

Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики

Кафедра программных систем

**Лабораторная работа 7**

**Построение и исследование моделей регрессионного анализа**

Выполнил: Гололобов

Владимир Владимирович

Группа № 3120

Проверила: Казанова П.П.

Санкт-Петербург

2020

**Цель работы:**

Изучить средства программы Microsoft Excel для регрессионного анализа данных

**Ход работы:**

1. Была создана таблица со столбцами “затраты на ТО” и “срок службы” и она была заполнена значениями. После чего по формуле =КОРРЕЛ(A2:A6;B2:B6) был подсчитан коэффициент корреляции между двумя массивами. В результате чего получилось значение очень близкое к единице, что значит – данные сильно связаны между собой

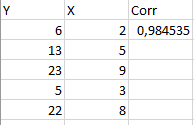


Рисунок 1 – Подсчет корреляции

1. Была создана таблица, показывающая распределение двух случайных величин, после чего был вычислен коэффициент корреляции по формуле =КОРРЕЛ(B8:F8;B9:F9). Он получился отрицательным, что значит, наши данные негативно связаны, то есть прямой с отрицательным коэффициентом, а небольшое значение означает, что данные плохо линейно связаны



Рисунок 2 – Подсчет корреляции

1. Была построена таблица данных из задания 3. И по ней была проведена линейная регрессия с помощью соответствующей функции из вкладки “анализ данных”. На вход ей было подана 2 входных интервала X и Y, а на выход ячейку, в которую необходимо записать значение

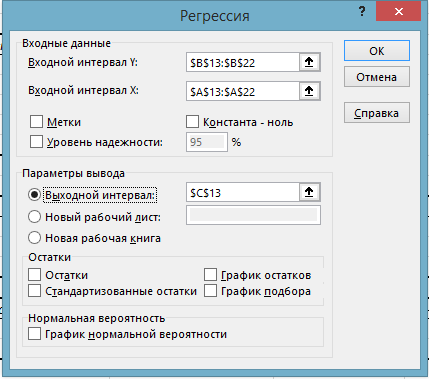


Рисунок 3 – Осуществление регрессии

1. В итоге получились следующие значения. Уравнение регрессии имеет вид y=41,1409x + 217,986. Значение коэффициента детерминации 0,89 показывает, что у нас получилась почти линейная зависимость. Значимость F=0,0005 меньше 0,05 и подтверждает значимость величины коэффициента детерминации. P-значение для количества работников равное 0,0005 и меньшее 0,05 подтверждает значимость коэффициента b1. P-значение коэффициента b0 равное 0,006 меньше 0,05. Это означает, что коэффициент b0 значимый

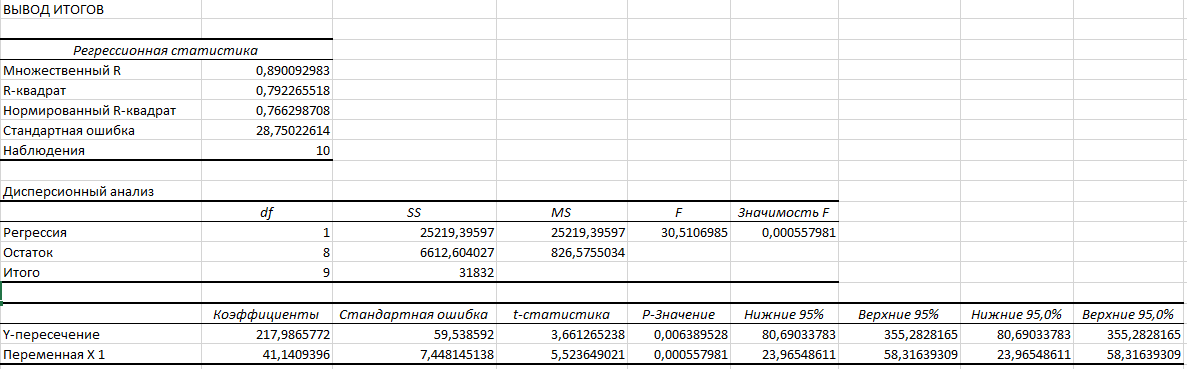


Рисунок 4 – Результаты регрессии

1. Была построена таблица данных, данная в задании 4. По ней была проведена линейная регрессия как и в предыдущем задании с отличием в том, что на входе было указано две зависимые переменные X1 и X2

