



Полякова Наталия Владимировна

**Кафедра ихтиологии и гидробиологии
факультета биологии СПбГУ, (16 линия ВО,
д 29)**

**Лаборатория поведения низших
позвоночных Института проблем
эволюции и экологии им. Северцова
(Москва)**

Зачем нужна статистика

- Что такое выборка
- Что такое среднее
- Что такое дисперсия
- Что такое статистическая ошибка
- Что такое доверительный интервал
- Что такое корреляция
- Что такое статистический критерий
- Что такое доверительная вероятность

Допустим, вы хотите опубликовать в своей статье что-нибудь простое, вроде

$$1 + 1 = 2.$$

Конечно, это будет выглядеть несолидно, поэтому правую часть можно представить в виде $\sum_{i=0}^{\infty} \frac{1}{2^i} = 2$, одну из единиц представим в виде $1 = \ln(e)$, а вторую единицу $1 = \cos^2(\sqrt{\pi}) + \sin^2(\sqrt{\pi})$, теперь получится что-то вроде

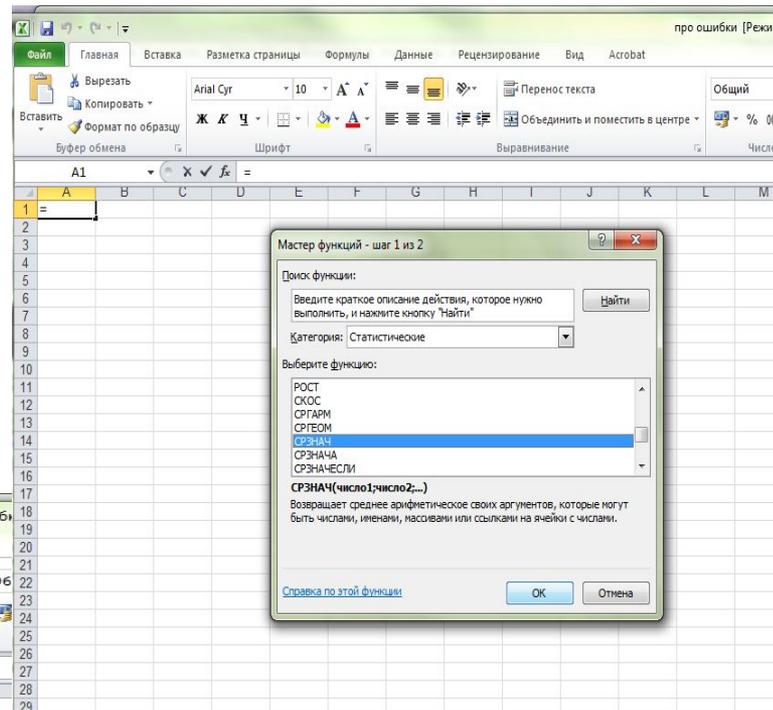
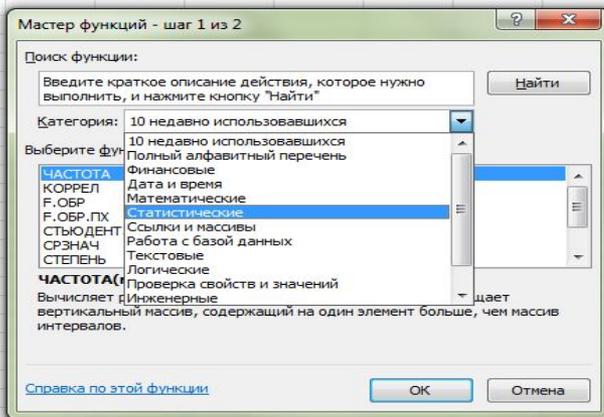
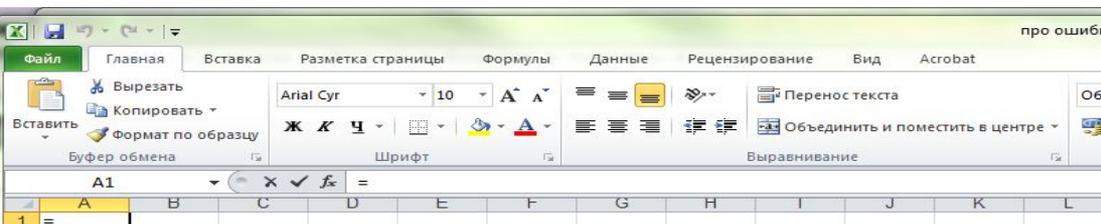
$$\ln(e) + \cos^2(\sqrt{\pi}) + \sin^2(\sqrt{\pi}) = \sum_{i=0}^{\infty} \frac{1}{2^i} = 2.$$

