Министерство образования и науки Российской Федерации

ФГБОУ ВО “Севастопольский государственный университет”

Кафедра ИС

Отчет

по лабораторной работе №1

«Корреляционный и регрессионный анализ данных. Множественный регрессионный анализ»

Выполнил:

ст. гр. ИTб-33д

Лисянский А. И.

Проверил:

Токарев А. И.

Севастополь

2015

**1.Цель работы**

Закрепить теоретические знания и приобрести практические навыки в построении регрессионных моделей объекта по экспериментальным данным с использованием программы Statistica.

По результатам наблюдений за функционированием объектов получены экспериментальные данные. Построить регрессионные модели объектов по заданным экспериментальным данным.

**2.Вариант задания**

Котировка юаня относительно других валют (доллар,евро,рубли)

**3.Ход работы**

Решение общей задачи разбивается на несколько этапов:

а) предварительная обработка данных с целью стандартизации результатов наблюдения;

б) оценка параметров регрессионных моделей;

в) проверка значимости коэффициентов регрессии;

г) оценка точности регрессионных моделей;

д) выводы о возможности применения составленных регрессионных моделей.

|  |  |  |  |
| --- | --- | --- | --- |
| **Месяц,год** | **Средняя цена за 10 юаней, руб.** | **Средняя цена за 10 юаней, USD** | **Средняя цена за 10 юаней, EUR** |
| Февраль 2015 | 104.856 | **1.5986** | **1.4090** |
| Январь 2015 | 104.936 | **0.1599** | **1.3254** |
| Декабрь 2014 | 87.7689 | **1.6123** | **1.3200** |
| Ноябрь 2014 | 75.6258 | **1.6277** | **1.3182** |
| Октябрь 2014 | 66.5377 | **1.6351** | **1.3005** |
| Сентябрь 2014 | 61.8445 | **1.6281** | **1.2650** |
| Август 2014 | 58.6739 | **1.6273** | **1.2348** |
| Июль 2014 | 55.4986 | **1.6188** | **1.2009** |
| Июнь 2014 | 55.224 | **1.6117** | **1.1875** |
| Май 2014 | 55.6978 | **1.6015** | **1.1770** |
| Апрель 2014 | 57.24 | **1.5984** | **1.1743** |
| Март 2014 | 58.8931 | **1.6078** | **1.1909** |
| Февраль 2014 | 58.1104 | **1.6314** | **1.2229** |
| Январь 2014 | 55.5971 | **1.6513** | **1.2207** |
| Декабрь 2013 | 54.1276 | **1.6497** | **1.2079** |
| Ноябрь 2013 | 53.6563 | **1.6410** | **1.2282** |
| Октябрь 2013 | 52.6294 | **1.6415** | **1.2129** |
| Сентябрь 2013 | 53.1277 | **1.6342** | **1.2446** |
| Август 2013 | 53.8939 | **1.6340** | **1.2321** |
| Июль 2013 | 53.2016 | **1.6307** | **1.2742** |
| Июнь 2013 | 52.6012 | **1.6282** | **1.2437** |
| Май 2013 | 51.0672 | **1.6310** | **1.2697** |
| Апрель 2013 | 50.6989 | **1.6220** | **1.2598** |
| Март 2013 | 49.5148 | **1.6104** | **1.2577** |
| Февраль 2013 | 48.3038 | **1.6058** | **1.2299** |
| Январь 2013 | 48.6386 | **1.6076** | **1.2308** |
| Декабрь 2012 | 49.3672 | **1.6047** | **1.2479** |
| Ноябрь 2012 | 50.4282 | **1.6055** | **1.2569** |
| Октябрь 2012 | 49.5774 | **1.6021** | **1.2388** |
| Сентябрь 2012 | 49.6242 | **1.5901** | **1.2563** |
| Август 2012 | 50.1365 | **1.5745** | **1.2865** |
| Июль 2012 | 51.1143 | **1.5677** | **1.2948** |
| Июнь 2012 | 51.5889 | **1.5736** | **1.2073** |
| Май 2012 | 48.9664 | **1.5735** | **1.2481** |
| Апрель 2012 | 46.7351 | **1.5865** | **1.2053** |
| Март 2012 | 46.4739 | **1.5876** | **1.2116** |
| Февраль 2012 | 47.5031 | **1.5870** | **1.2190** |
| Январь 2012 | 49.8254 | **1.5815** | **1.2525** |
| Декабрь 2011 | 49.5917 | **1.5874** | **1.2112** |
| Ноябрь 2011 | 48.563 | **1.5675** | **1.1715** |
| Октябрь 2011 | 49.0938 | **1.5736** | **1.1878** |
| Сентябрь 2011 | 47.6146 | **1.5630** | **1.1543** |
| Август 2011 | 45.1815 | **1.5673** | **1.1021** |
| Июль 2011 | 43.1337 | **1.5523** | **1.1100** |
| Июнь 2011 | 43.1844 | **1.5468** | **1.0926** |
| Май 2011 | 42.9857 | **1.5424** | **1.0973** |
| Апрель 2011 | 43.071 | **1.5405** | **1.0758** |
| Март 2011 | 43.2916 | **1.5246** | **1.1014** |
| Февраль 2011 | 44.5068 | **1.5221** | **1.1198** |
| Январь 2011 | 45.4414 | **1.5140** | **1.1296** |
| Декабрь 2010 | 46.1868 | **1.5149** | **1.1344** |
| Ноябрь 2010 | 46.5468 | **1.5004** | **1.1150** |
| Октябрь 2010 | 45.5883 | **1.4979** | **1.09081** |
| Сентябрь 2010 | 45.5777 | **1.4958** | **1.1646** |
| Август 2010 | 44.8138 | **1.4690** | **1.1638** |
| Июль 2010 | 45.1091 | **1.4764** | **1.1756** |
| Июнь 2010 | 45.6126 | **1.4716** | **1.2271** |
| Май 2010 | 44.6423 | **1.4644** | **1.2014** |
| Апрель 2010 | 42.7844 | **1.4649** | **1.1098** |
| Март 2010 | 43.2406 | **1.4650** | **1.0987** |
| Февраль 2010 | 44.0874 | **1.4649** | **1.0547** |
| Январь 2010 | 43.532 | **1.4648** | **1.0173** |

