**Лабораторная работа №1  
по курсу «Имитационное моделирование экономических процессов»  
на повторение теории вероятностей и математической статистики**

**Исходные данные:**

Рассматриваются две нормально распределенные случайные величины (СВ) *X* и *Y*, для которых заданы *истинное* математическое ожидание (МО) *M* и среднеквадратическое отклонение (СКО) .

Для каждой СВ имеется выборка объемом *n* и *p* наблюдений соответственно:

,

.

**Задание:**

1. Сгенерировать случайные величины X и Y с заданными характеристиками.
2. Рассчитать числовые характеристики X и Y: выборочное среднее *m* и стандартное отклонение *S*.
3. Построить гистограмму распределения и график теоретической плотности распределения для X и Y.

Построить график кумуляты и интегральной плотности распределения для X и Y.

1. Построить доверительный интервал оценки *mX* при заданном уровне значимости . Проверить попадание истинного значения в построенный доверительный интервал.
2. Оценить вероятность того, что а) X < W; б) Y < W
3. Сделать вывод по работе. Сравнить полученные результаты (между X и Y, между выборочными и теоретическими показателями).

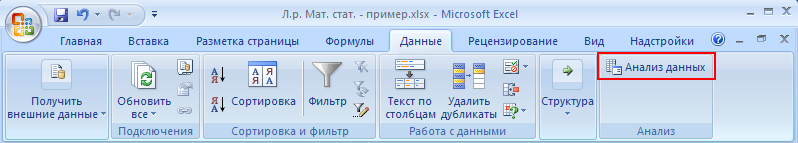
|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Вариант** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| ***n*** | 15 | 20 | 25 | 30 | 15 | 20 | 25 | 30 | 20 | 30 |
| ***p*** | 170 | 175 | 200 | 215 | 170 | 175 | 200 | 215 | 190 | 210 |
| ***M*** | 21 | 8 | 31 | 45 | 35 | 48 | 4 | 22 | 54 | 10 |
| **σ** | 5 | 4 | 6 | 12 | 15 | 8 | 1 | 9 | 6 | 3 |
|  | 0,01 | 0,05 | 0,1 | 0,01 | 0,05 | 0,1 | 0,01 | 0,05 | 0,1 | 0,001 |
| W | 18 | 0 | 15 | 30 | 5 | 20 | 0 | 16 | 35 | 1 |

## Указания к выполнению работы

Лабораторная работа выполняется на приготовленном шаблоне “Л.р. Имитац. мод. 1.xlsx”.

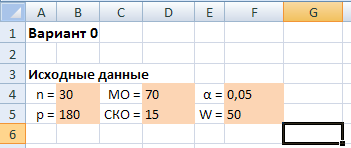
1. Генерация случайных чисел.

Для генерации случайных чисел с разными законами распределения в Excel используется «Пакет анализа».



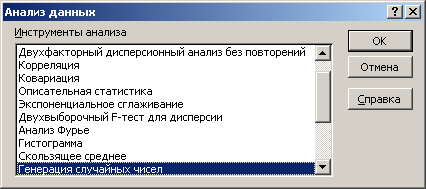
По умолчанию «Пакет анализа» отключен. В приложении описано, как его включить.

Прежде, чем приступить к генерации, введите исходные данные для своего варианта на лист.

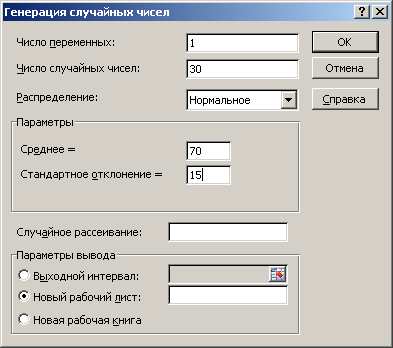


В примере требуется сгенерировать 30 значений X с нормальным распределением с МО = 70 и СКО = 15 и 180 значений Y с теми же параметрами.