Конечно, такая сложность подойдёт для некоторых статей, но для продвинутого журнала и это следует немного усложнить. Пусть $e = \lim_{q \rightarrow \infty} \left(1 + \frac{1}{q}\right)^q$, $\sqrt{\pi} = \int_{-\infty}^{+\infty} e^{-x^2} dx$, теперь окончательно:

$$\ln\left(\lim_{q \rightarrow \infty} \left(1 + \frac{1}{q}\right)^q\right) + \cos^2\left(\int_{-\infty}^{+\infty} e^{-x^2} dx\right) + \sin^2\left(\int_{-\infty}^{+\infty} e^{-x^2} dx\right) = \sum_{i=0}^{\infty} \frac{1}{2^i}.$$

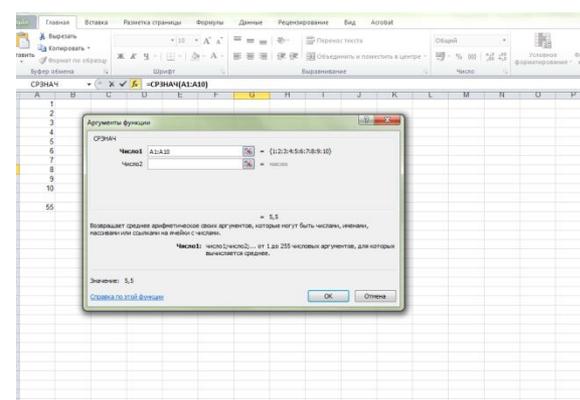
СРЕДНЯЯ АРИФМЕТИЧЕСКАЯ

$$\bar{x} = \frac{\sum x_i}{n}$$



- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 55

=55/10



Средняя= 5,5

*** СТРУКТУРНЫЕ СРЕДНИЕ ***

Мода

величина, наиболее часто встречающаяся в вариационном ряду

Медиана

это средняя, относительно которой ряд распределения делится на две равные части

1 2 3 4 5

1 2 2 3 3 3 4 4 5

1 2 2 2 2 2 2 2 2 2 3 3 3 4 4 5

*** ПОКАЗАТЕЛИ ВАРИАЦИИ ***

РАЗМАХ ВАРИАЦИИ

$R = X_{\max} - X_{\min}$

10 15 20 25 30 35 40 45 50

10 28 28 30 30 30 32 32 50

ДИСПЕРСИЯ
(ВАРИАНСА)

$$s^2 = \frac{\sum (x_i - M)^2}{n - 1}$$

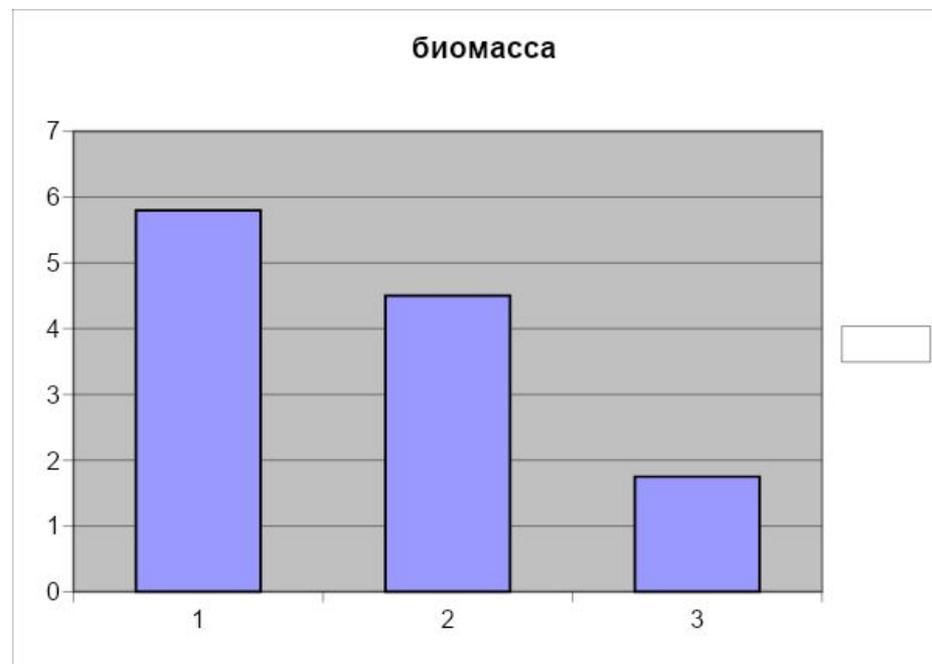
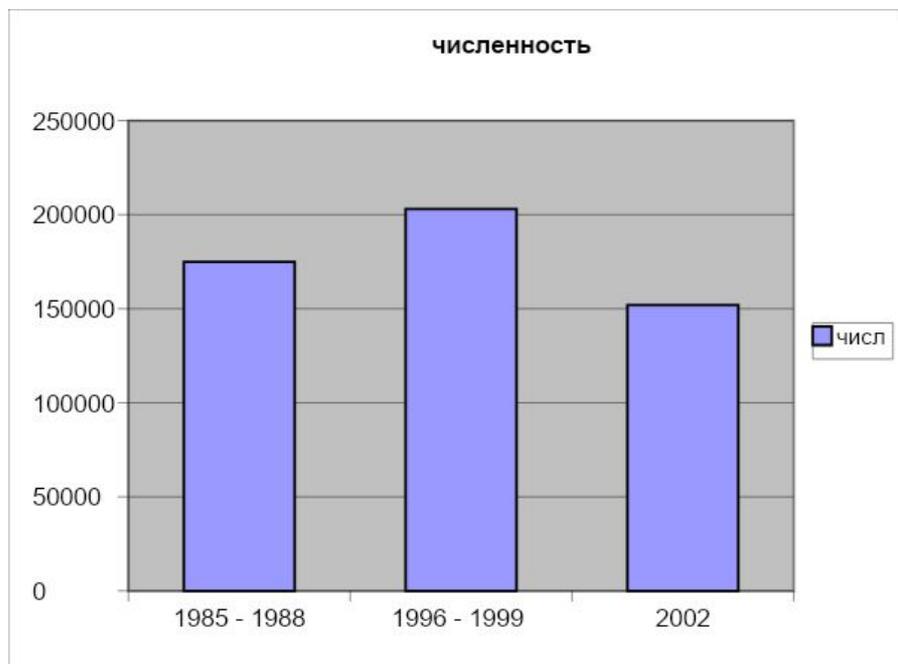
СРЕДНЕЕ КВАДРАТИЧЕСКОЕ
ОТКЛОНЕНИЕ
(стандартное отклонение)

