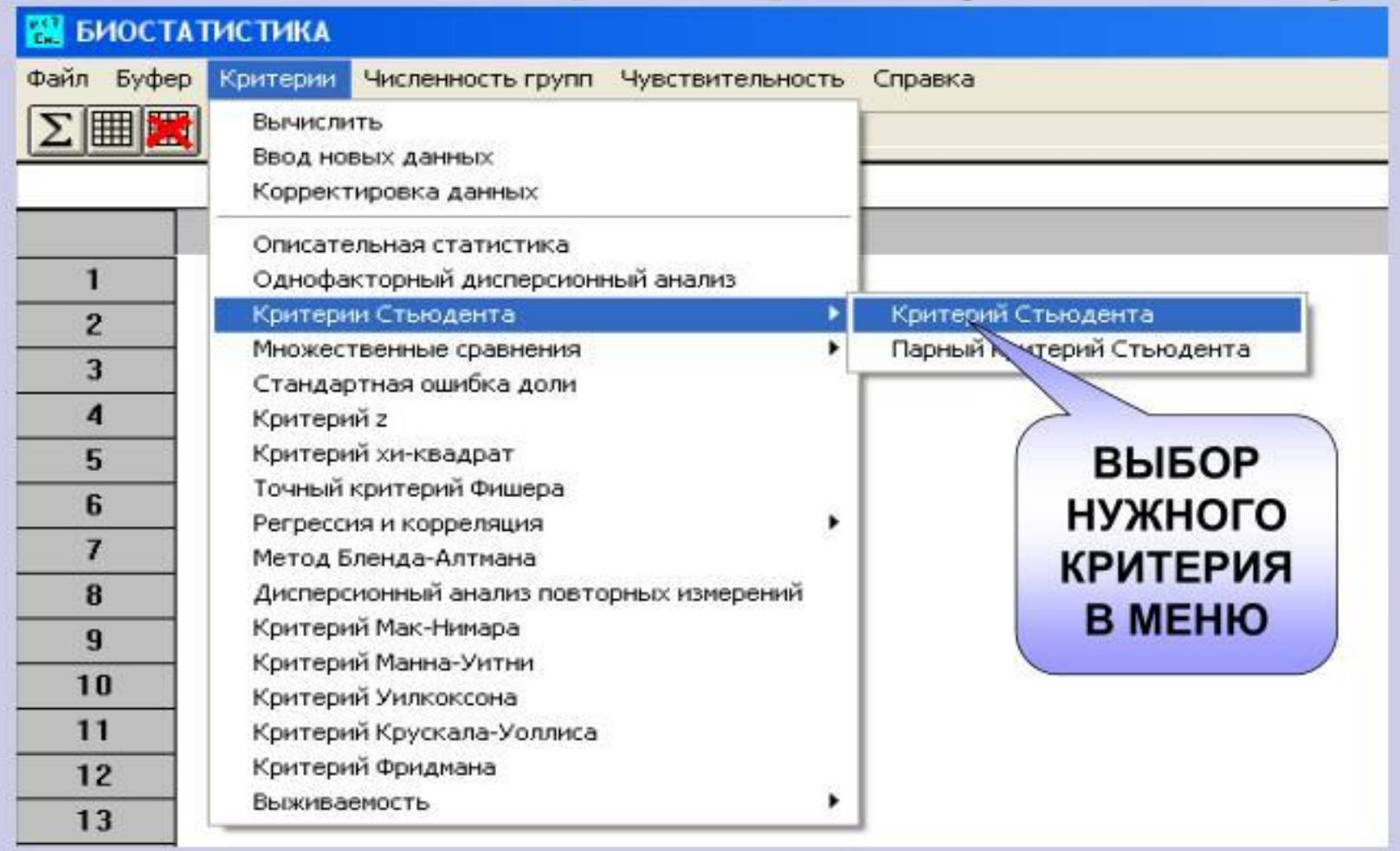
Издательство: Практика

ISBN 5-89816-009-4, 0-07-024268-2; 1998 г.