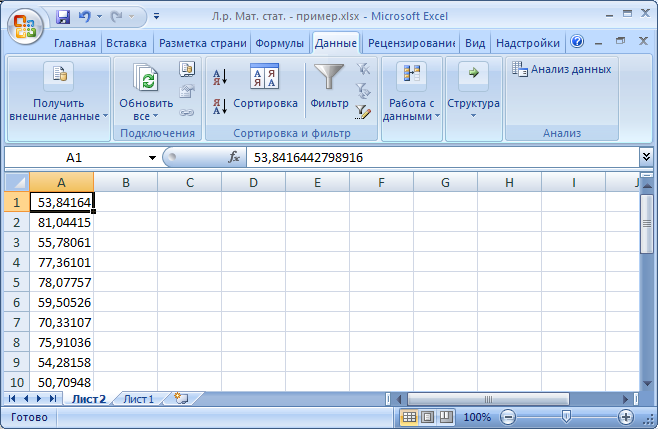
Нажмите кнопку «Анализ данных». В открывшемся окне выберите «Генерация случайных чисел».



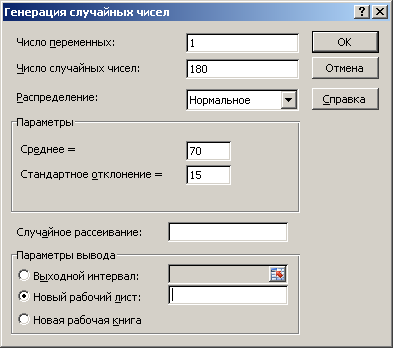
Введите исходные данные для X из своего варианта.



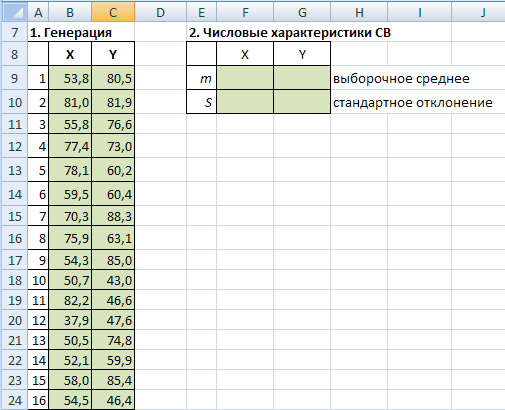
После нажатия OK будет создан новый лист со значениями X. Аналогично выполним генерацию значений Y.



Аналогично сгенерируйте значения Y.



Скопируйте полученные значения в шаблон. Используйте «Вставку значений», чтобы не нарушать оформление. Пронумеруйте значения в столбце А.



1. Расчет **выборочного среднего** и **стандартного отклонения**:

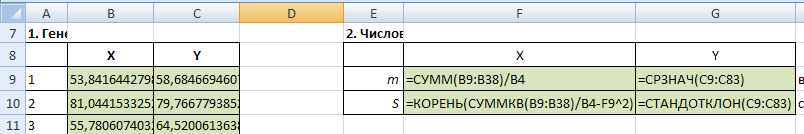
,

.

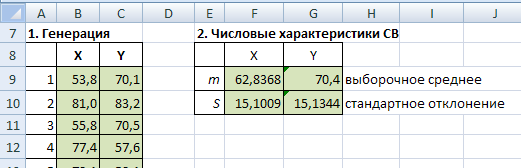
Стандартное отклонение является мерой разброса значений выборки относительно среднего. Однако это **не** означает, что, например, при  и  все значения будут находиться в диапазоне от 130 до 170.

*m* является выборочной оценкой МО, *S* – оценкой СКО.

Рассчитайте числовые характеристики двумя способами: через суммирование и с использованием встроенных функций СРЗНАЧ и СТАНДОТКЛОН.



В примере получим:



В данном случае более близкое к истинному значение среднего получилось у Y, а более близкое к истинному S – у X, но это исключение. В большинстве случаев, при увеличении числа генерируемых значений их характеристики приближаются к заданным. Разница была бы более заметной при сравнении, скажем 100 и 100 000 генераций случайного числа.