$$s = \sqrt{s^2}$$

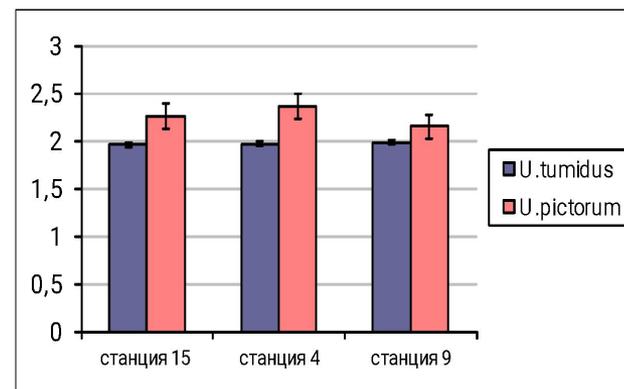
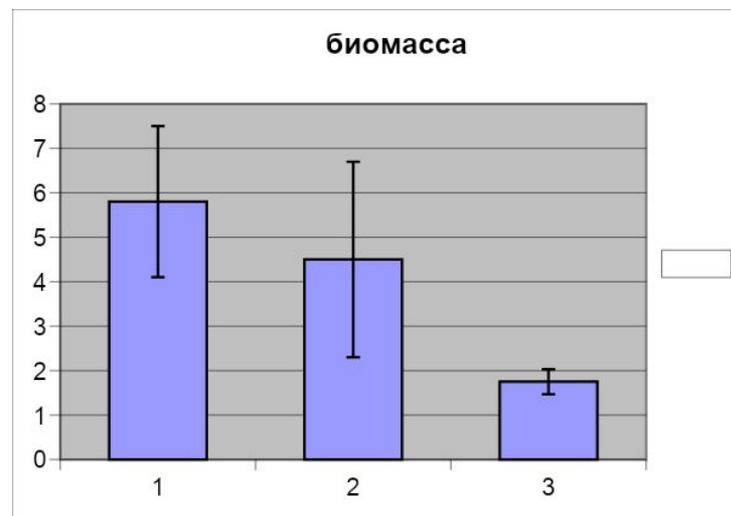
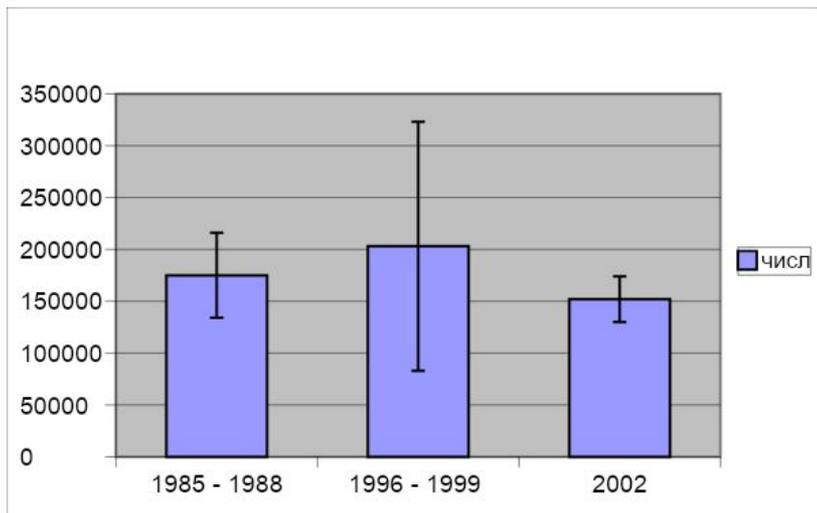
КОЭФФИЦИЕНТ
ВАРИАЦИИ

$$C_v = \frac{s}{M} \cdot 100\%$$

период	числ										биомасса
1985 - 198	175000									1985 - 198	5,8
1996 - 199	203000									1996 - 199	4,5
2002	152000									2002	1,75