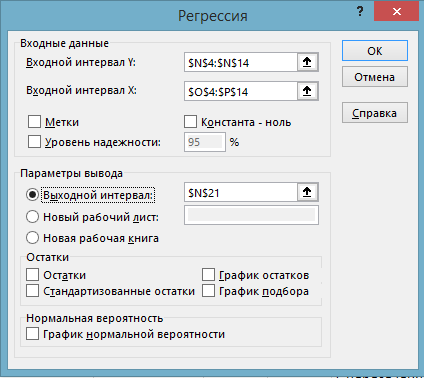


Рисунок 5 – Осуществление регрессии

1. В итоге получились следующие значения. Уравнение получилось y=348,762x1-5453,944x2+30189,61. Значение коэффициента детерминации получилось равных 0,84, что означает, что данные линейно зависимы. Рассчитанных множитель значимости F = 9,67 больший 0,05 означает, что величина детерминации незначима и следует выбрать другую метрику. Значение P для x1 равно 0,049 что меньше 0,05 означает значимость b1, однако значение P для x2 равно 0,087, что больше 0,05 следовательно коэффициент b2 незначим. Аналогично коэффициент b0 значим

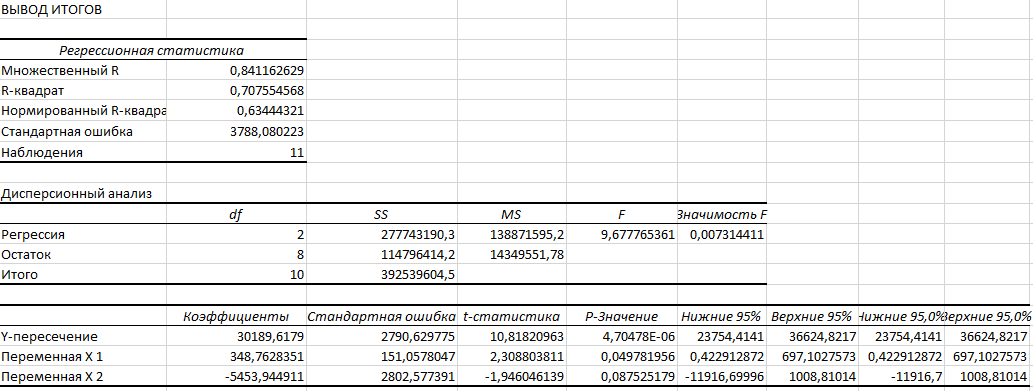


Рисунок 6 – Результаты регрессии

1. Была построена таблица данная в 5 задании, после чего для нее была произведена регрессия по методу, описанному в пункте выше. В результате чего получились следующие результаты. Уравнение получилось y=0,202x1+0,694x2-1,943. Коэффициент детерминации получился равным 0,99, что означает, что наша модель почти идеально линейна. Рассчитанный уровень значимости F получился равным 273,74, что больше 0,05, следовательно коэффициент детерминации незначимый. Значения P для x1 и x2 получились равными 0,038 и 0,01, что меньше 0,05, следовательно коэффициенты b1 и b2 значимые, а значение P для y получилось равным 0,81, что больше 0,05, а следовательно коэффициент b0 незначимый

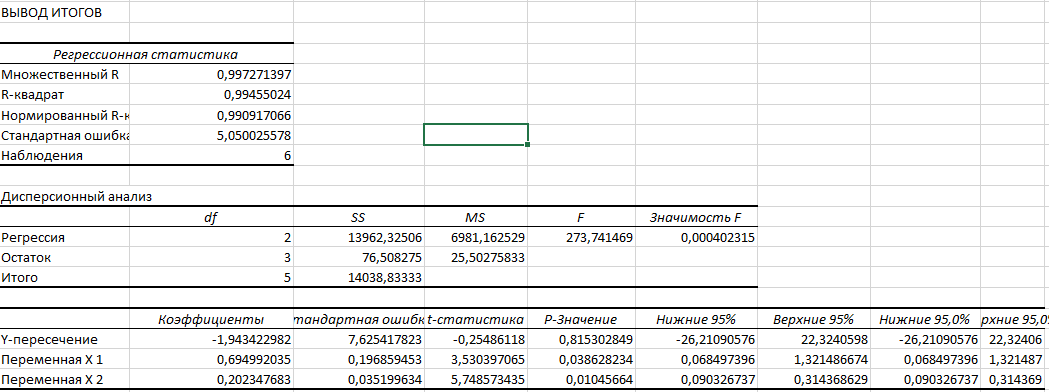


Рисунок 7 – Результаты регрессии

**Вывод:**

В ходе выполнения лабораторной работы были изучены средства Microsoft Excel для регрессивного анализа данных

**Ответы на контрольные вопросы:**

1. Что называется регрессивной моделью?

Регрессивная модель – это функция, описывающая зависимость между количественными характеристиками сложных систем

1. Привести общий вид регрессионной модели
2. Каким образом можно проверить значимость коэффициента регрессии?

Если соответствующее P значение меньше 0,05, то коэффициент значимый, иначе незначимый

1. Какой метод обычно используется при определении коэффициента регрессии?

Метод наименьших квадратов