1. Построение **гистограммы распределения.**

Гистограмма распределения показывает, насколько часто встречаются те или иные значения СВ (выборочный аналог плотности распределения). Каждый столбец гистограммы показывает частоту попадания значения выборки в интервал значений.

Аналогом интегральной функции распределения является **кумулята** **–** гистограмма «с накоплением».

Для нее рассчитывается не число попаданий в каждый интервал, а число значений, меньших, чем верхняя граница интервала, т.е. суммарное (накопленное, кумулятивное) число попаданий во все интервалы с первого по текущий:



В первую очередь необходимо разбить весь диапазон исходных значений на *m* интервалов и определить их ширину.

Число столбцов гистограммы распределения (формула Стерджесса):

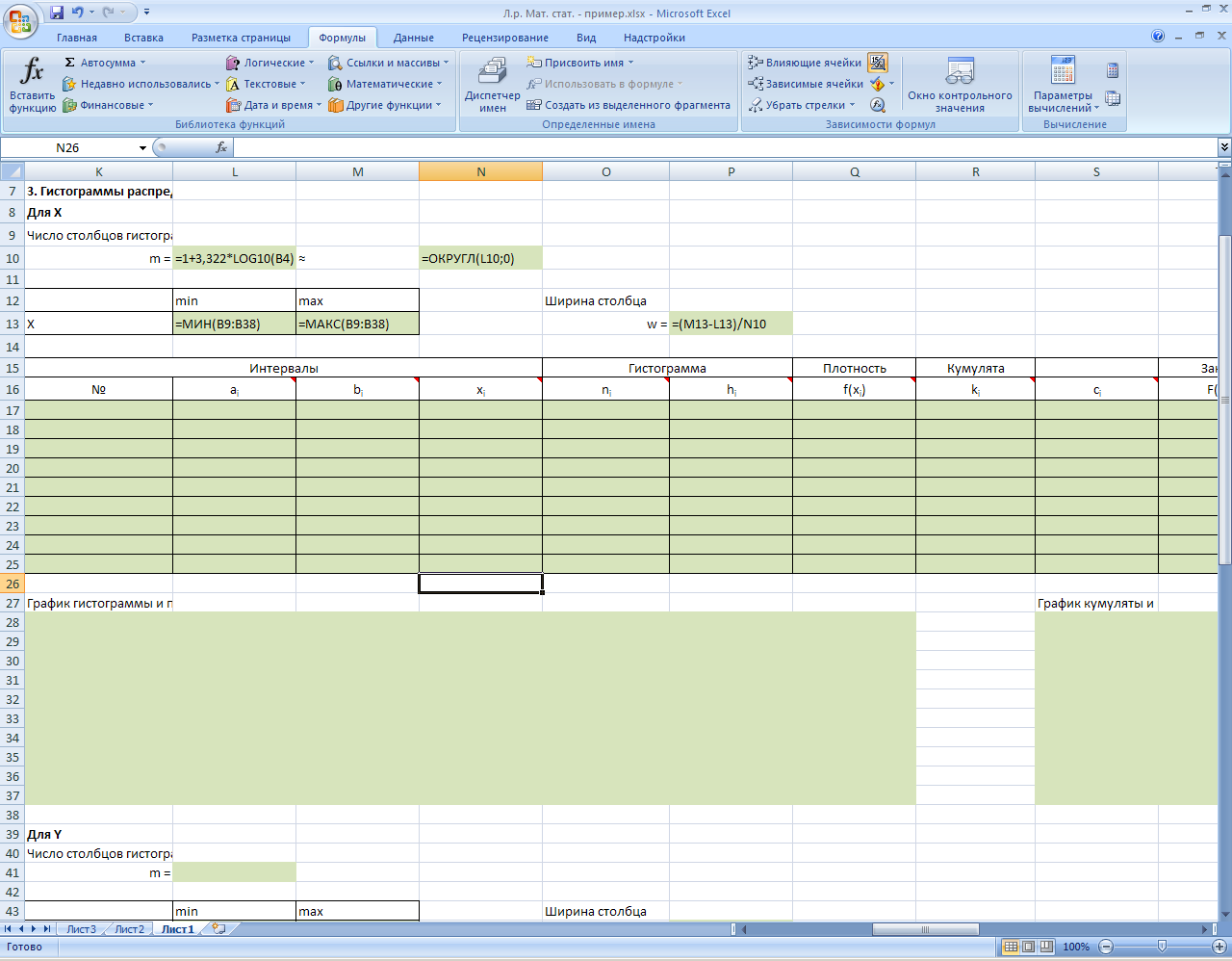
.