Таблица 1- Котировка юаня за 5 лет относительно других валют

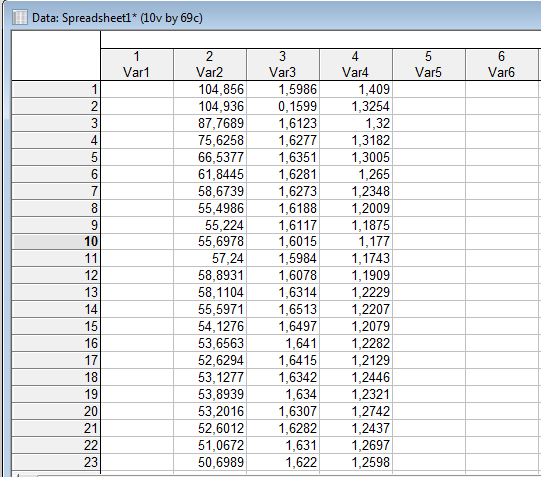


Рисунок 1 Создание электронной таблицы с исходными данными

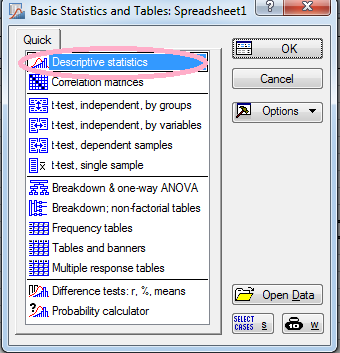


Рисунок 2 - Для получения дескриптивных статистик в стартовой панели модуля **Основные статистики и таблицы** выберем команду Descriptive statistics.

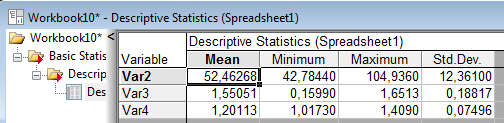


Рисунок 3 - Дескриптивная статистика

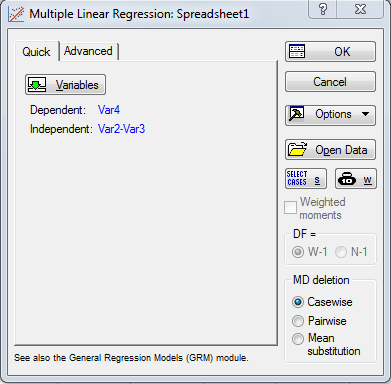


Рисунок 4 – При построении была выявлена одна зависимая переменная Var 4 и две независимых переменных Var 2 и Var 3

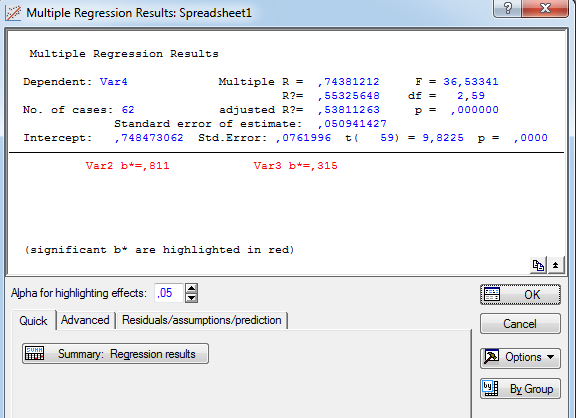


Рисунок 5 - Окно результатов анализа

Рассмотрим вначале информационную часть окна. В ней содержится краткая информация о проведенном анализе:

• *Dep. Var. —* Имя зависимой переменной*.*  — **VAR4***.*

• *No. of Cases* — Число случаев, по которым построена регрессия*.* В примере число равно **62**.