период	числ	ошибка									биомасса		
1985 - 1988	175000	41000									1985 - 1988	5,8	1,7
1996 - 1999	203000	120000									1996 - 1999	4,5	2,2
2002	152000	22000									2002	1,75	0,28



Ошибка средней арифметической (M) (стандартная ошибка)

$$m_x = \frac{s}{\sqrt{n}}$$

 Сохранить Сохранить как Открыть Закреть

Сведения

Последние

Создать

Печать

Сохранить и отправить

Справка

 Параметры Выход

Параметры Excel

Общие

Формулы

Правописание

Сохранение

Язык

Дополнительно

Настройка ленты

Панель быстрого доступа

Настройки

Центр управления безопасностью



Управление надстройками Microsoft Office.

Надстройки

Имя ^	Расположение	Тип
ABBYY PDF Transformer 2.0 MSEXCEL COM Add-In	C:\...ddInMSO.dll	Надстр
Пакет анализа	C:\...NALYS32.XLL	Надстр

Надстройка: Пакет анализа
Издатель: Microsoft Corporation
Совместимость: Отсутствуют сведения о совместимости
Расположение: C:\Program Files\Microsoft Office\Office14\Library\Analysis\ANALYS32.XLL
Описание: Содержит инструменты для анализа научных и финансовых данных

Управление: Надстройки Excel

Перейти...

OK

Отмена



Поиск проблем

Подготовить к общему доступу

Перед распространением общего доступа к этому файлу необходимо учесть, что он содержит:

- Свойства документа, путь к принтеру и имя автора
- Содержимое, в котором невозможно выполнить проверку читаемости типа текущего файла

 Сохранить Сохранить как Открыть Закреть

Сведения

Последние

Создать

Печать

Сохранить и отправить

Справка

 Параметры Выход

Параметры Excel

Общие

Формулы

Правописание

Сохранение

Язык

Дополнительно

Настройка ленты

Панель бы

Надстрой

Центр уп



Управление надстройками Microsoft Office.

Надстройки

Имя ^	Расположение	Тип
ABBYY PDF Transformer 2.0 MSExcel.COM Add-In	C:\...ddToMSO.dll	Надстр
Пакет анализа	C:\...NALYS32.XLL	Надстр

Надстройка: Пакет анализа
Издатель: Microsoft CorporationСведения о совместимости
s\Microsoft Office\Office14\Library\Analysis\
рументы для анализа научных и финансовых

Перейти...

OK

Отмена

Надстройки

Доступные надстройки:

- Инструменты для евро
- Пакет анализа
- Пакет анализа - VBA
- Поиск решения

OK

Отмена

Обзор...

Автоматизация...