Дополнительные характеристики ▾

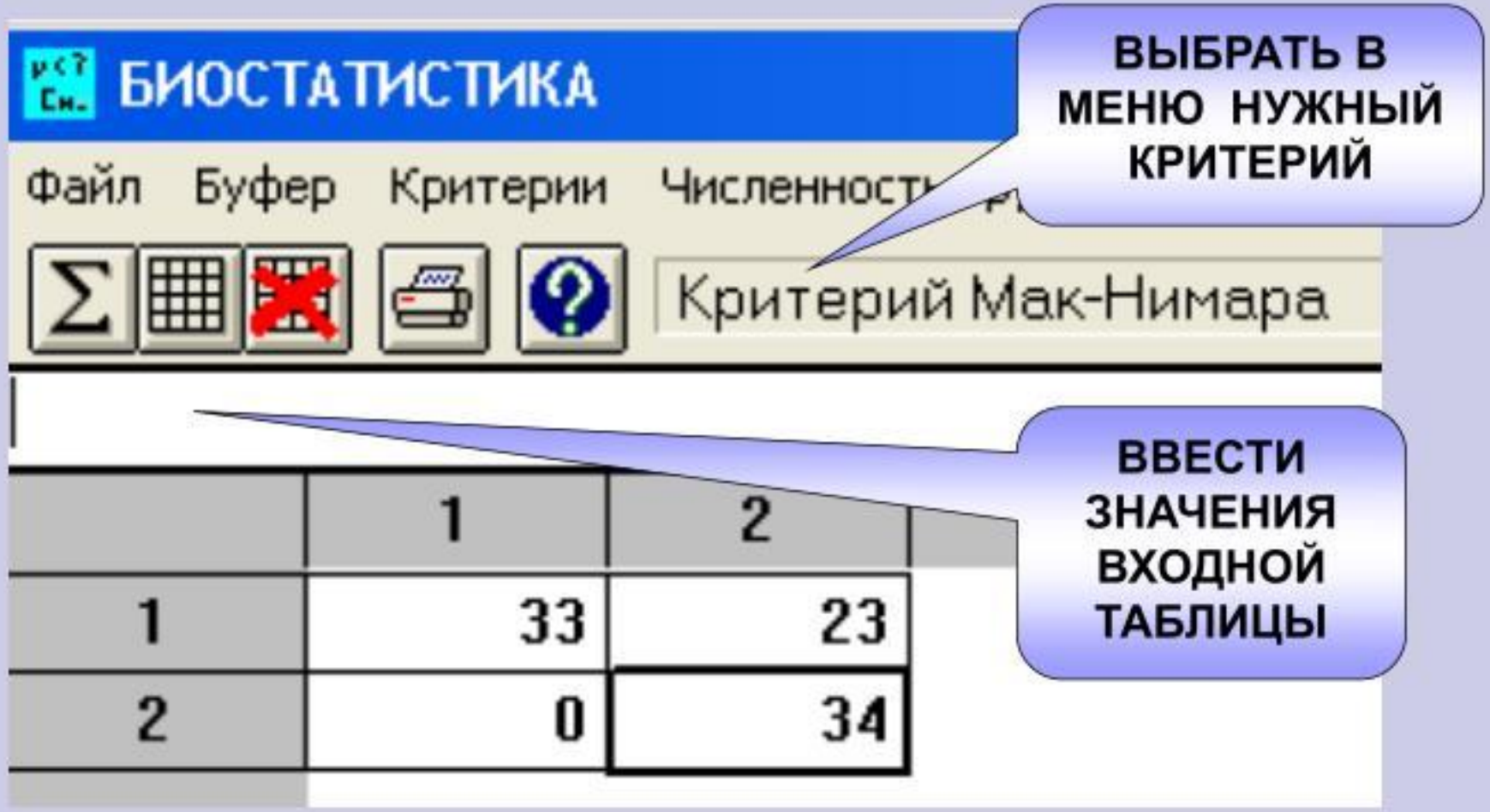

БИОСТАТИСТИКА
Файл Буфер Критерии Численность групп Чувствительность Справка
Критерий не выбран
Рекомендации по выбору критерия
Признак
Количественный
Качественный
Порядковый
Выживаемость
Вид эксперимента
Две группы
Более двух групп
Одна группа до и после лечения
Одна группа, несколько видов лечения
Связь признаков
Рекомендация...
Критерий Стьюдента (или Дисперсионный анализ)
Выбор
Выход
14
15
16
17

