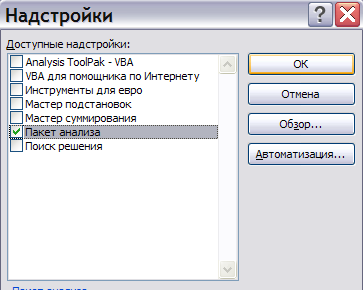
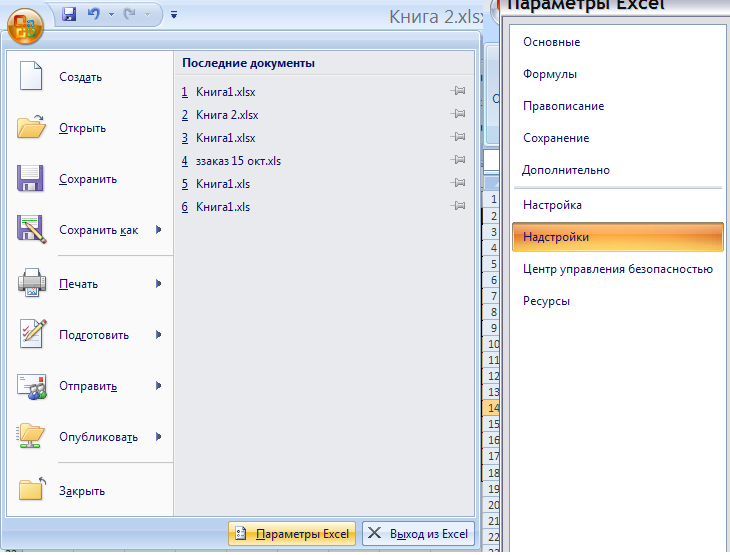
**Лаборатория работа 2**

При создании использовалась монография «**Тестирование учебных достижений»** **В.С. КИМ** <http://uss.dvfu.ru/static/kim_testing_monograph/index.html>

1. Скопируем нашу матрицу на новый лист, чтобы было удобнее работать
2. Вычислить ковариацию

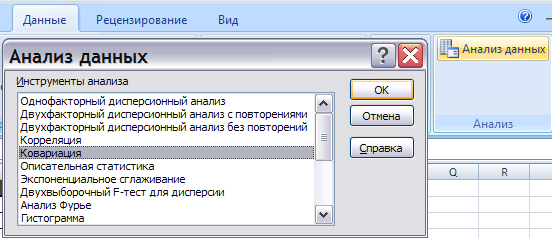
Ковариация (корреляционный момент) - в теории вероятностей и математической статистике мера линейной зависимости двух случайных величин.

Чтобы посчитать ковариацию и построить ковариационную матрицу в Excel необходимы предварительно установить пакет анализа. Для этого открываем **меню (кнопка Office)** и идем в **Параметры Excel**. Затем слева выбираем **Надстройки,** затем уже в открывшемся разделе внизу находим **Управление**: выбираем (если не выбрано)**Надстройки Excel** и нажимаем **Перейти**.

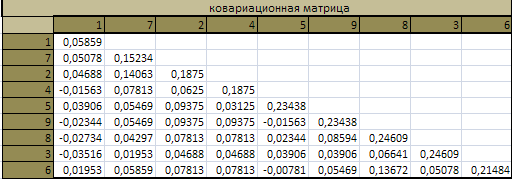
 

В открывшемся окне выберем **Пакет Анализа** и Окей. Начнется установка. Ждем и по окончании приступаем непосредственно к заданию.

Вкладка **Данные** в правом углу **Анализ Данных.** В появившемся меню **Ковариация.**



В появившемся меню Ковариация **Входной Интервал** метим собственно нашу матрицу, только без номеров строк и столбцов, **Параметры вывода**, для удобства выберем **новый рабочий лист**. Потом просто скопируем на наш лист. Просто можно выходной интервал, но тогда придется выделить нужный нам интервал, что не очень удобно. Вот что у нас должно получиться: матрица заполнилась не вся, так как относительно главной диагонали все симметрично. Можно теперь перебить названия и вместо «*Столбец»* написать номера, только номера нашей упорядоченной матрицы. Только необходимо понимать, что номера столбцов вверху должны совпадать с номерами строк новой ковариационной матрицы.