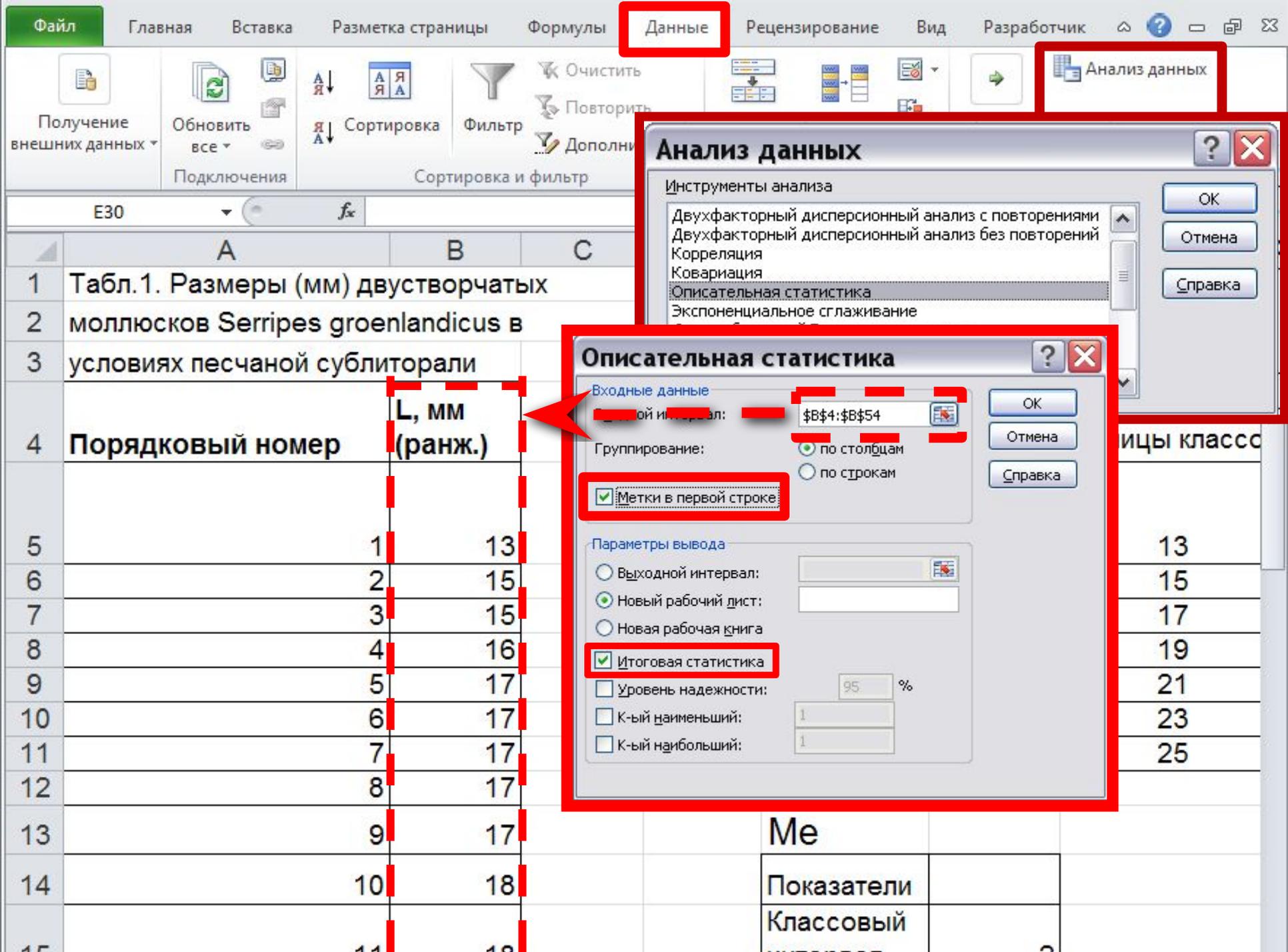
Пакет анализа

Содержит инструменты для анализа научных и финансовых данных

Файл Главная Вставка Разметка страницы Формулы **Данные** Рецензирование Вид Разработчик

Получение внешних данных Обновить все Подключения Сортировка Фильтр Сортировка и фильтр Очистить Повторить Дополнительно Текст по столбцам Удалить дубликаты Работа с данными Структура **Анализ данных** Анализ

	A	B	C	D	E	F	G
1	Табл.1. Размеры (мм) двустворчатых						
2	моллюсков <i>Serripes groenlandicus</i> в						
3	условиях песчаной сублиторали				Таблица 3. Интервальный вариационный		
4	Порядковый номер	L, мм (ранж.)			Границы классов		верхние границы классов
5		1	13		12-14	12-13	13
6		2	15		14-16	14-15	15
7		3	15		16-18	16-17	17
8		4	16		18-20	18-19	19
9		5	17		20-22	20-21	21
10		6	17		22-24	22-23	23
11		7	17		24-26	24-25	25
12		8	17				
13		9	17		Me		
14		10	18		Показатели		
15		11	18		Классовый		



Получение внешних данных

Обновить все



Сортировка

Фильтр

Очистить
Повторить



Анализ данных

Анализ данных

Инструменты анализа

- Двухфакторный дисперсионный анализ с повторениями
- Двухфакторный дисперсионный анализ без повторений
- Корреляция
- Ковариация
- Описательная статистика**
- Экспоненциальное сглаживание

OK

Отмена

Справка

Описательная статистика

Входные данные

Выходной интервал:

Группирование:

по столбцам

по строкам

Метки в первой строке

OK

Отмена

Справка

Параметры вывода

Выходной интервал:

Новый рабочий лист:

Новая рабочая книга

Итоговая статистика

Уровень надежности:

%

К-ый наименьший:

К-ый наибольший:

1 Табл.1. Размеры (мм) двустворчатых
2 моллюсков *Serripes groenlandicus* в
3 условиях песчаной сублиторали