Полученное значение округляется до целых.

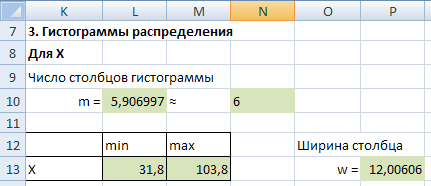
Тогда ширина каждого интервала:

.

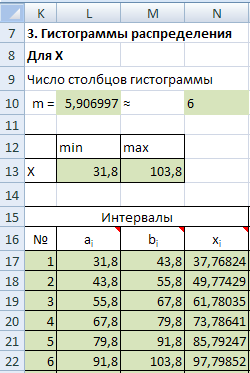
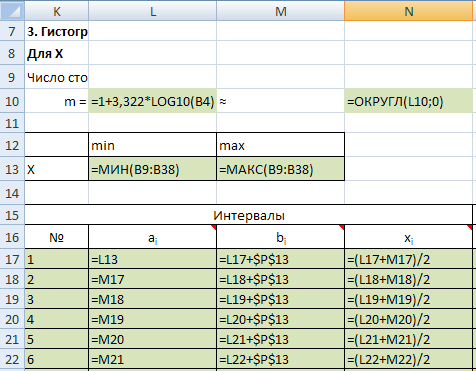
Выполним расчет для X:



Получим:



Теперь необходимо сформировать интервалы. Первый интервал начинается в минимальном значении X, последний заканчивается в максимальном. Сразу рассчитаем середины интервалов xi.



Далее необходимо подсчитать число значений X, попадающих в каждый интервал. Воспользуемся функцией СЧЁТЕСЛИМН. Условие попадания в интервал:

.

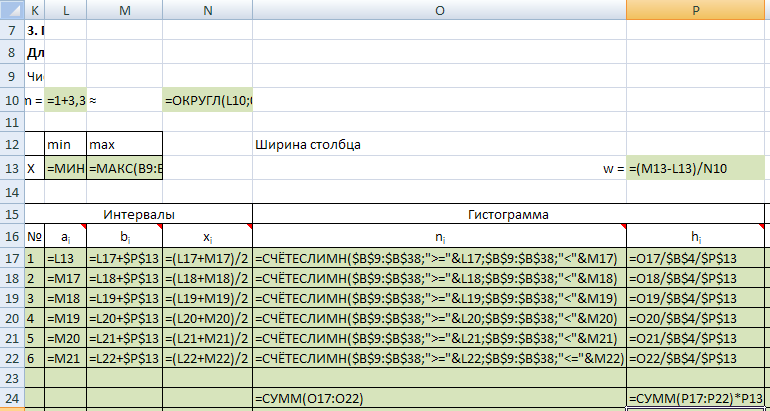
Проверьте себя:

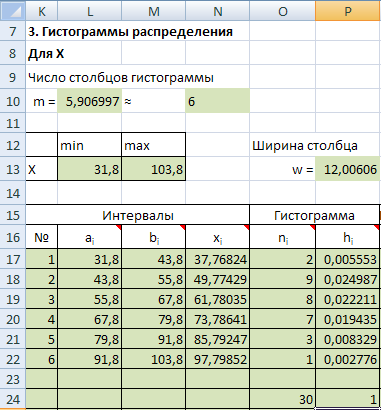
.

Полученные частоты ni пересчитываются в высоты столбцов гистограммы hi, так, чтобы выполнялось условие нормировки:

.