Применение стат. пакетов для вычисления критерия (Биостат):



Ввод данных для расчета критерия Стьюдента:

Ввод данных в пакете Биостат:



Решение. Ввод данных в пакете Биостатистика:

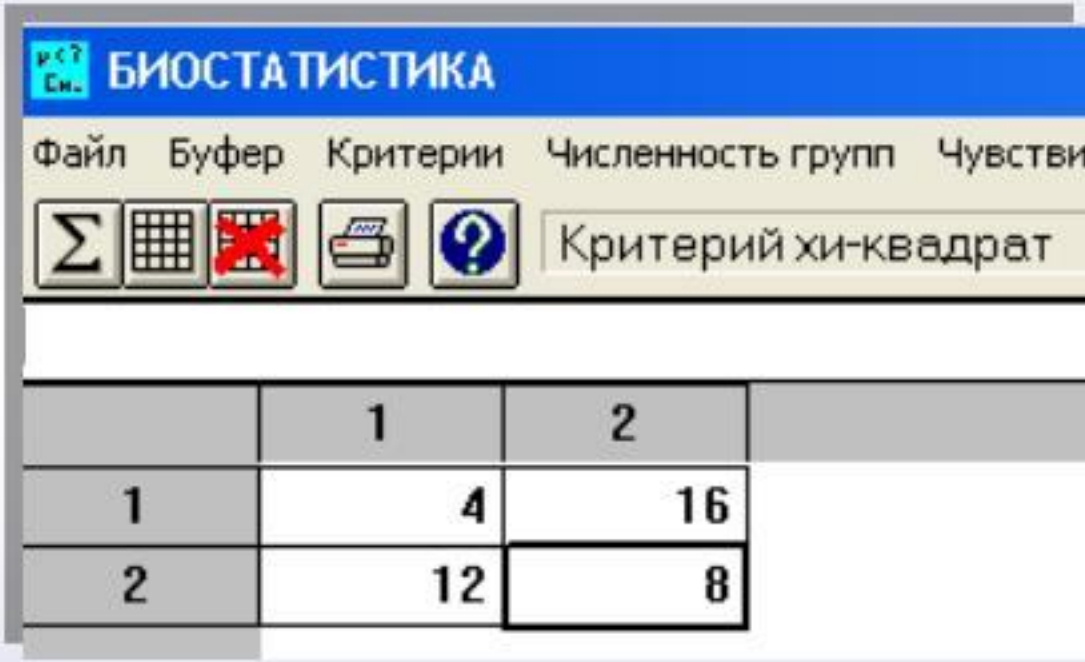
	X	Y
1	19.5	0
2	15	38.5
3	13.5	59
4	23.3	97.4
5	6.3	119.2
6	2.5	129.5
7	13	191.7
8	1.8	248.7
9	6.5	318
10	1.8	438.5

Решение задачи в пакете Биостатистика:

Признак качественный («заболел» с возможными значениями «да» или «нет»), две независимых группы. Следовательно, необходимо воспользоваться критерием χ2.

Важно: в условии задачи дана численность всей группы и число заболевших.

Во входную таблицу для пакета «Биостатистика» необходимо по группам ввести число *заболевших* и *не заболевших* (как разность между общей численностью группы и числом заболевших).



	1	2
1	4	16
2	12	8

Для таблиц сопряженности 2×2 в пакете автоматически вычисляется поправка Йейтса

Основная философская мысль