4 **Порядковый номер** **L, мм (ранж.)**

5	1	13
6	2	15
7	3	15
8	4	16
9	5	17
10	6	17
11	7	17
12	8	17

ицы классо

13

15

17

19

21

23

25

Me

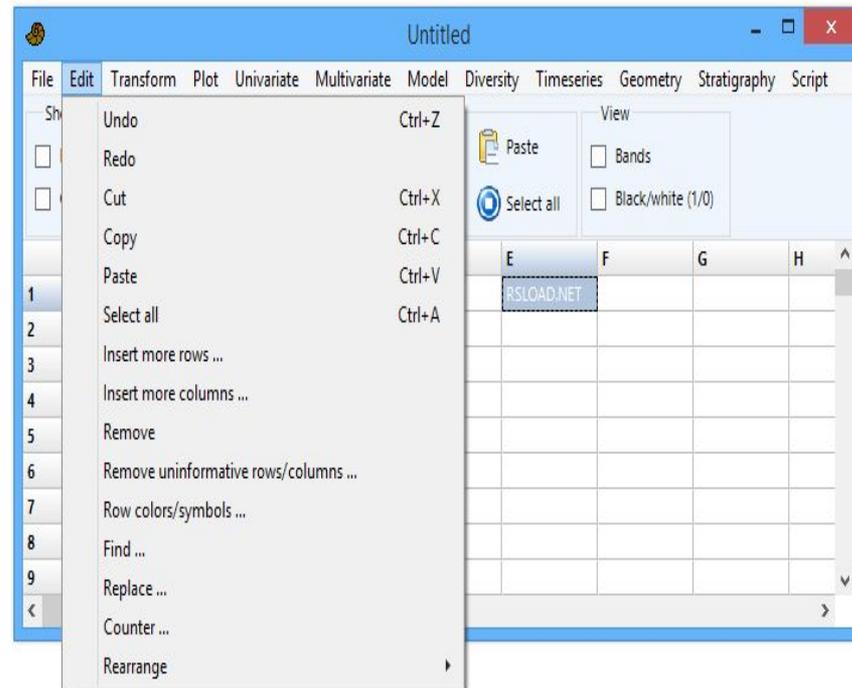
Показатели

Классовый

Столбец1

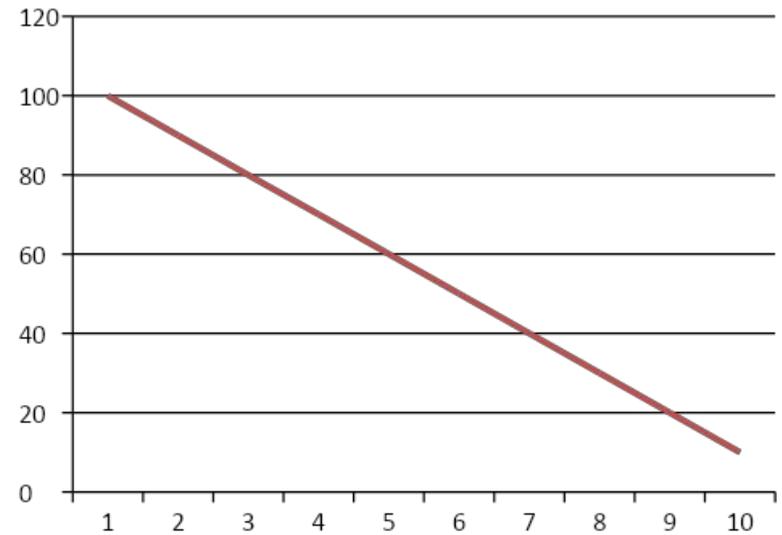
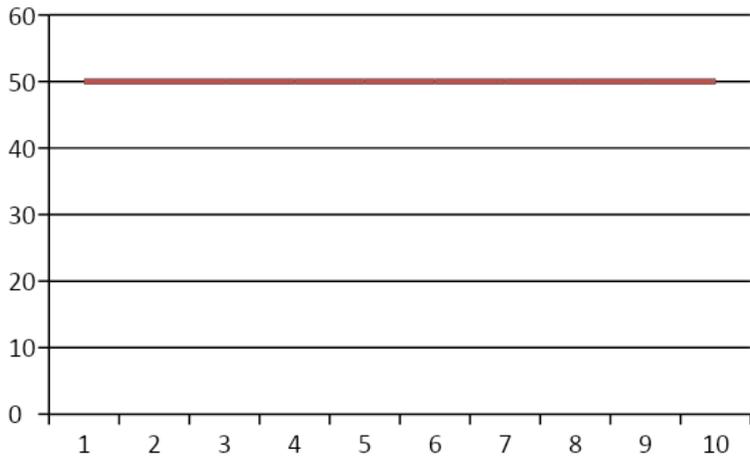
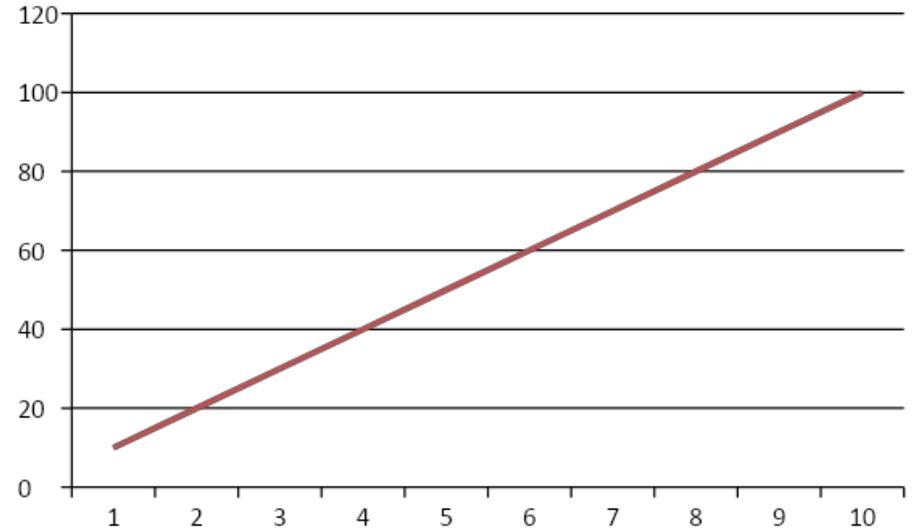
Past 3.x