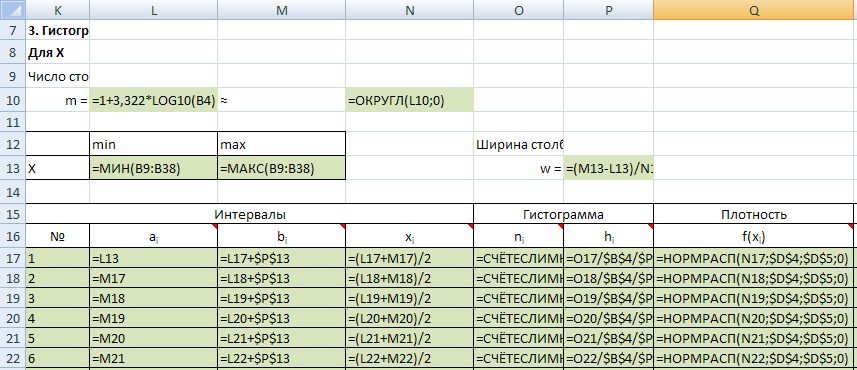
Обратите внимание, для последней строки формула СЧЁТЕСЛИМН немного отличается, иначе максимальное значение не будет учтено:

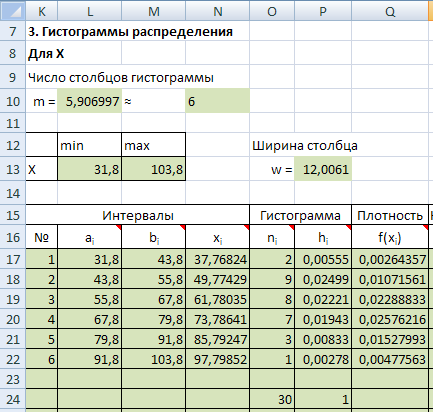




Условие нормировки выполнено.

Сравним полученную гистограмму с теоретическим законом. Для его вычисления воспользуемся формулой НОРМРАСП(xi;МО;СКО;0). 0 означает, что нам нужна плотность распределения, а не интегральный закон.





Сравним hi и f(xi) графике. Добавьте на лист столбчатую гистограмму и измените тип ряда 2 на «График с маркерами».

Видим, что расхождения достаточно заметные. Гистограмма несимметрична относительно центра и смещена влево.

Осталось построить кумулятивную гистограмму и сравнить его с интегральным законом распределения.

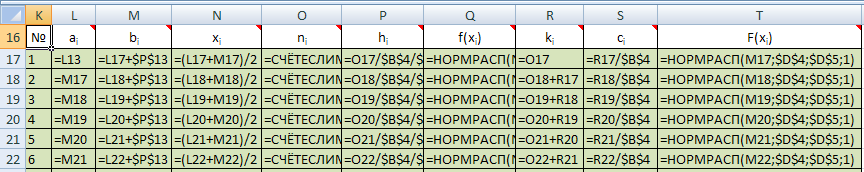
Каждый столбец кумулятивная гистограмма («с накоплением») показывает число попаданий ki в интервал [xmin; bi). Но строить ее удобнее как сумму попаданий в интервалы гистограммы плотности распределения.

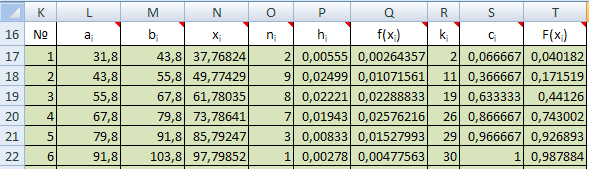
Высота столбца рассчитывается по формуле:



В последнем столбце должна получиться 1.

Для расчета интегральной функции также воспользуемся функцией НОРМРАСП, но рассчитывать будем не по серединам, а по концам интервалов bi.