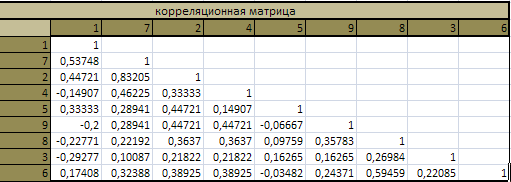


В Excel 2003 установка: **Сервис – Надстройки – Пакет Анализа – Ок**.

В Excel 2003 принцип построение абсолютно такой же, только **Анализ Данных** находится в меню **Сервис** (он появится там после установки).

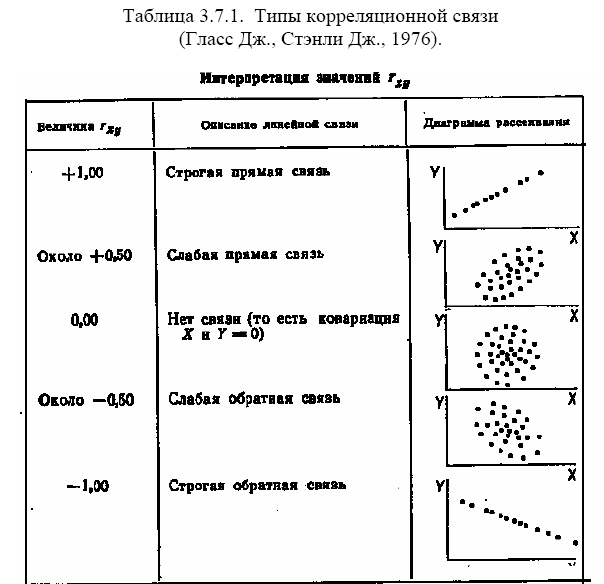
1. Вычислить коэффициенты корреляции Пирсона

**Данные** – **Анализ данных** – **Корреляция** (все, так же как и в предыдущем пункте строим матрицу, только вместо Ковариации выбираем Корреляция). Получается матрица, в которой главную диагональ занимают единицы. Матрица так же выведется не вся, так как она симметрична относительно главной диагонали.



1. Построить корреляционную матрицу. Сделать выводы.

Мы уже построили её. Просто в ней у нас стоят как раз коэффициенты корреляции Пирсона. Теперь надо посмотреть на типы связи, которые мы определим по значению корреляций. Будем оценивать по таблице. Понятно, что столбец с самим собой будет иметь прямую связь. Нам надо посмотреть на близкие к 1 значения. Это значит что данные задания хорошо коррелируют между собой. Для этого метим нашу корреляционную матрицу, выбираем **Условное форматирование** на вкладке **Главная – Правила выделения ячеек** (как выделяли нули в первой лабораторной). Для начала выделим, где связи хорошая – можно взять правило **Больше** и там указать способ заливки и значение скажем больше 0,6, раз у нас близкое к 0,5 по таблице – слабая связь. В итоге мы получим, что у нас хорошо коррелирует 2 задания с 7. Единицы по главной диагонали мы в расчет не берем. Затем слабая прямая связь близкое к 0,5. Действуем по такой же схеме. Можно так же использовать **«Между»**. Выделив ячейки для наглядности и для удобства, мы сделаем выводы и напишем их прямо в Excel напротив матрицы. Условно можно разделить интервалы, с учетом таблицы. Это разделение является условным.