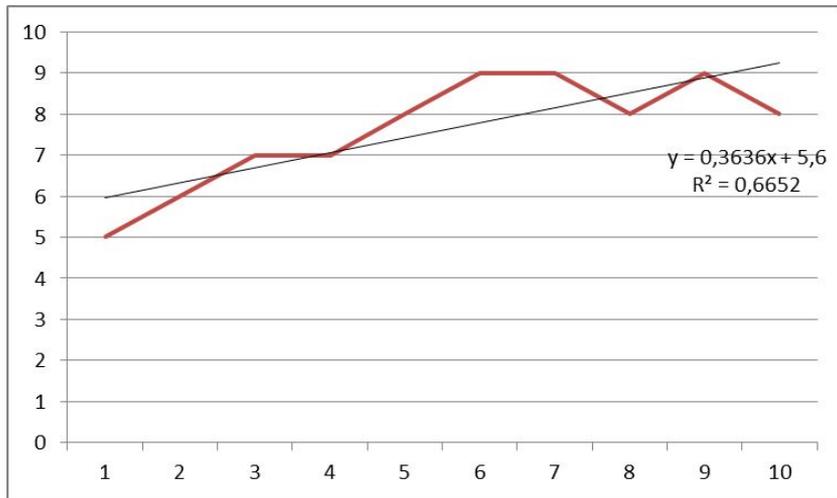
Среднее	3,5
Стандартная ошибка	0,763762616
Медиана	3,5
Мода	#Н/Д
Стандартное отклонение	1,870828693
Дисперсия выборки	3,5
Экссесс	-1,2
Асимметричность	0
Интервал	5
Минимум	1
Максимум	6
Сумма	21
Счет	6



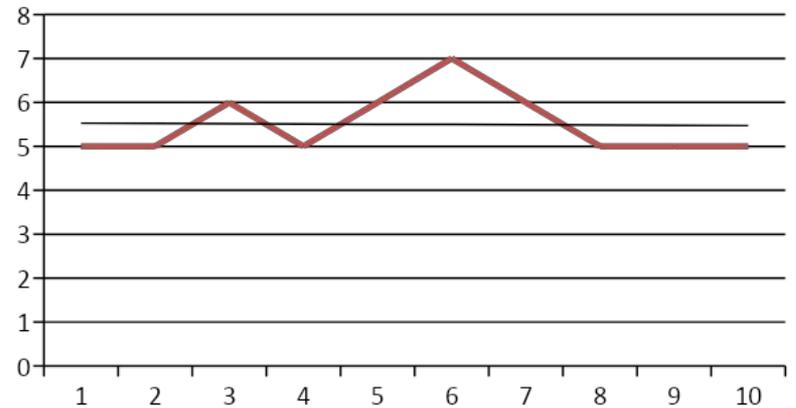
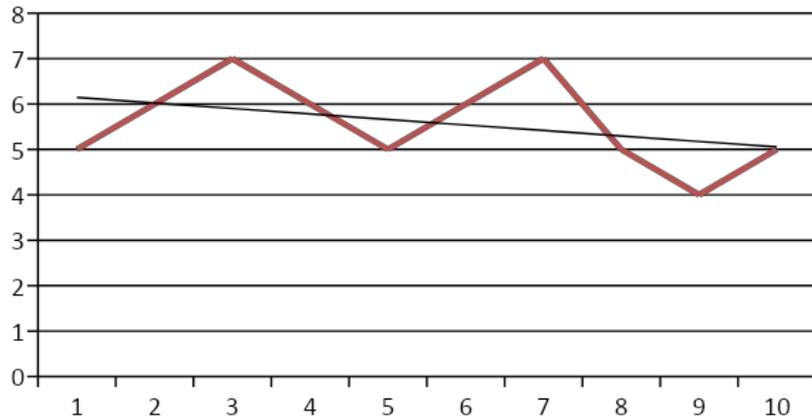
X	у1	у2	у3
1	10	100	50
2	20	90	50
3	30	80	50
4	40	70	50
5	50	60	50
6	60	50	50
7	70	40	50
8	80	30	50
9	90	20	50
10	100	10	50