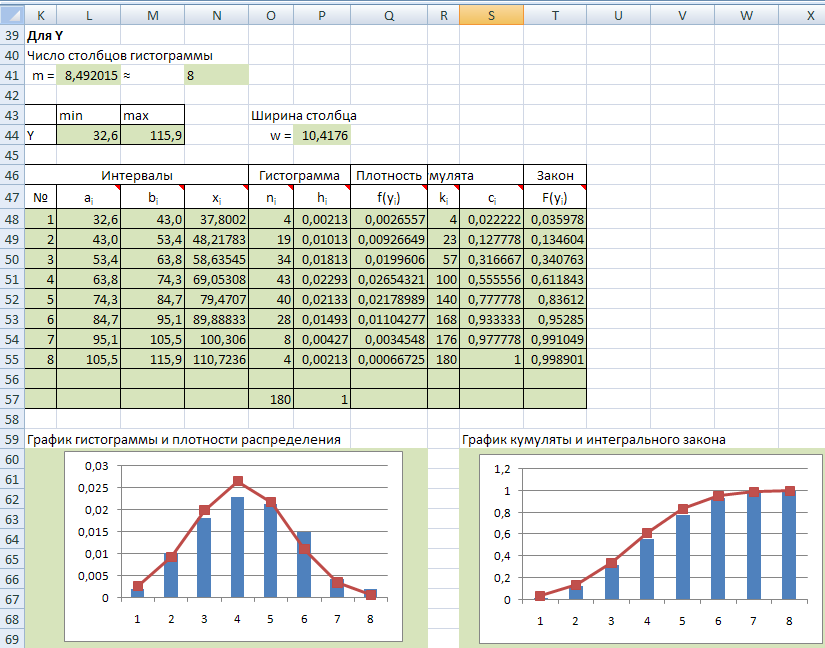




По аналогии построим график гистограммы:

На кумулятивном графике расхождения не так заметны, но также наблюдается смещение влево.

Полностью аналогично строятся гистограммы для Y.



1. **Доверительный интервал** математического ожидания.

Доверительный интервал – это диапазон значений, в который рассматриваемая числовая характеристика попадает с заданной вероятностью (90%, 95%, 99% и т.п.).

Доверительный интервал математического ожидания нормально распределенной СВ:

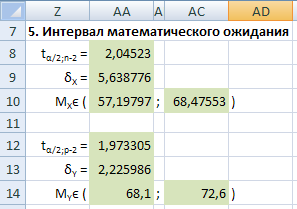
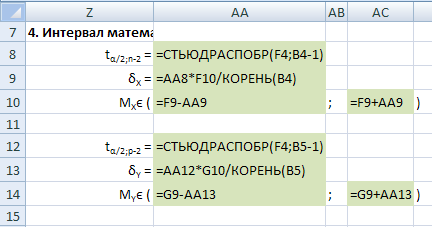
,

,

где  – обратная функция распределения Стьюдента с  степенями свободы,  – критический уровень.

Для вычисления  воспользуемся функцией СТЬЮДРАСПОБР(*α; n‑*1). Делить альфа пополам **не нужно**, Excel сделает это сам!

Затем рассчитаем δ по формуле и постоим границы интервалов.



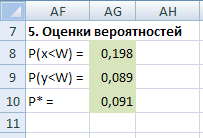
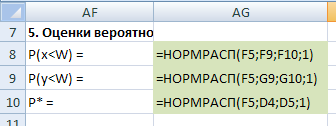
Видим, что при увеличении объема выборки ширина интервала δ уменьшается. При этом для X истинное значение МО не попадает в доверительный интервал с заданной вероятностью 0,95, т.е. выборочное среднее при малом числе генераций получается слишком грубым.

1. **Вероятность** получения того или иного значения СВ оценивается через интегральный закон распределения:

,

.

Вычислим *P*(*x*<*W*) и *P*(*y*<*W*) через выборочные *m* и *S* и сравним их с «точной» вероятностью *P*\*, рассчитанной через истинные МО и СКО (предполагается, что на практике истинные значения неизвестны).



Видим, что при малом числе генераций (X) вероятность получилась завышенной более чем в два раза. При большом числе генераций (Y) вероятность оказалась очень близкой к истинному значению.