[0,8…1] – сильная корреляция

[0,6…0,8) – близкая к сильной

[0,35…0,6) – близкая к слабой корреляции

[0…0,35) – близкая к нулевой связи (связи практически нет)

[0…-0,35) – близкая к нулевой обратной

[-0,35…-0,6) – слабая обратная связь

[-0,6…-0,8) – близкая к сильной обратной связи (связи практически нет)

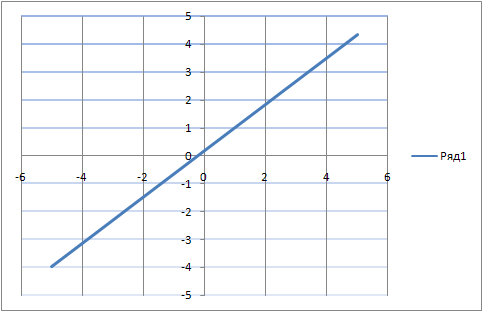
[-0,8…-1] – сильная обратная связь

1. Построить уравнение регрессии

Уравнение регрессии имеет вид у-уср = rXY(x-xcp). Где rXY – коэффициент корреляции между двумя заданиями (столбцами). Лучше выбрать наибольший из корреляционной матрицы, не равный единице. В нашем случае это 0,83205, который находится на пересечении 2 и 7 заданий. Значит yср – это среднее значение по заданию (по столбику) 7, а соответственно xcp – среднее по столбику 2. (учитываем нашу нумерацию). В итоге получаем наше уравнение: y-0,8125=0,83205(x-0,75). Затем построим график, это будет прямая.

Для этого в ячейки скажем (у нас они свободны) Q1:Q21 необходимо ввести значение аргумента при помощи автозаполнения, в данном случае с шагом 0,5. В ячейку R1 вводится значение функции, вычисляемое по формуле =0,83205\*(Q1-0,75)+0,8125. Ячейки R2:R21 заполняются копированием формулы из ячейки R1.

Далее выделим диапазон Q1:R21 и воспользуемся "Мастером диаграмм". Для построения графика функции лучше выбрать точечную диаграмму.



1. Вычислить коэффициент ϕ

Этот коэффициент совпадает с коэффициентом корреляции Пирсона, они дают одно и тоже значение. Результаты занесены в нашу матрицу.

1. Вычислить бисериальный коэффициент

бисериальный коэффициент корреляции Пирсона вычисляется по формуле:

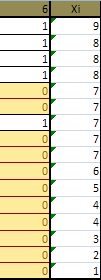
C:\Мои документы\Downloads\e9a7777cfcd0df959d8d6292b834ba41.png

где  - средний индивидуальный балл испытуемых, справившихся с данным заданием,  - средний индивидуальный балл испытуемых, не справившихся с данным заданием,

n1 – число испытуемых, выполнивших данное задание (Rj), n0 – число испытуемых, не выполнивших его (Wj). n =n1 + n0 – общее количество испытуемых (обще количество строк (только учтем, что после редуцирования)); Sx - стандартное отклонение для индивидуальных баллов всех испытуемых – корень из дисперсии (σ)

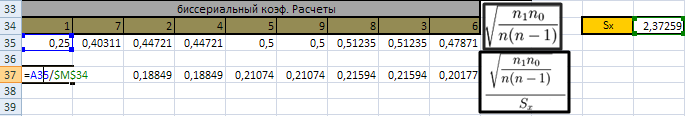
Например:

В нашем случае, для задания 6: n1 = 5, n0 = 11, n=16 (общее для всех).

S x = 2,37259 (одинаково для всех). - считается так: (9+8+8+8+7)/5.  = (7+7+7+7+6+5+4+4+3+2+1)/11. Если вы найдете в Excel способ быстро посчитать бисериальный коэффициент, значит вы молодец. Как можно облегчить себе расчеты: выражение под корнем, посчитать таким образом. Вбить в ячейку формулу , где В18 – содержит Rj, а B19 – Wj. 240 – это 16\*15 в нашем случае - всего 16 испытуемых. Протянем по горизонтали и получим значения корня. И считать вручную его не придется.

(В ячейках В18-J18 у нас содержатся количество верных ответов, а в B19-J19 – количество неверных ответов - мы считали их в предыдущей работе).

Затем ниже делим наш корень на стандартное отклонение. Получаем вот такие расчеты:



Средне - арифметическое придется считать вручную. Затем разность и потом умножение.

Так же можно создать массив, чтобы удобно было считать разность и умножение.