Корреляция

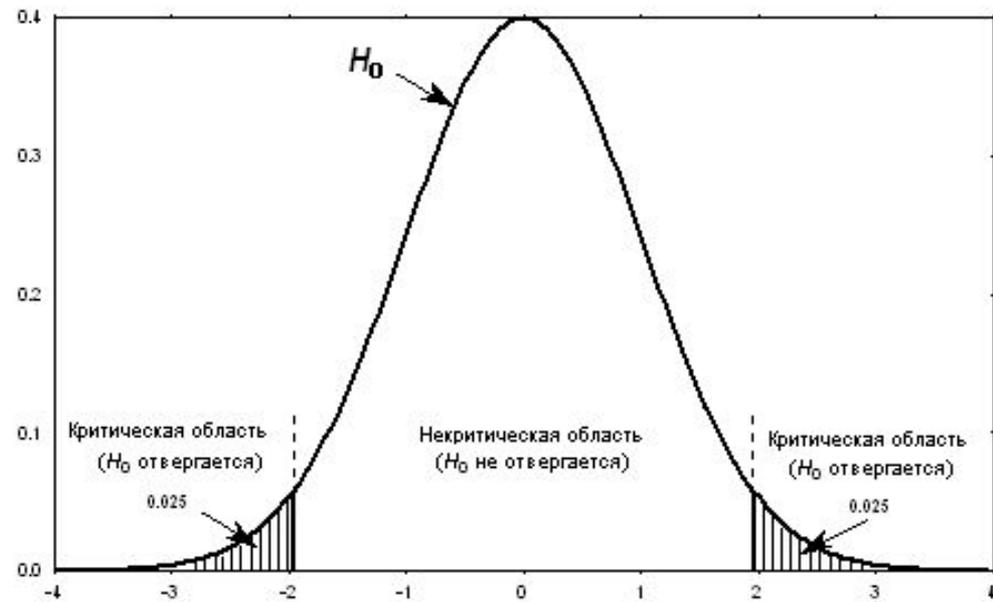
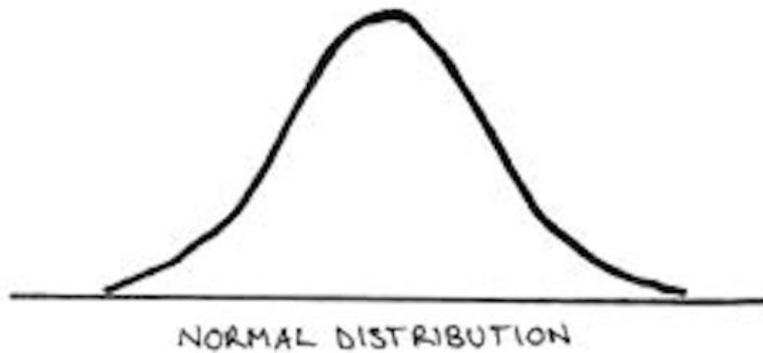




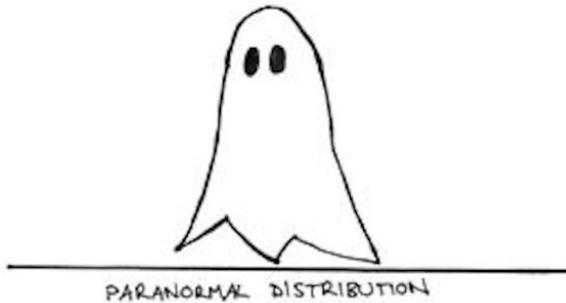
X	y1	y2	y3
1	5	5	5
2	6	6	5
3	7	7	6
4	7	6	5
5	8	5	6
6	9	6	7
7	9	7	6
8	8	5	5
9	9	4	5
10	8	5	5



Нормальное распределение

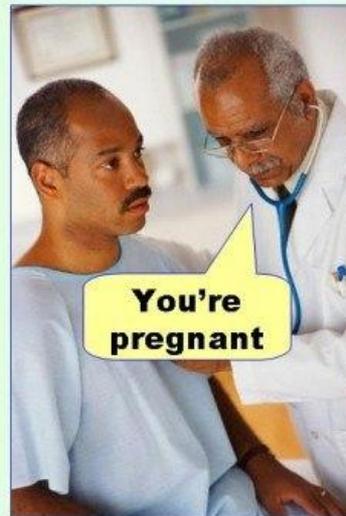


Паранормальное распределение

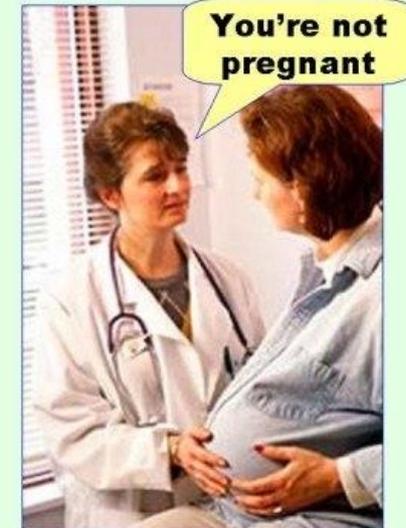


Freeman.

Type I error
(false positive)



Type II error
(false negative)



учебники