1. Самостоятельно сделайте выводы по работе, ориентируясь на пояснения, приведенные в конце каждого пункта задания.

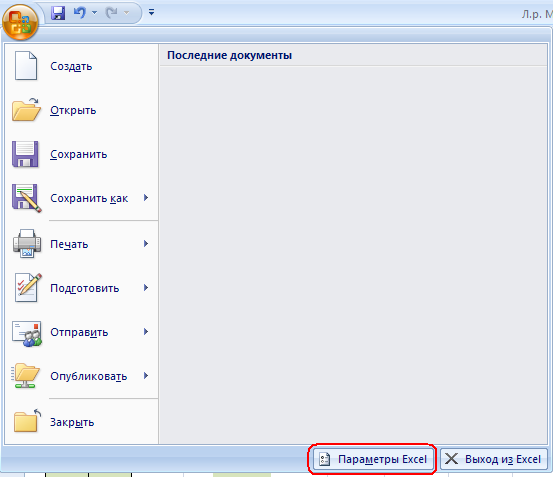
Почему возникают такие расхождения в результатах при разном числе генераций?

Как вы считаете, сколько генераций нужно для достаточно точного результата?

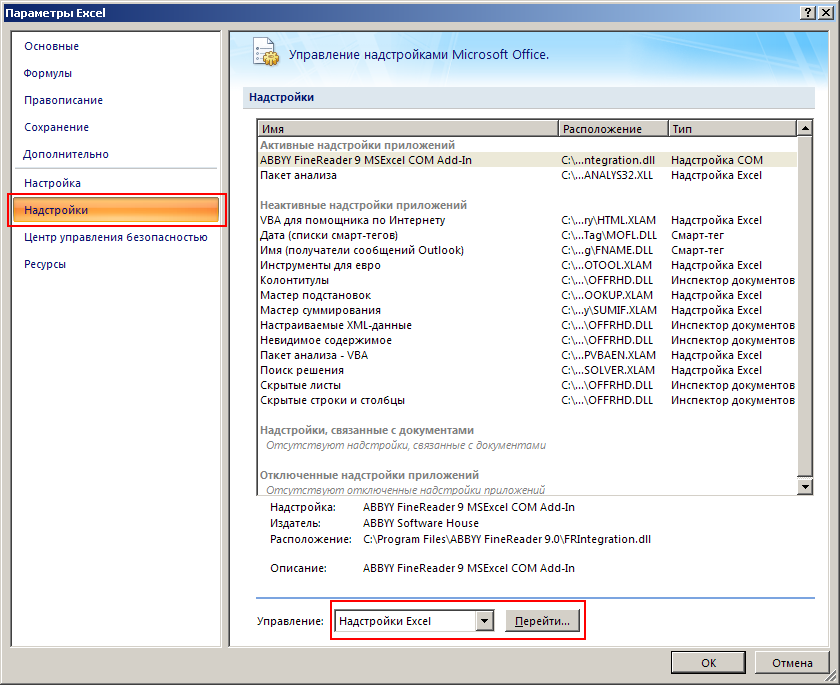
# Приложение

Как включить «Пакет анализа»

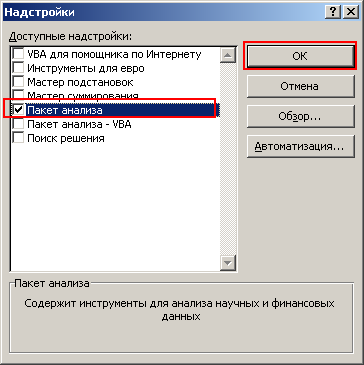
1. В меню Office выбрать пункт «Параметры Excel».



1. На вкладке «Надстройки» в выпадающем списке выбрать «Надстройки Excel» и нажать кнопку «Перейти…»



1. В открывшемся окне установить флажок «Пакет анализа» и нажать OK.



После этого кнопка «Анализ данных» появится на вкладке «Данные».