• *Multiple R —* Коэффициент множественной корреляции (эта статистика полезна во множественной регрессии, когда вы хотите описать зависимости между переменны­ми) – характеризует тесноту связи зависимой переменной с совокупностью независимых переменных: **0,74381212**

• *R2 —* Коэффициент детерминации - является одной из основных статистик в данном окне, он показывает долю общего разброса (относительно выборочного среднего зависимой переменной), которая объясняется построенной регрессией: **0,55325648**

•Adjusted *R2**—* Скорректированный коэффициент детерминации (применяется при небольших объемах выборочных данных, контролирует число степеней свободы): **0,53811263**

• *Std. Error of estimate* — Стандартная ошибка оценки. Эта статистика является ме­рой рассеяивания наблюдаемых значений относительно регрессионной прямой: **0,50941427**

• *Intercept* — Оценка свободного члена регрессии. Значение коэффициента b0 в уравнении регрессии: **0,748473062**

•*Std. Error* — Стандартная ошибка оценки свободного члена. Стандартная ошибка коэффициента b0 в уравнении регрессии: **0,761996**

• *t(df) and p-value* — Значение t-критерия и уровня р. t-критерий используется для проверки гипотезы о равенстве 0 свободного члена регрессии.**t=9,8225**

•*F* — расчетный критерий Фишера: **36,53341**

*•df*— Число степеней свободы F критерия: **2,59**

•*р* — Уровень значимости уравнения множественной регрессии (вероятность, что Fрасч < Fтабл, если p > α, то H0 принимаем): **0,0000**

В информационной части значение коэффициента множественной корреляции: **R = 0,74381212**. Коэффициент детерминации: **R2 = 0,53811263**, это означает, что доля влияния независимых переменных на зависимую значительна (53%). Гипотезу о том, что коэффициенты при независимых переменных равны нулю, мы отвергаем, т.к. **р = 0,0000 < α = 0,05**. Т.е. множественная регрессия высоко значима.

Стандартная ошибка оценки равна : 0,50941427. Стандартная ошибка свободного члена = 0,761996. Здесь же оценивается значимость свободного члена: р = 0,0000 < α = 0,05, это означает, что нулевая гипотеза о незначимости свободного члена отвергается.

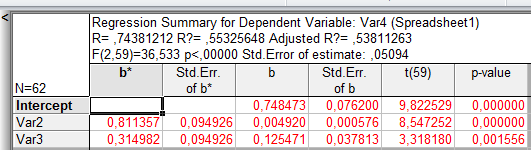


Рисунок 6 Краткие результаты регрессии для исходных данных

* 1. Оценка статистической надежности уравнения регрессии в целом с помощью F-критерия Фишера

Известно, что значение F-критерия Фишера может быть рассчитано по формуле:

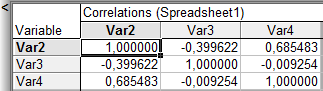
 ,

Где у нас  = -0,009 (для зависимости второго столбца от 3-ей строки(зависимость котировки юаня от доллара)).

Тогда, если n=62, а k=3, то

F = (0,000081/0,999919) \*19,3 = 0,0016

Если мы сравним полученное значение с табличным, то получим, что , а следовательно, уравнение регрессии в целом статистически значимо. Fтабл=0,001



**Выводы**

В ходе выполнения данной лабораторной работы были закреплены теоретические знания и приобретены практические навыки в построении регрессионных моделей объекта по экспериментальным данным с использованием программы Statistica.

Для этого были выполнены следующие действия:

А) С помощью команды Open Data был осуществлен ввод исходных данных в систему STATISTICA

Б) Были получены дескриптивные статистики по каждому из 8ми признаков.

В) Было составлено уравнение множественной регрессии

Г) Была составлена матрица парных коэффициентов корреляции

Д) С помощью F-критерия Фишера была оценена статистическая надежность уравнения регрессии в целом

Таким образом, были выполнены все цели и задачи данного лабораторного исследования.